US 20140135526A1

(19) **United States**

(12) **Patent Application Publication**

Lynch

(10) **Pub. No.: US 2014/0135526 A1**

(43) **Pub. Date: May 15, 2014**

(54) **COMPOSITIONS AND METHODS FOR 3-HYDROXYPROPIONATE BIO-PRODUCTION FROM BIOMASS**

(71) Applicant: **OPX Biotechnologies, Inc.**, Boulder, CO (US)

(72) Inventor: **Michael D. Lynch**, Boulder, CO (US)

(73) Assignee: **OPX Biotechnologies, Inc.**, Boulder, CO (US)

(21) Appl. No.: **14/067,838**

(22) Filed: **Oct. 30, 2013**

Related U.S. Application Data

(63) Continuation of application No. 13/284,337, filed on Oct. 28, 2011, now Pat. No. 8,652,816, which is a continuation of application No. 12/328,588, filed on Dec. 4, 2008, now Pat. No. 8,048,624.

(60) Provisional application No. 60/992,290, filed on Dec. 4, 2007.

Publication Classification

(51) **Int. Cl.**

C12P 7/40 (2006.01)

C07C 57/04 (2006.01)

(52) **U.S. Cl.**

CPC .. ***C12P 7/40*** (2013.01); ***C07C 57/04*** (2013.01)

USPC **562/598**; 435/136

(57) **ABSTRACT**

Methods of obtaining mutant nucleic acid sequences that demonstrate elevated oxaloacetate a-decarboxylase activity are provided. Compositions, such as genetically modified microorganisms that comprise such mutant nucleic acid sequences, are described, as are methods to obtain the same.

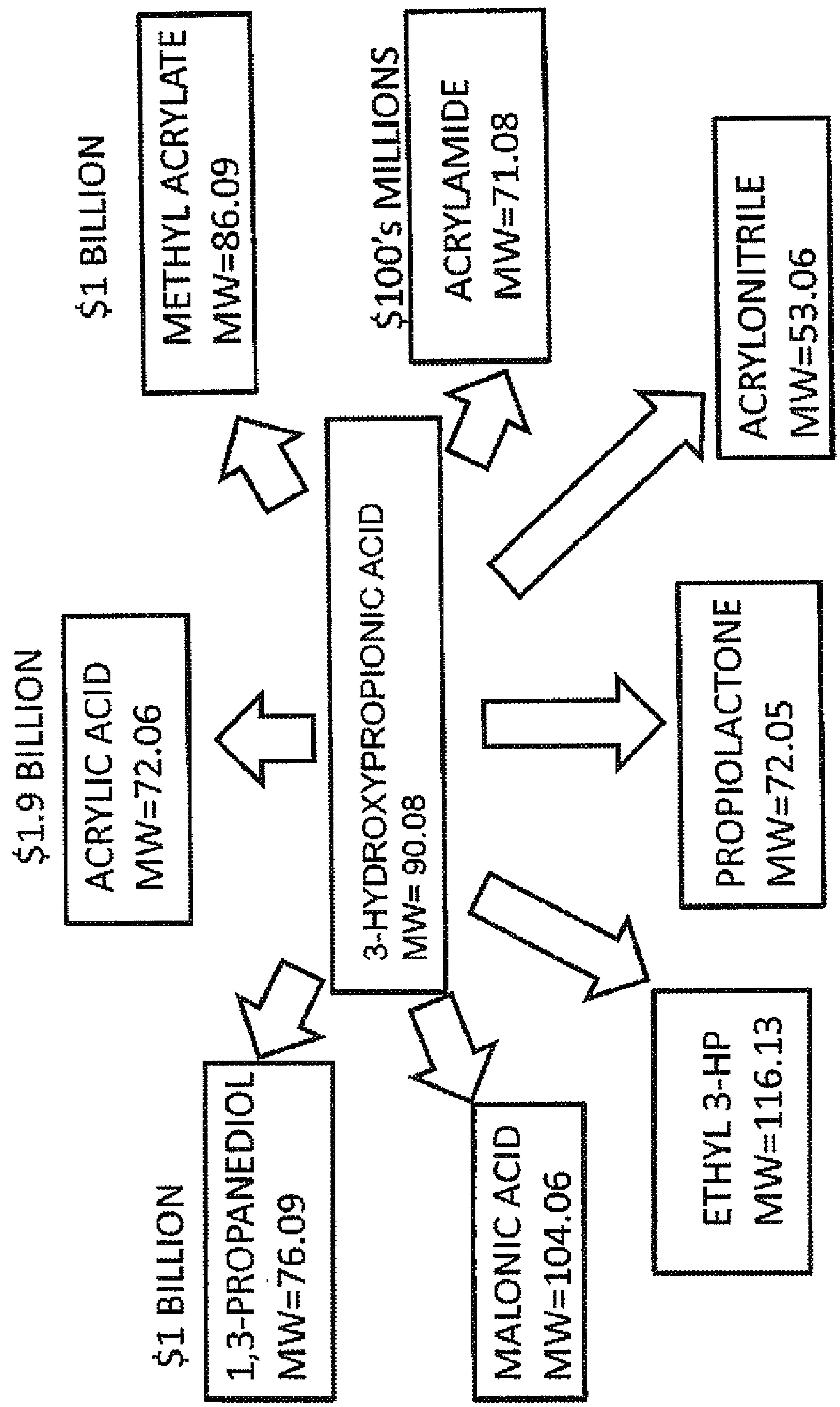

3-HYDROXYPROPIONIC ACID
MW= 90.08
$1 BILLION
METHYL ACRYLATE
MW=86.09
$100's MILLIONS
ACRYLAMIDE
MW=71.08
ACRYLONITRILE
MW=53.06
$1.9 BILLION
ACRYLIC ACID
MW=72.06
PROPIOLACTONE
MW=72.05
$1 BILLION
1,3-PROPANEDIOL
MW=76.09
MALONIC ACID
MW=104.06
ETHYL 3-HP
MW=116.13

FIG. 1

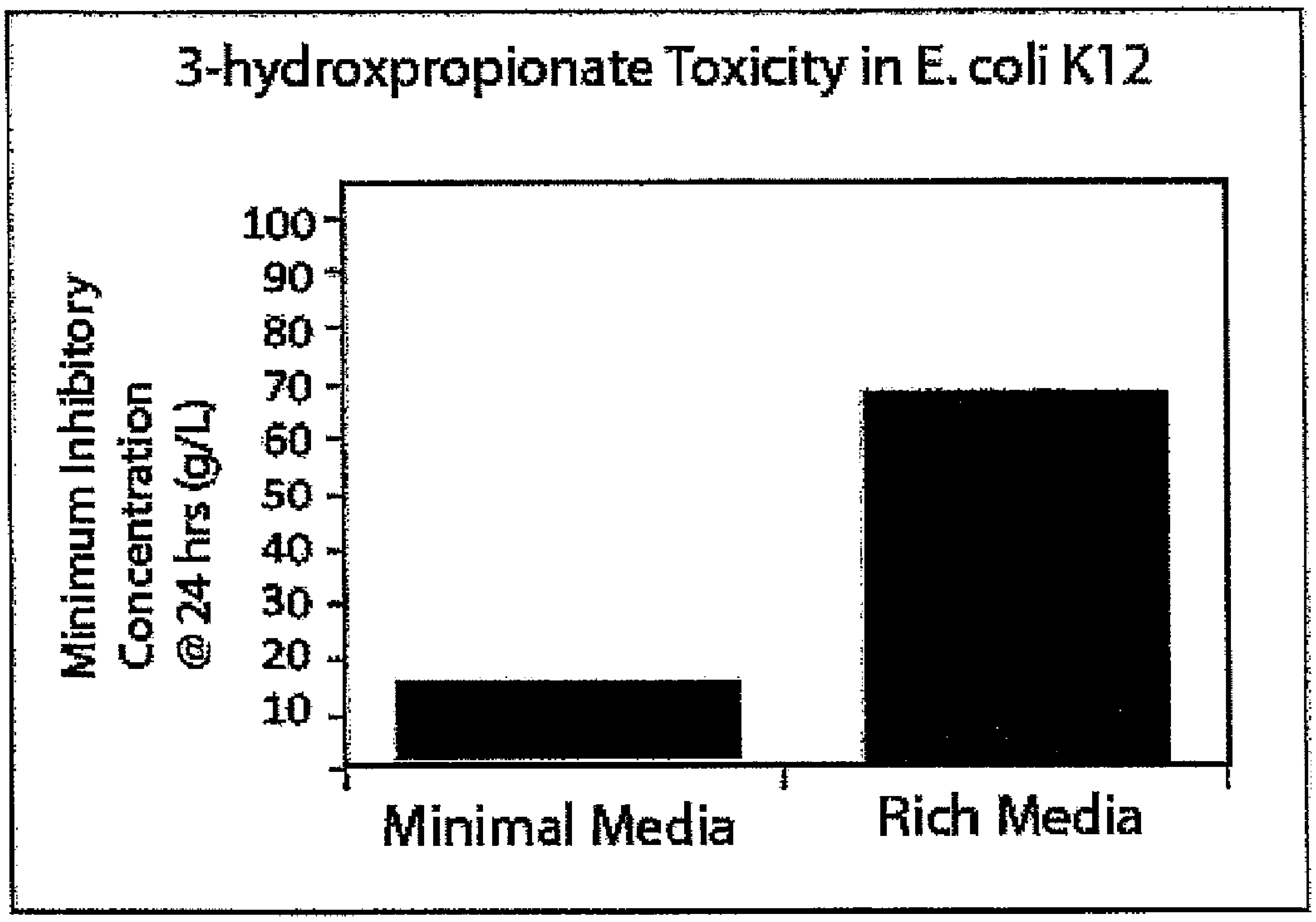

3-hydroxpropionate Toxicity in E. coli K12
Minimum Inhibitory Concentration @ 24 hrs (g/L)
100
90
80
70
60
50
40
30
20
10
Minimal Media
Rich Media

FIG 2

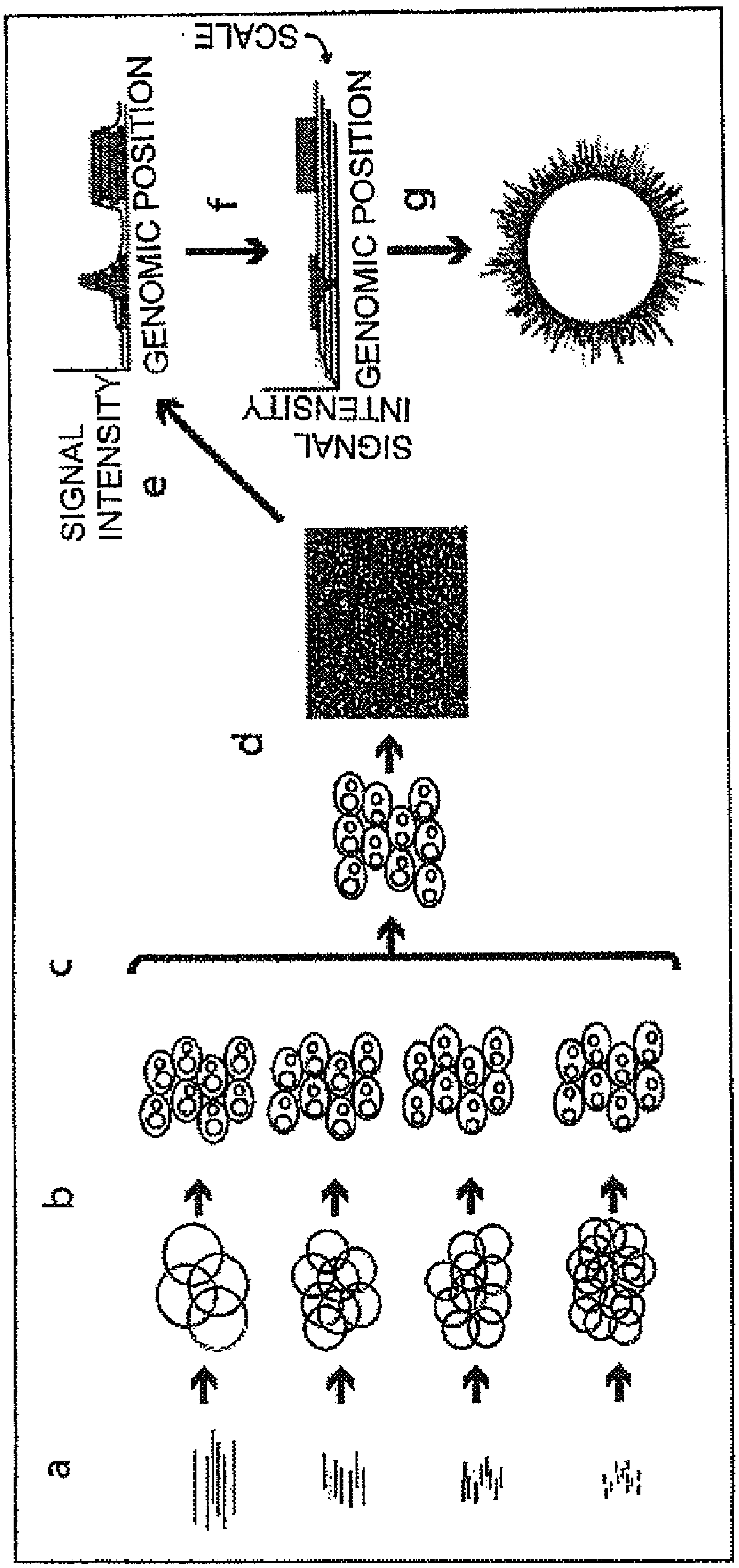

a
b
c
d
e
f
g
SIGNAL INTENSITY
GENOMIC POSITION
SIGNAL INTENSITY
GENOMIC POSITION
SCALE

FIG. 3

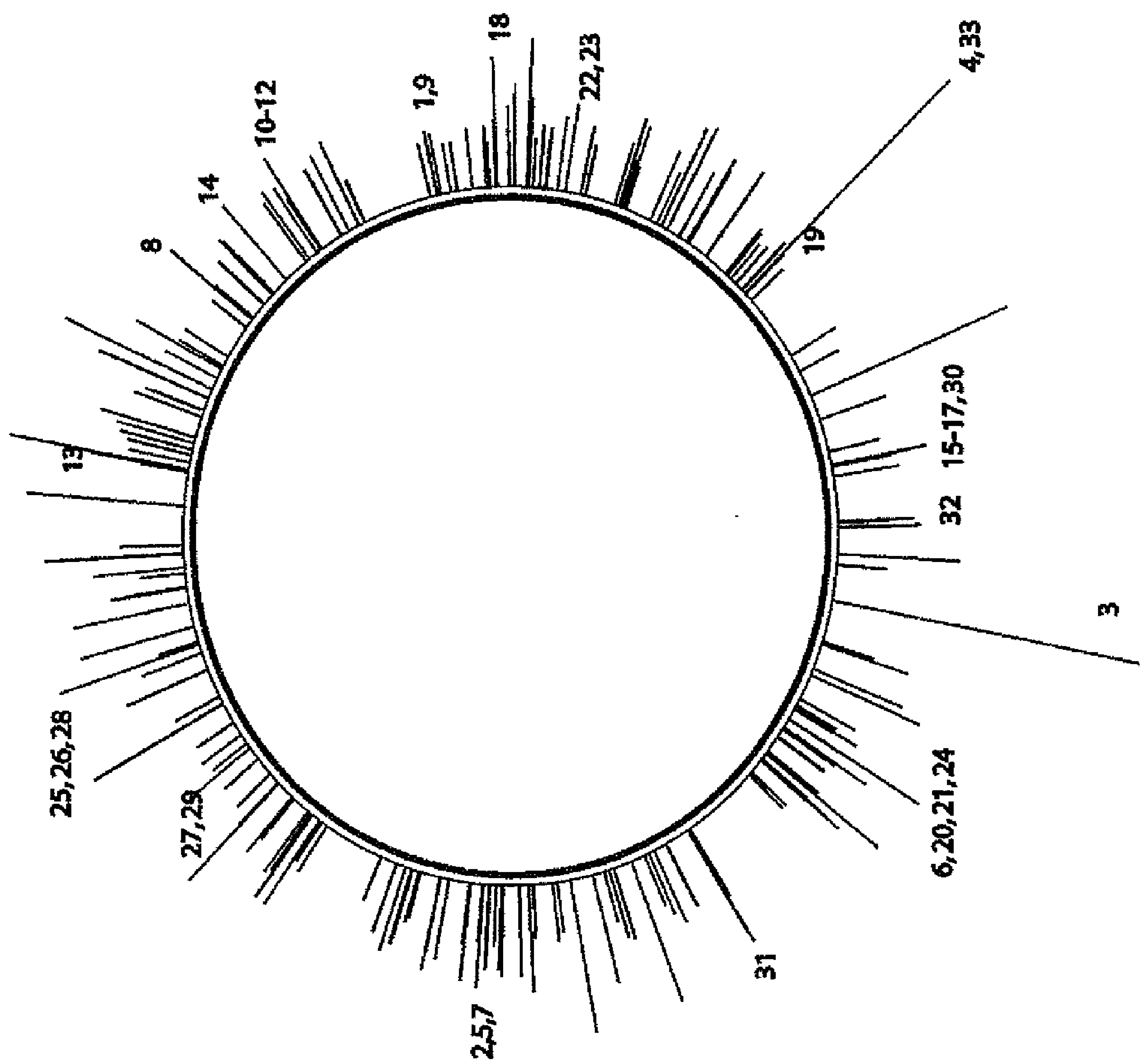

18
22,23
1,9
10-12
14
8
4,33
19
15-17,30
32
3
13
25,26,28
27,29
6,20,21,24
31
2,5,7

FIG. 4A

FIG. 4B

NUMBER	GENE
1	aroA
2	aroB
3	aroC
4	aroD
5	aroE
6	aroF
7	aroK
8	aroL
9	aspC
10	entC
11	entD
12	entE
13	folA
14	folD
15	menD
16	menE
17	menF
18	pabA
19	pabB
20	pheA
21	purN
22	trpA
23	trpD
24	tyrA
25	tyrB
26	ubiA
27	ubiB
28	ubiC
29	ubiE
30	ubiG
31	ubiH
32	ubiX
33	ydiB

FIG. 4C-1

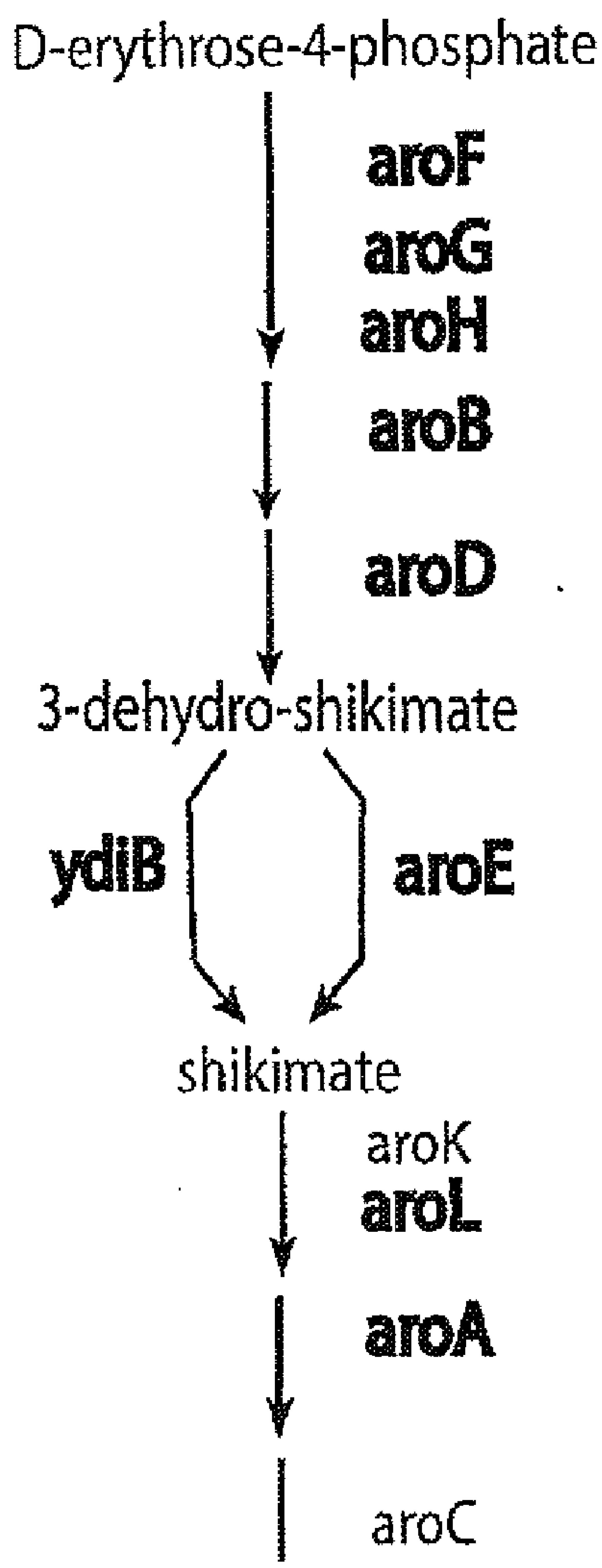
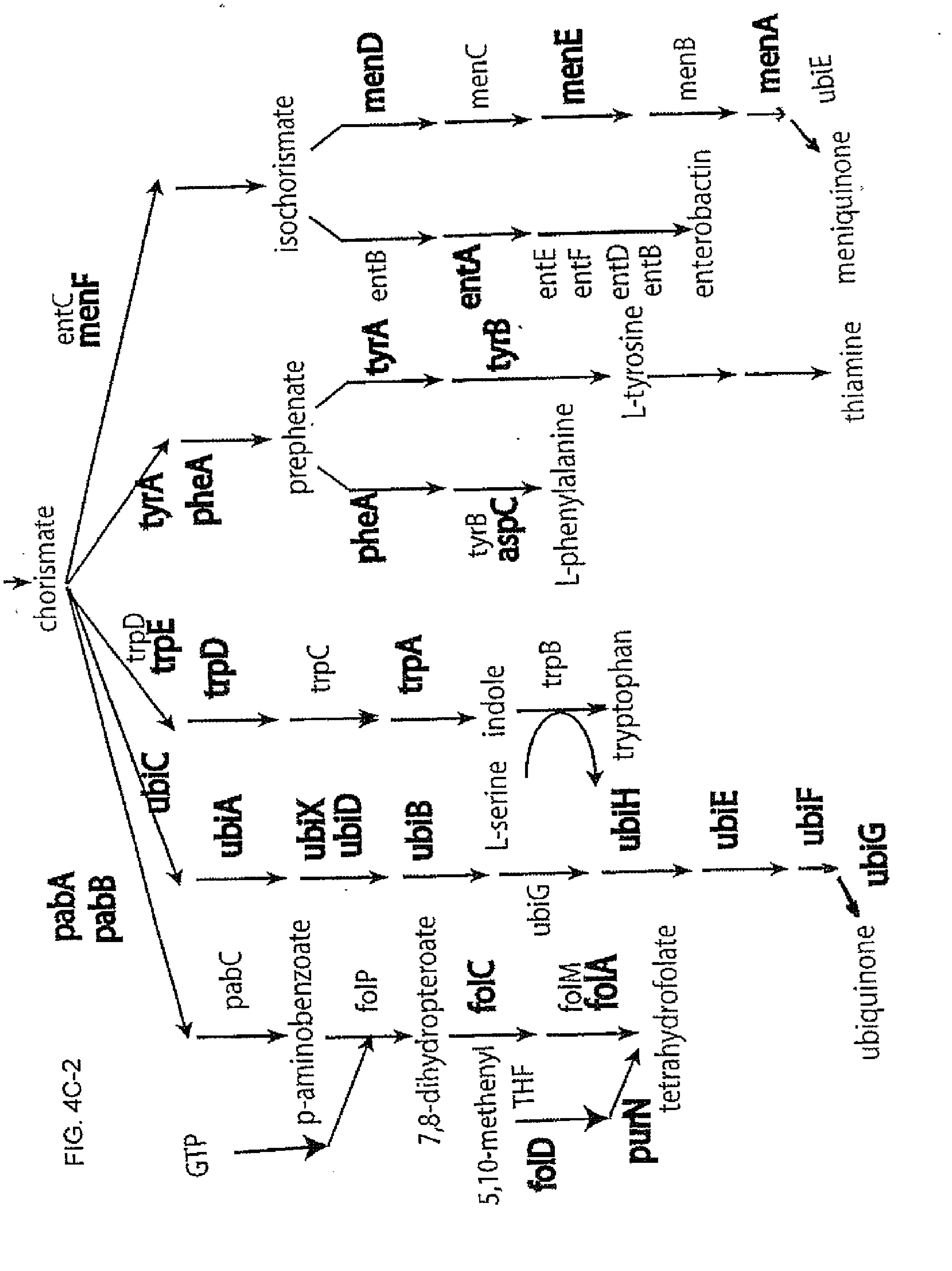

D-erythrose-4-phosphate
aroF
aroG
aroH
aroB
aroD
3-dehydro-shikimate
ydiB
aroE
shikimate
aroK
aroL
aroA
aroC

FIG. 4C-2


chorismate
pabA
pabB
pabC
p-aminobenzoate
GTP
folP
7,8-dihydropteroate
folC
folM
folA
tetrahydrofolate
5,10-methenyl THF
folD
purN
ubiC
ubiA
ubiX
ubiD
ubiB
ubiG
L-serine
ubiH
ubiE
ubiF
ubiG
ubiquinone
trpD
trpE
trpD
trpC
trpA
indole
trpB
tryptophan
tyrA
pheA
prephenate
pheA
tyrB
aspC
L-phenylalanine
tyrA
tyrB
L-tyrosine
thiamine
entC
menF
isochorismate
entB
entA
entE
entF
entD
entB
enterobactin
menD
menC
menE
menB
menA
ubiE
meniquinone

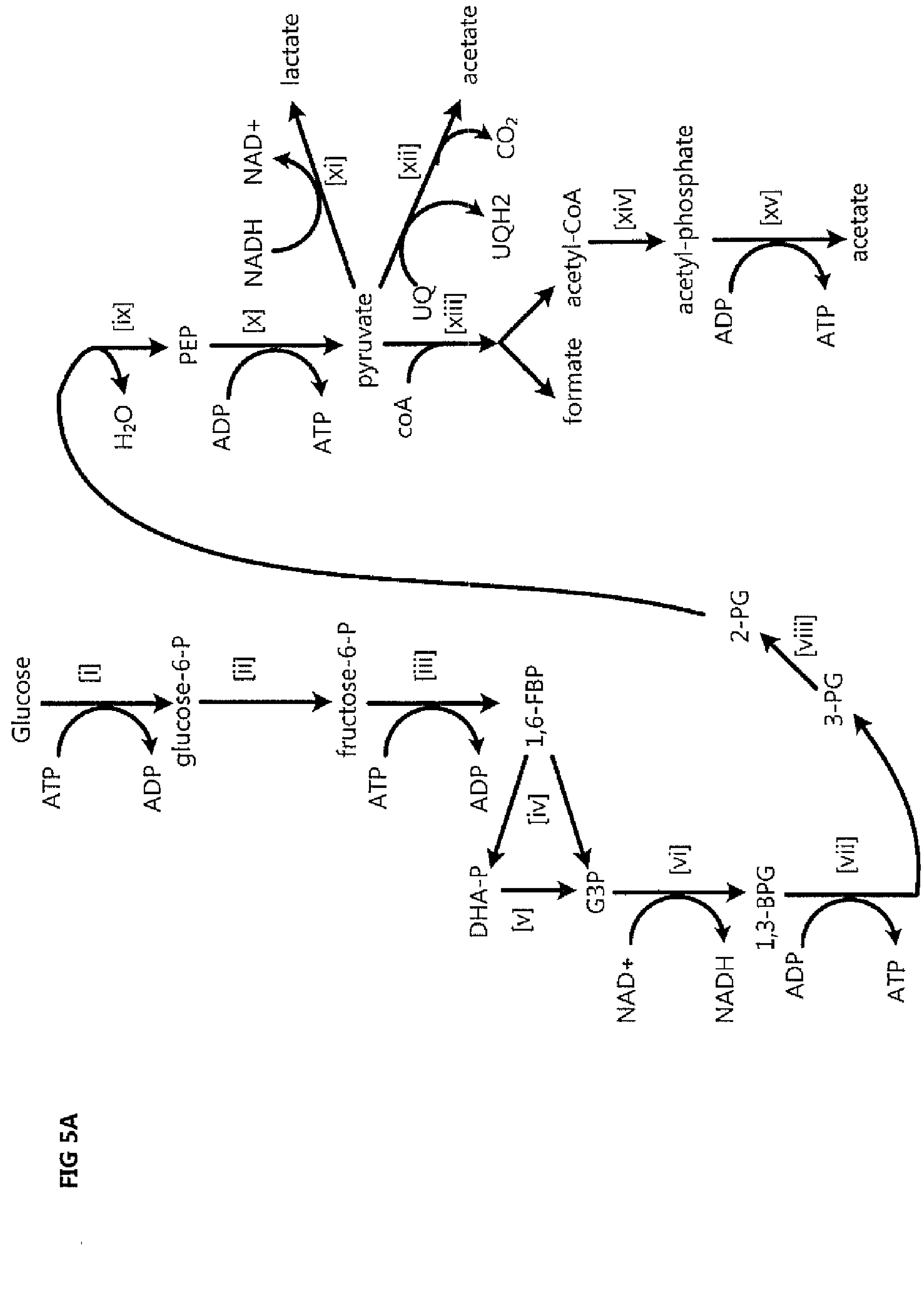
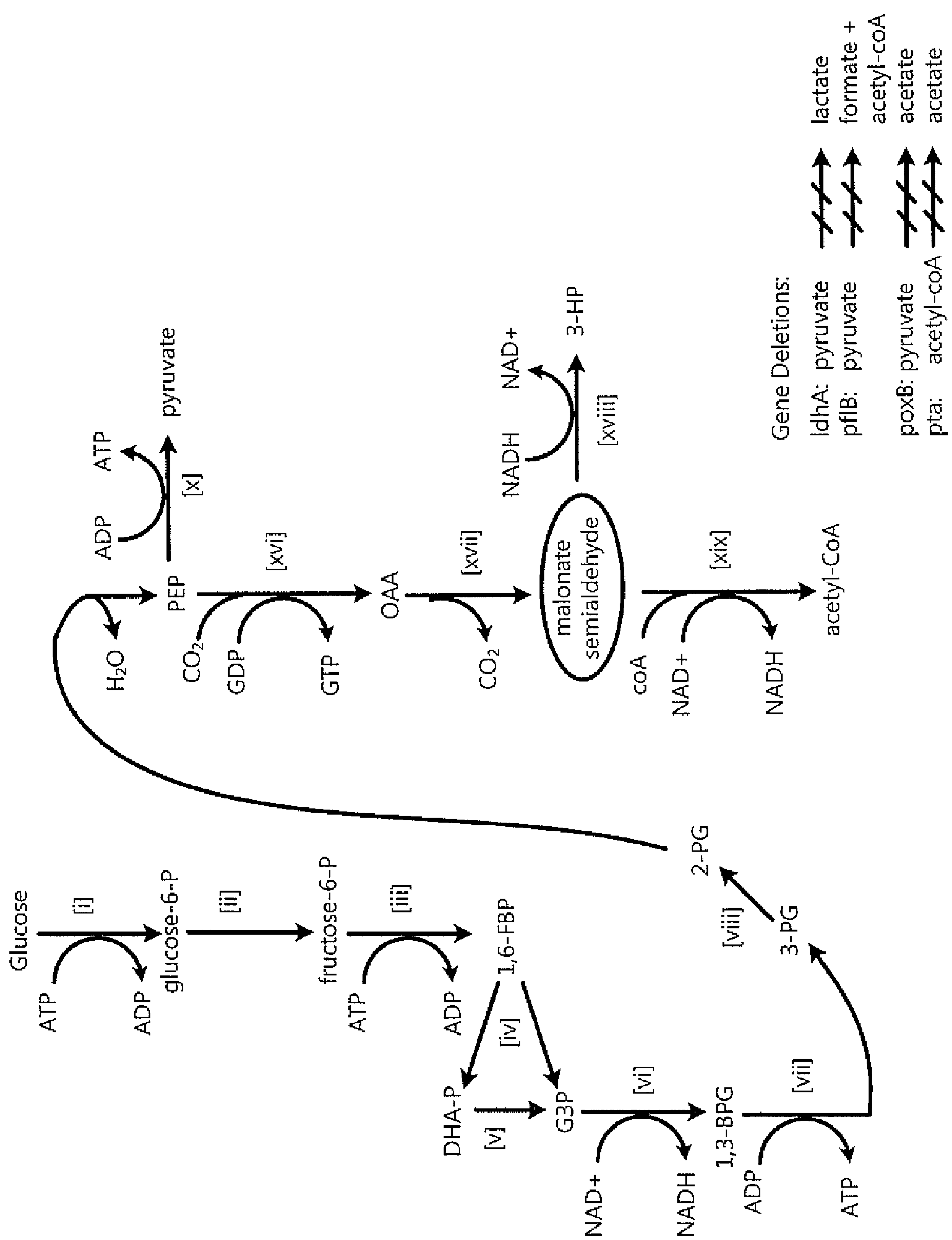

Glucose
ATP
ADP
[i]
glucose-6-P
[ii]
fructose-6-P
ATP
ADP
[iii]
1,6-FBP
[iv]
DHA-P
[v]
G3P
[vi]
NAD+
NADH
1,3-BPG
ADP
ATP
[vii]
3-PG
[viii]
2-PG
[ix]
H_2O
PEP
[x]
ADP
ATP
pyruvate
NADH
NAD+
[xi]
lactate
UQ
UQH2
[xii]
CO_2
acetate
coA
[xiii]
formate
acetyl-CoA
[xiv]
acetyl-phosphate
ADP
ATP
[xv]
acetate

FIG 5A

FIG 5B


Glucose
ATP
ADP
[i]
glucose-6-P
[ii]
fructose-6-P
[iii]
1,6-FBP
[iv]
DHA-P
[v]
G3P
[vi]
NAD+
NADH
1,3-BPG
[vii]
3-PG
[viii]
2-PG
H_2O
PEP
[x]
pyruvate
CO_2
GDP
GTP
[xvi]
OAA
[xvii]
malonate semialdehyde
NADH
NAD+
[xviii]
3-HP
coA
[xix]
acetyl-CoA
Gene Deletions:
ldhA: pyruvate → lactate
pflB: pyruvate → formate + acetyl-coA
poxB: pyruvate → acetate
pta: acetyl-coA → acetate

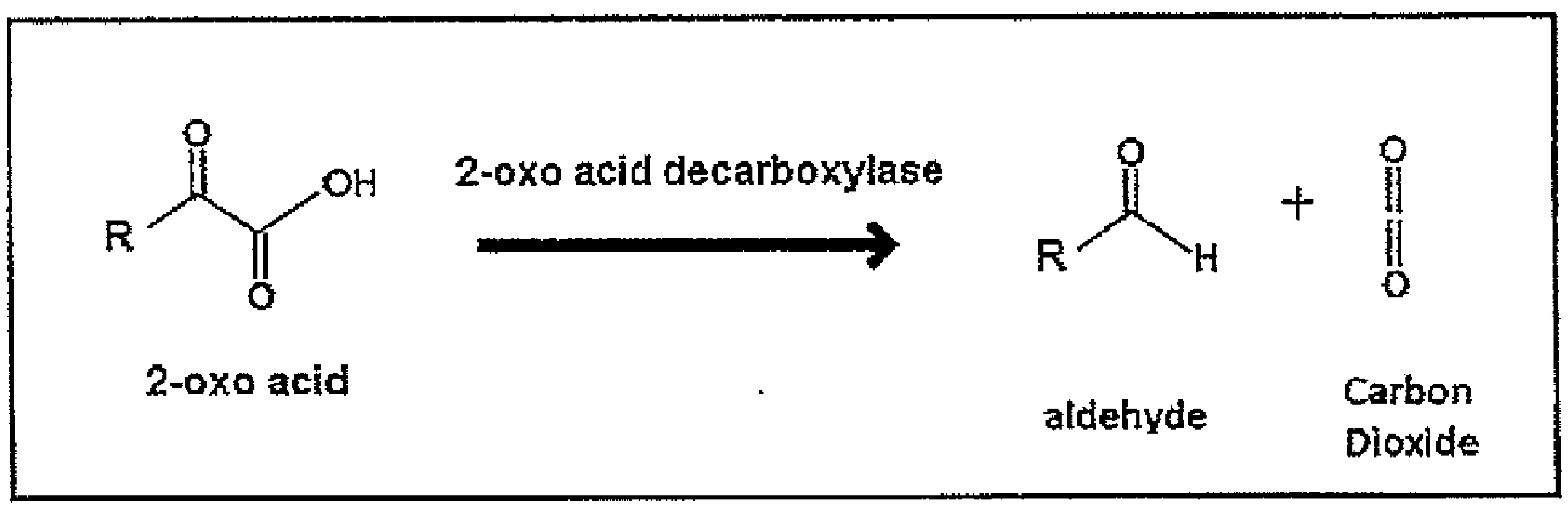

2-oxo acid decarboxylase
2-oxo acid
aldehyde
Carbon
Dioxide

FIG 6

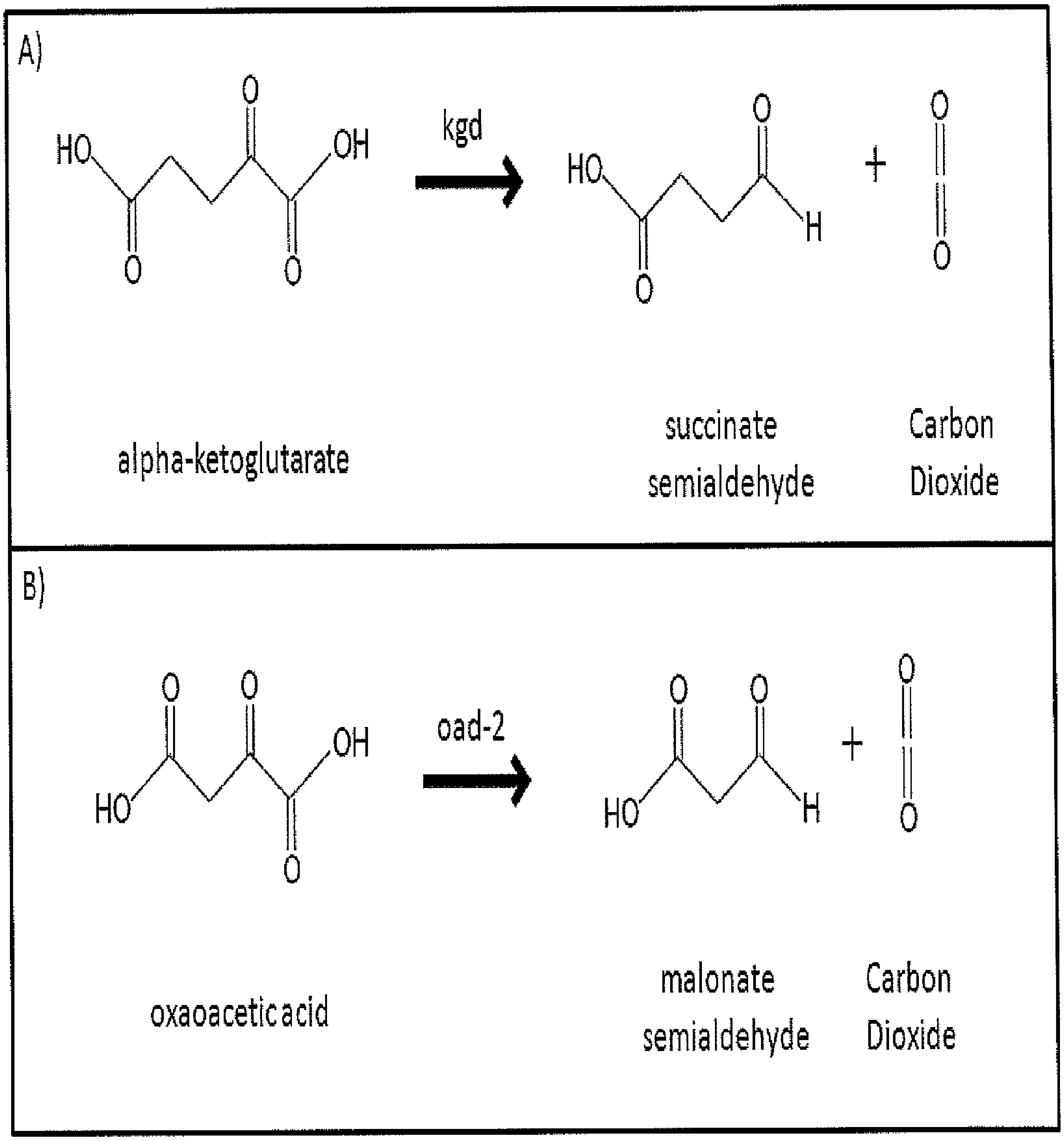

A)
HO
O
OH
O
O
kgd
HO
O
H
O
+
O
O
alpha-ketoglutarate
succinate
semialdehyde
Carbon
Dioxide
B)
O
O
OH
HO
O
oad-2
O
O
HO
H
+
O
O
oxaoaceticacid
malonate
semialdehyde
Carbon
Dioxide

FIG. 7A and 7B

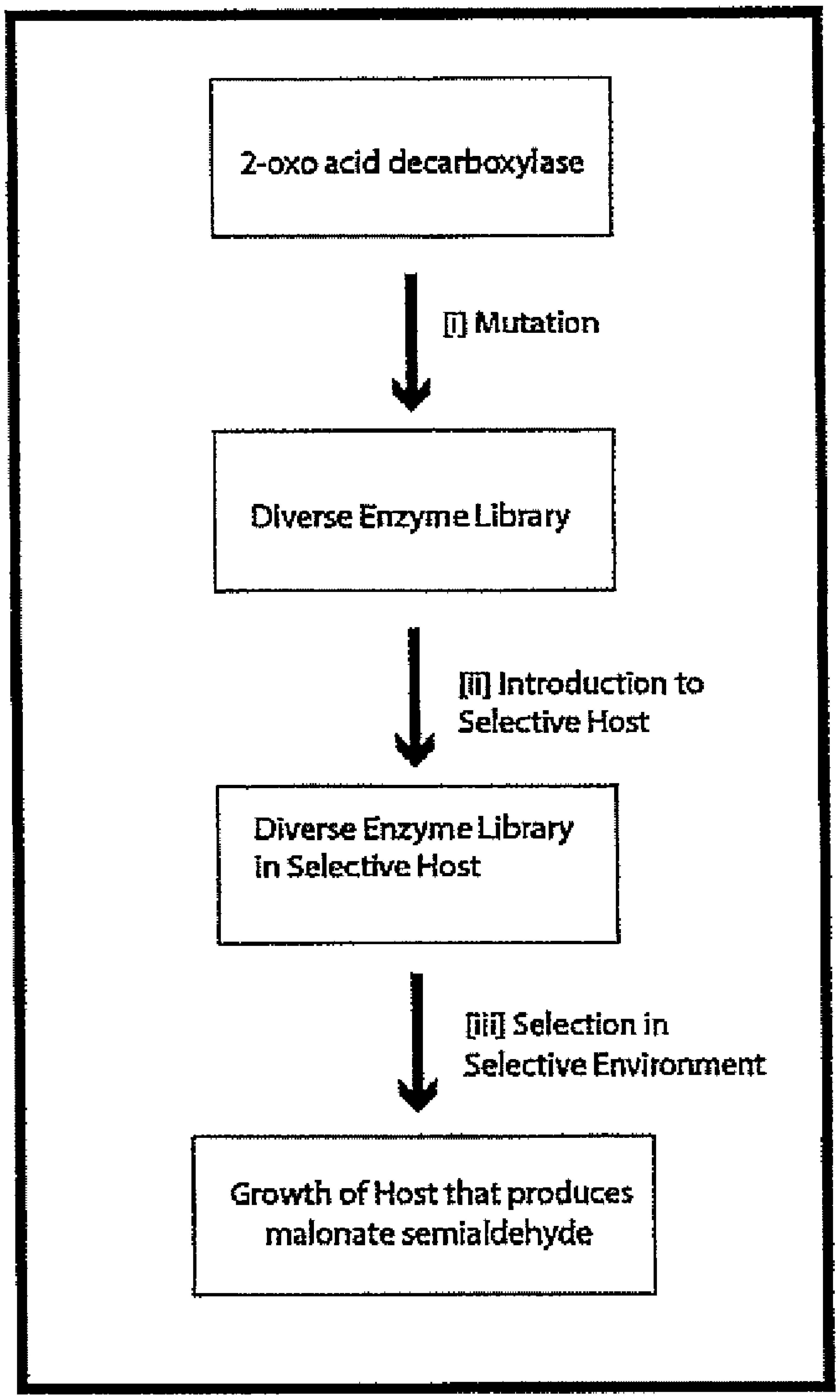

2-oxo acid decarboxylase
[i] Mutation
Diverse Enzyme Library
[ii] Introduction to Selective Host
Diverse Enzyme Library in Selective Host
[iii] Selection in Selective Environment
Growth of Host that produces malonate semialdehyde

FIG 8

FIG 9A

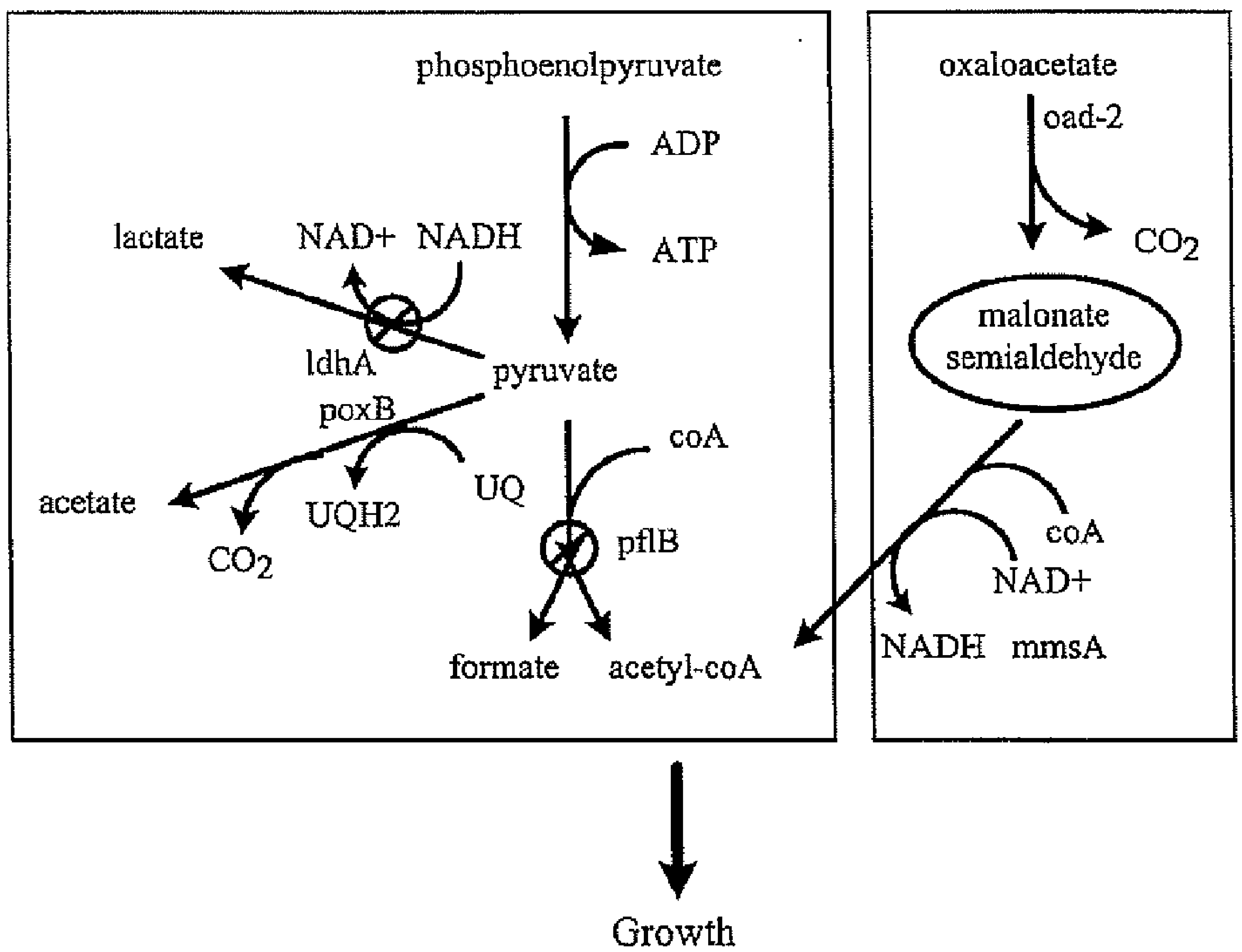

phosphoenolpyruvate
ADP
ATP
lactate
NAD+
NADH
ldhA
pyruvate
poxB
acetate
UQH2
UQ
CO2
coA
pflB
formate
acetyl-coA
oxaloacetate
oad-2
CO2
malonate semialdehyde
coA
NAD+
NADH
mmsA
Growth

FIG 9B

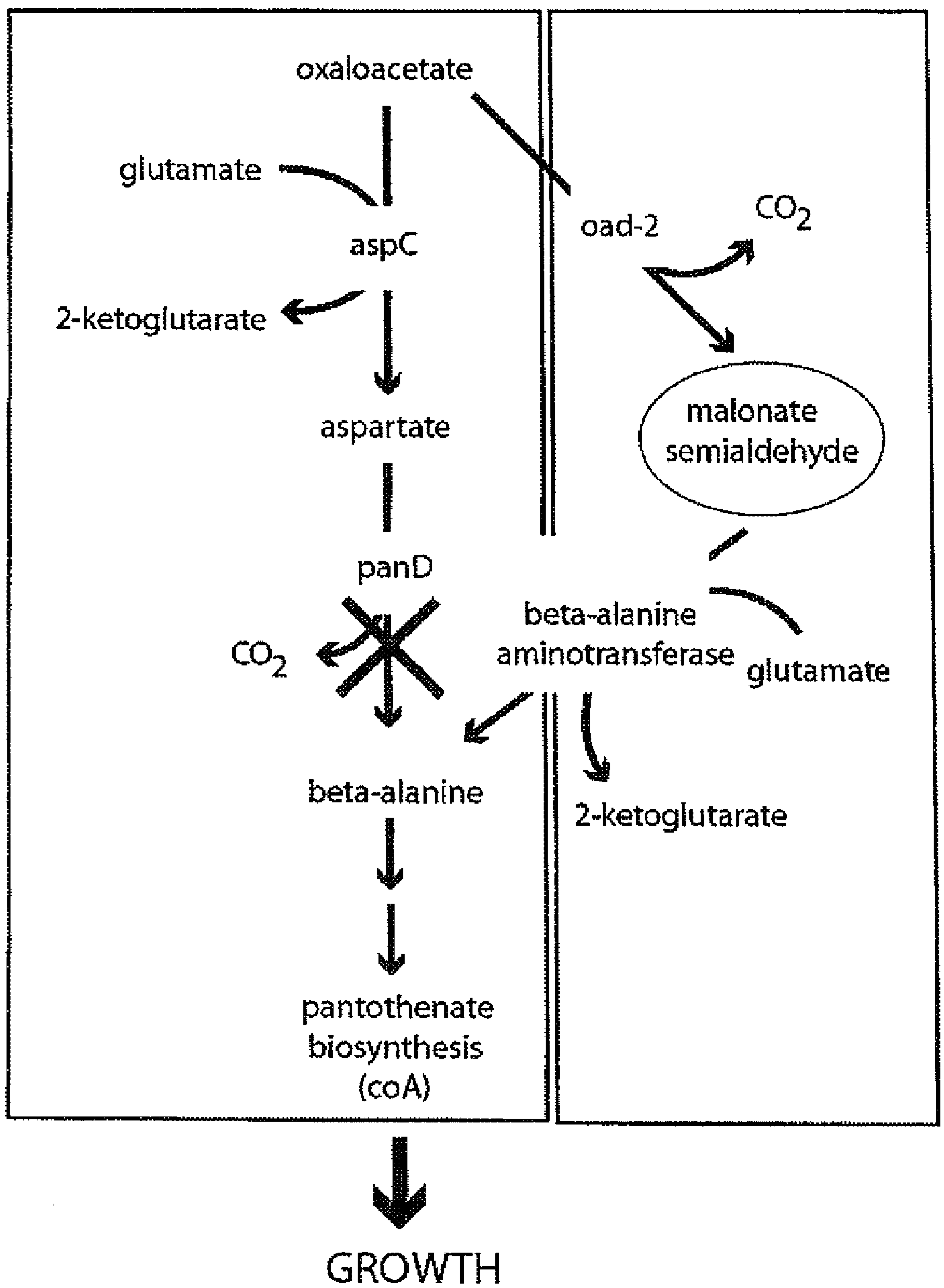

oxaloacetate
glutamate
aspC
2-ketoglutarate
aspartate
panD
CO_2
beta-alanine
pantothenate biosynthesis (coA)
oad-2
CO_2
malonate semialdehyde
beta-alanine aminotransferase
glutamate
2-ketoglutarate
GROWTH

FIG. 10A

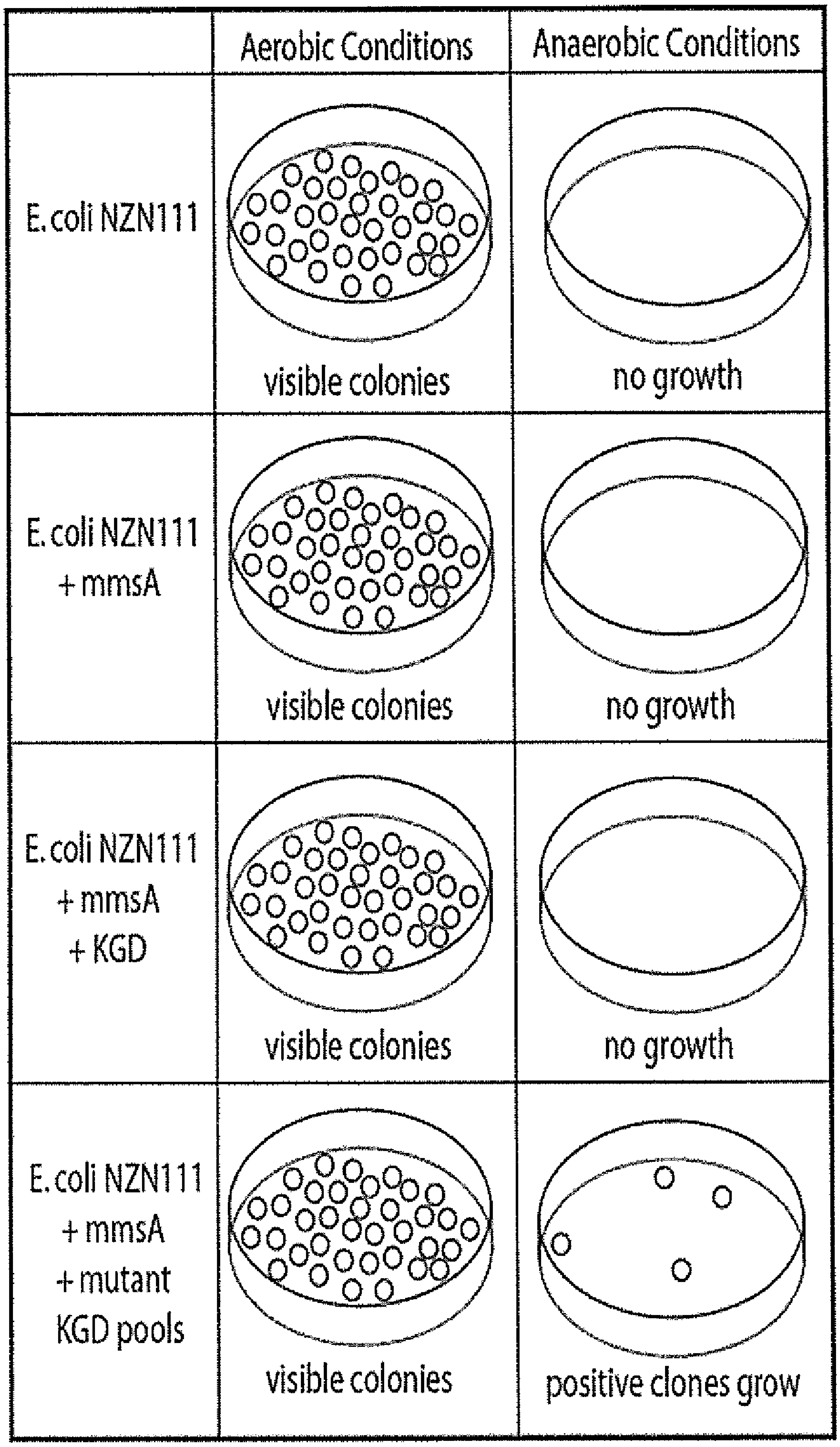
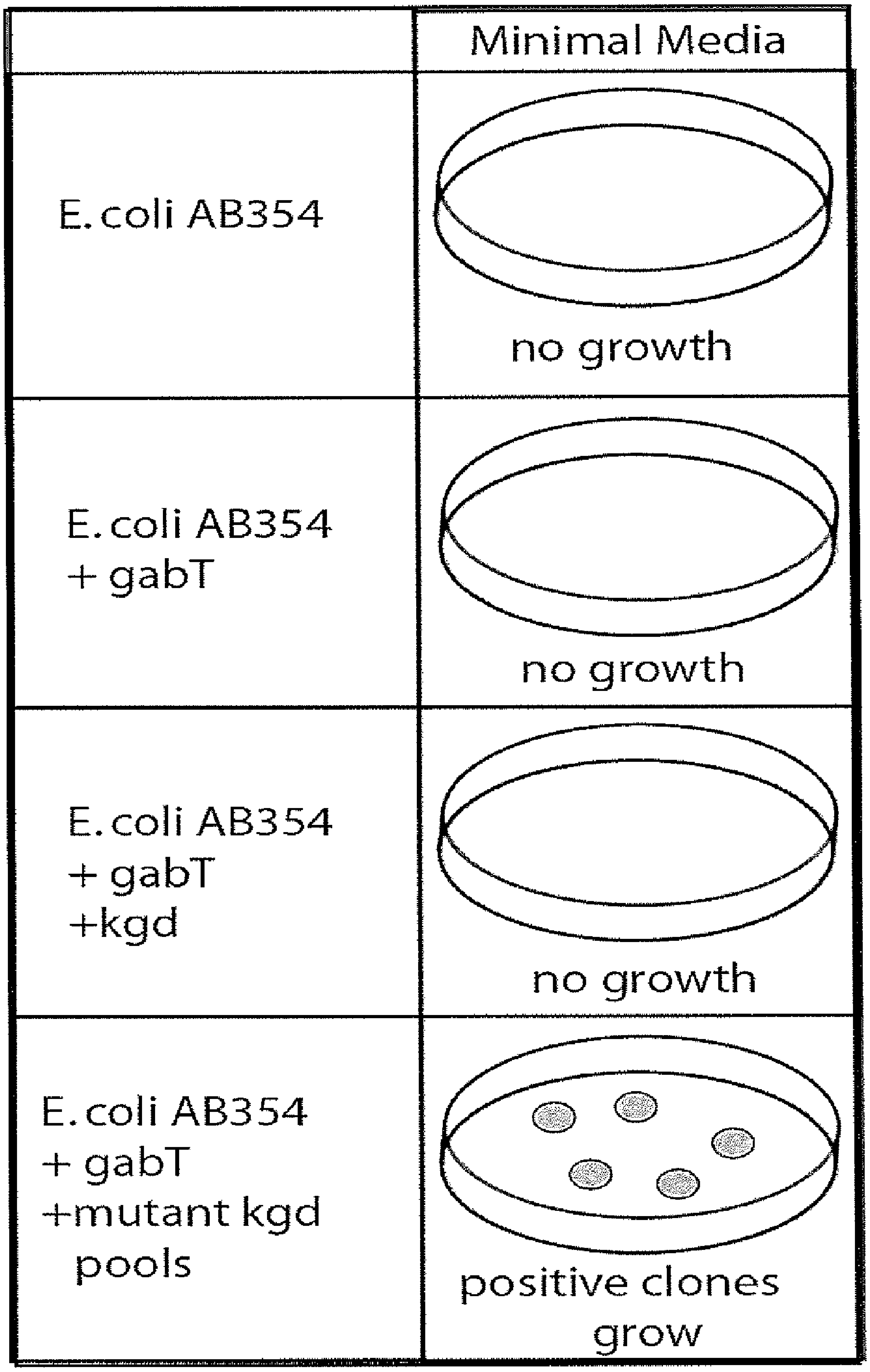

Aerobic Conditions
Anaerobic Conditions
E. coli NZN111
visible colonies
no growth
E. coli NZN111
+ mmsA
visible colonies
no growth
E. coli NZN111
+ mmsA
+ KGD
visible colonies
no growth
E. coli NZN111
+ mmsA
+ mutant
KGD pools
visible colonies
positive clones grow

FIG. 10B


Minimal Media
E. coli AB354
no growth
E. coli AB354
+ gabT
no growth
E. coli AB354
+ gabT
+kgd
no growth
E. coli AB354
+ gabT
+mutant kgd
pools
positive clones
grow

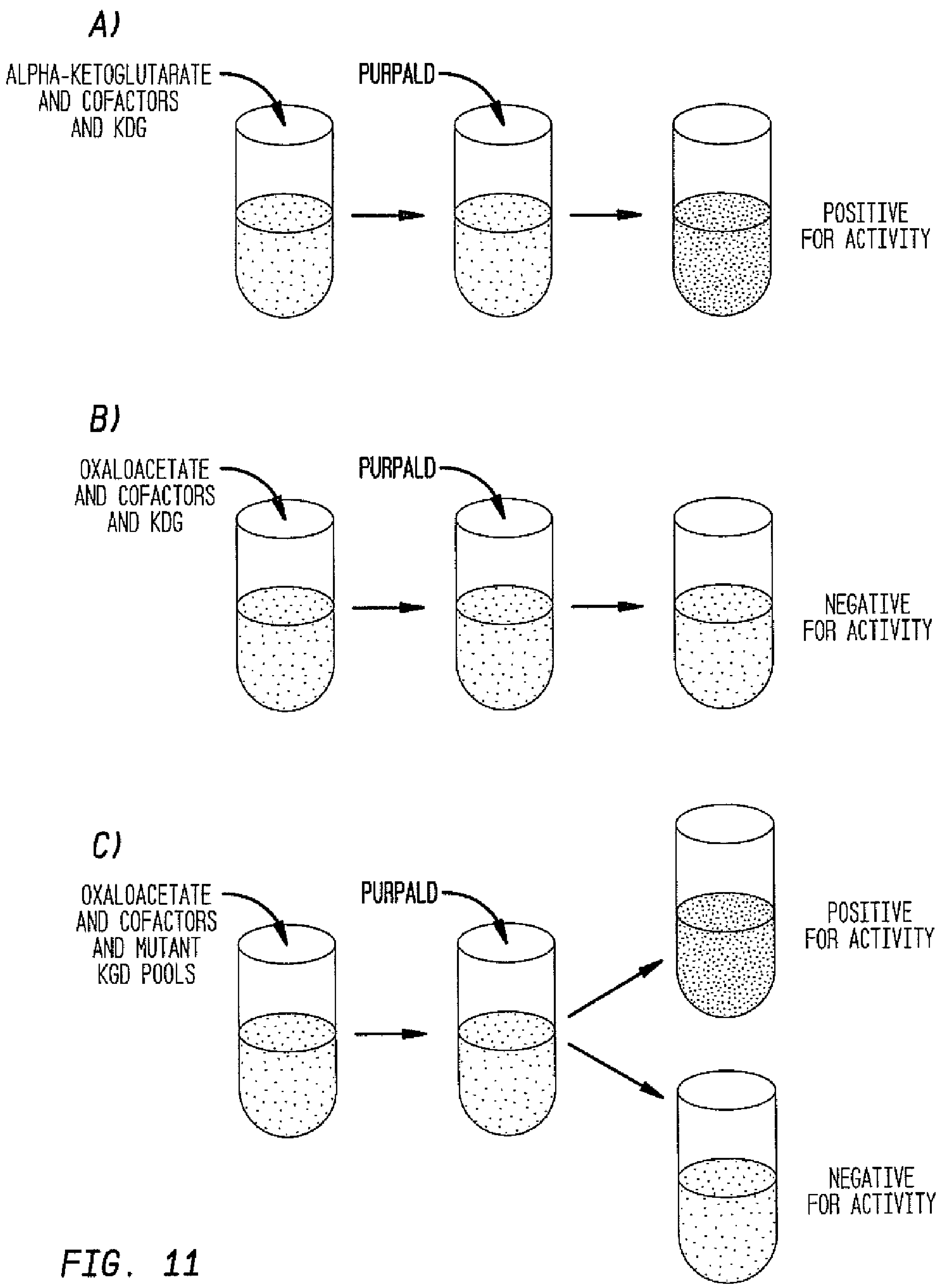

A)
ALPHA-KETOGLUTARATE AND COFACTORS AND KDG
PURPALD
POSITIVE FOR ACTIVITY
B)
OXALOACETATE AND COFACTORS AND KDG
PURPALD
NEGATIVE FOR ACTIVITY
C)
OXALOACETATE AND COFACTORS AND MUTANT KGD POOLS
PURPALD
POSITIVE FOR ACTIVITY
NEGATIVE FOR ACTIVITY

FIG. 11

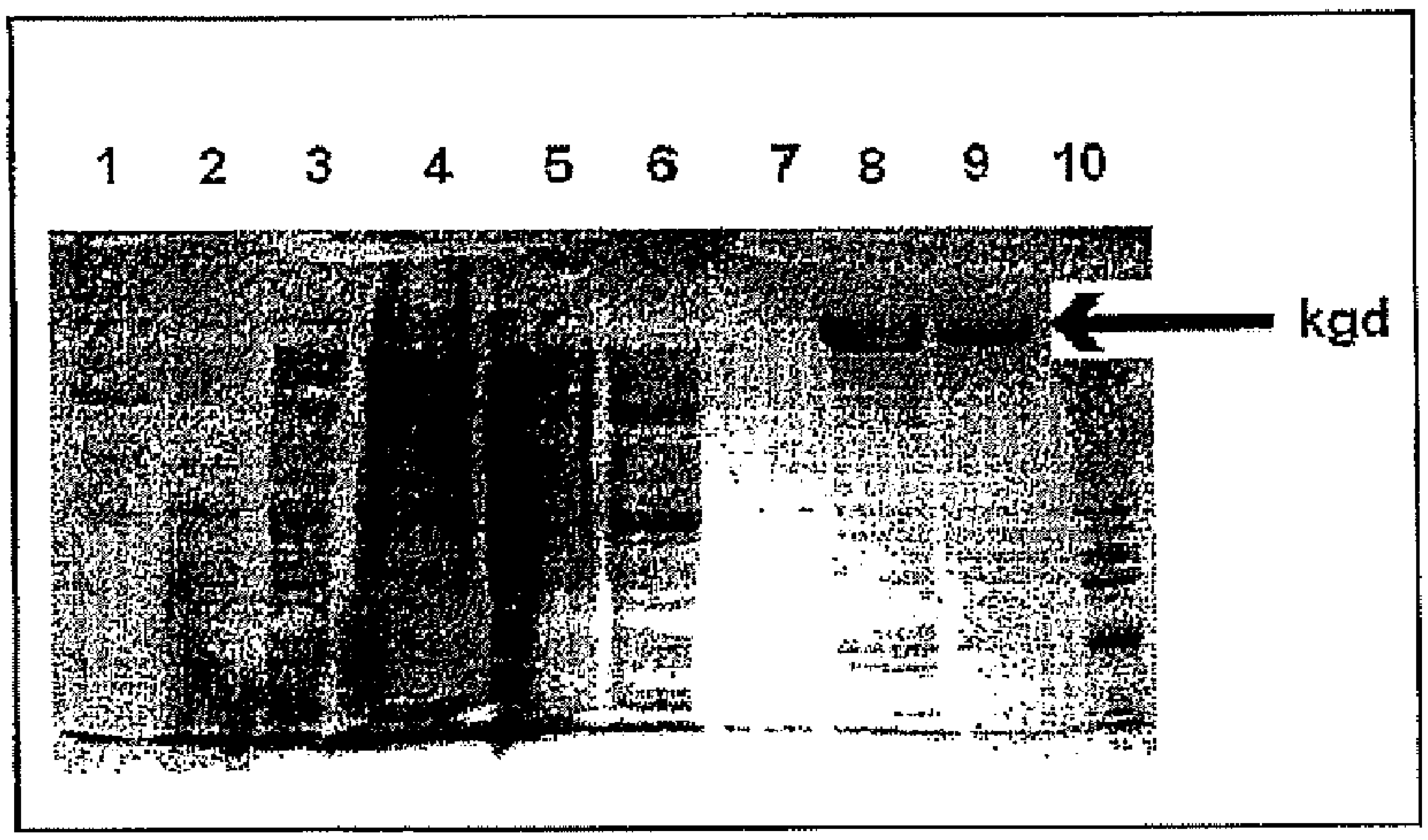


FIG 12

COMPOSITIONS AND METHODS FOR 3-HYDROXYPROPIONATE BIO-PRODUCTION FROM BIOMASS

RELATED APPLICATIONS

[0001] This application is a continuation application which claims priority under 35 USC 120 to U.S. patent application Ser. No. 13/284,337, filed Oct. 28, 2011, which is a continuation of Ser. No. 12/328,588, filed Dec. 4, 2008, and this application also claims priority under 35 USC 119 to U.S. Provisional Patent Application No. 60/992,290, filed Dec. 4, 2007. Both referenced patent applications are incorporated by reference in their respective entireties herewith.

REFERENCE TO A SEQUENCE LISTING

[0002] An electronically filed sequence listing is provided herewith which has been submitted in ASCII format and is hereby incorporated by reference in its entirety. Said ASCII copy, created on Dec. 27, 2013, is named 34246-703.302_SL.txt and is 72,182 bytes in size.

FIELD OF THE INVENTION

[0003] The present invention relates to methods, systems and compositions, including genetically modified microorganisms, i.e., recombinant microorganisms, adapted to exhibit elevated oxaloacetate alpha-oxo decarboxylase activity (also referred to herein as oxaloacetate alpha-decarboxylase activity).

BACKGROUND OF THE INVENTION

[0004] 3-hydroxypropionate ("3-HP", CAS No. 503-66-2) has been identified as a highly attractive potential chemical feedstock for the production of many large market commodity chemicals that are currently derived from petroleum derivatives. For example, commodity products that can be readily produced using 3-HP include acrylic acid, 1,3-propanediol, methyl-acrylate, and acrylamide, as shown in FIG. **1**. The sum value of these commodity chemicals is currently estimated to exceed several billions of dollars annually in the US. However, the current petrochemical manufacturing techniques for these commodities adverse impact the environment via the pollutants generated and the energy used in their production. Manufacture of these same commodities via the clean, cost-effective, production of 3-HP from biomass will simultaneously reduce toxic waste and substitute renewable feed stocks for non-renewable resources. In addition to the environmental benefits associated with bio-based production of 3-HP, if the production cost of the derived commodities is substantially reduced relative to petroleum-based production, this would make a biorefining industry not only environmentally beneficial but also a very attractive investment.

[0005] Previous attempts to produce 3-HP via biological pathways provide product titers which have been low and these processes have required the use of expensive, rich media. Both of these factors limit commercial feasibility and profitability. The use of rich media was necessary due to the toxicity of 3-HP when fermented with the more economical minimal media. For example, in wild type *E. coli*, metabolic activity is significantly inhibited at levels of 3-HP that are 5-10 times lower than the approximate 100 g/L titer needed for economic feasibility using the more economical minimal media. In fact, toxic effects have also been observed in rich media at product titers which are approximately two times lower than desired titers for commercial feasibility (Refer to FIG. **2**). Further, the fermentative pathways reported by other investigators have not addressed and resolved the toxicity mechanisms of 3-HP to the host organisms. Further to issues related to commodity chemical production, which largely relies on petroleum-based starting materials, there is an increasing need to reduce the domestic usage of petroleum and natural gas. The numerous motivating factors for this increasing need include, but are not limited to: pollutant reduction (such as greenhouse gases), environmental protection, and reducing the dependence on foreign oil. These issues not only impact fuel markets, but also the markets of numerous other products that are currently derived from oil. Biorefining promises the development of efficient biological processes allowing for the conversion of renewable sources of carbon and energy into large volume commodity chemicals.

[0006] A biosynthetic route to 3-HP as a platform chemical would be of benefit to the public, not only in terms of reduced dependence on petroleum, but also by a reduction in the amount of pollutants that are generated by current non-biosynthetic processes. Because 3-HP is not currently used as a building block for the aforementioned commodity chemicals, technical hurdles must be surmounted to achieve low cost biological routes to 3-HP. These hurdles include the development of a new organism that not only has a metabolic pathway enabling the production of 3-HP, but is also tolerant to the toxic effects of 3-HP thus enabling the sustained production of 3-HP at economically desired levels.

[0007] There are numerous motivating factors to reduce the domestic usage of petroleum. These factors include, but are not limited to: 1) the negative environmental impacts of petroleum refining such as production of greenhouse gases and the emission of a wide variety of pollutants; 2) the national security issues that are associated with the current dependence on foreign oil such as price instability and future availability; and 3) the long term economic concerns with the ever-increasing price of crude oil. These issues not only impact fuel markets, but also the multi-billion dollar commodity petro-chemical market

[0008] One potential method to alleviate these issues is the implementation of bioprocessing for the conversion of renewable feed stocks (e.g. agricultural wastes) to large volume commodity chemicals. It has been estimated that such bioprocesses already account for 5% of the 1.2 trillion dollar US chemical market. Furthermore, some experts are projecting that up to 50% of the total US chemical market will ultimately be generated through biological means.

[0009] While the attractiveness of such bioprocesses has been recognized for some time, recent advances in biological engineering, including several bio-refining success stories, have accelerated interest in the large scale production of chemicals through biological routes. However, many challenges still remain for the economical bio-production of commodity chemicals. These challenges include the need to convert biomass into usable feed stocks, the engineering of microbes to produce relevant chemicals at high titers and productivities, the improvement of the microbes' tolerance to the desired product, and the need to minimize the generation of byproducts that might affect downstream processes. Finally, the product must be economically competitive in the marketplace.

[0010] The contributions of bioprocessing are expected to grow in the future as existing biological methods become more efficient and as new bioprocesses are developed. A recent analysis by the U.S. Department of Energy identified a list of the Top Value Added Chemicals from Biomass that are good candidates for biosynthetic production. Eight of the top value added chemicals were organic acids, including 3-hydroxypropionic acid (3-HP). As depicted in FIG. **1**, 3-HP is considered to be a platform chemical, capable of yielding valuable derivative commodity chemicals including acrylic acid and acrylic acid polymers, acrylate esters, acrylate polymers (plastics), acrylamide, and 1,3-propanediol. Presently, these high value chemicals are produced from petroleum.

[0011] One method to efficiently generate 3-HP by a bioprocess approach would be the microbial biosynthesis of renewable biomass sugars to 3-HP. According to the DOE Report (Werpy, T.; Petersen, G. Volume 1: Results of Screening for Potential Candidates from Sugars and Synthetic Gas. Oak Ridge, Tenn., U.S. Department of Energy; 2004. Top Value Added Chemicals from Biomass), a number of factors will need to be addressed, including: identifying the appropriate biosynthetic pathway, improving the reactions to reduce other acid co-products, increasing microbial yields and productivities, reducing the unwanted salts, and scale-up and integration of the system. Additionally, as noted above, it is critical to engineer the microbial organism to be tolerant to the potential toxicity of the desired product at commercially significant concentrations.

[0012] The production of acrylic acid from 3-HP is of particular interest because of the high market value of acrylic acid and its numerous derivatives. In 2005, the estimated annual production capacity for acrylic acid was approximately 4.2 million metric tons, which places it among the top 25 organic chemical products. Also, this figure is increasing annually. The demand for acrylic acid may exceed $2 billion by 2010. The primary application of acrylic acid is the synthesis of acrylic esters, such as methyl, butyl or ethyl acrylate. When polymerized, these acrylates are ingredients in numerous consumer products, such as paints, coatings, plastics, adhesives, dispersives and binders for paper, textiles and leather. Acrylates account for 55% of the world demand for acrylic acid products, with butyl acrylate and ethyl acrylate having the highest production volumes. The other key use of the acrylic acid is through polymerization to polyacrylic acid, which is used in hygiene products, detergents, and waste water treatment chemicals. Acrylic acid polymers can also be converted into super absorbent materials (which account for 32% of worldwide acrylic acid demand) or developed into replacement materials for phosphates in detergents. Both of these are fast growing applications for acrylic acid.

[0013] Today, acrylic acid is made in a two step catalytic oxidation of propylene (a petroleum product) to acrolein, and acrolein to acrylic acid, using a molybdenum/vanadium based catalyst, with optimized yields of approximately 90%. It should be noted that several commercial manufacturers of acrylic acid are exploring the use of propane instead of propylene. The use of propane is projected to be more environmentally friendly by reducing energy consumption during production. However, propane is petroleum based, and while its use is a step in the right direction from an energy consumption standpoint, it does not offer the benefits afforded by the bioprocessing route.

[0014] In addition to acrylic acid, acrylates, and acrylic acid polymers, another emerging high value derivative of 3-HP is 1,3-propanediol (1,3-PD). 1,3-PD has recently been used in carpet fiber production for carpets. Further applications of 1,3-PD are expected to include cosmetics, liquid detergents, and anti-freeze. The market for 1,3-PD is expected to grow rapidly as it becomes more routinely used in commercial products.

[0015] Pursuing a cleaner, renewable carbon source route to commodity chemicals through 3-HP will require downstream optimization of the chemical reactions, depending on the desired end product. 3-HP production through bioprocesses directly, or through reaction routes to the high-value chemical derivatives of 3-HP will provide for large scale manufacture of acrylic acid, as well reduction of environmental pollution, the reduction in dependence on foreign oil, and the improvement in the domestic usage of clean methods of manufacturing. Furthermore, the products produced will be of the same quality but at a competitive cost and purity compared to the current petroleum based product.

[0016] Thus, notwithstanding various advances in the art, there remains a need for methods that identify and/or provide, and compositions directed to recombinant microorganisms that have improved 3-HP production capabilities, so that increased 3-HP titers are achieved.

BRIEF DESCRIPTION OF THE DRAWINGS

[0017] FIG. **1** depicts how biomass derived 3-HP can serve as a chemical feedstock to many major chemical commodities worth billions of dollars. (Adapted from Werpy et al. US Dept. Energy, 2004).

[0018] FIG. **2** depicts 3-hydroxpropionate toxicity in *E. coli* K12. The minimum concentration of 3-HP that is required to inhibit visible growth after 24 hrs in minimal media is shown for wild type *E. coli* K12 grown in both minimal media and rich media which contains more complex nutrients.

[0019] FIG. **3** is an overview of SCALEs. (a) Genomic DNA is fragmented to several specific sizes and ligated into vectors creating libraries with defined insert sizes. (b) These libraries are individually transformed into the host cell line used for selections. (c) The pools of transformants are mixed and subjected to selection. Clones bearing inserts with increasing fitness in a given selection have a growth advantage. (d) Enriched plasmids are purified from the selected population, prepared for hybridization, and applied to a microarray. (e) The processed microarray signal is analyzed as a function of genomic sequence position. (f) A nonlinear multi-scale analysis decomposition gives signal not only as a function of position but as a function of scale or library size. (g) Data are visualized and analyzed as a function of genomic position and scale. (for the circular chromosome of *E. coli* shown, genomic position correlates to position around the circle and scale is represented by color. The height of the peak above the circle correlates to population frequency or fitness of a given scale at a given position.)

[0020] FIGS. **4A-4C-2** depict the SCALEs data identifying the chorismate superpathway as a 3-HP target. (A) Fitness data for positions and scales conferring increased fitness of *E. coli* in the presence of 3-HP. Genomic position correlates to position around the circle and scale is represented by color (red=500-1000 bp, yellow=1000-2000 bp, green=4000-8000 bp, blue=8000-10000 bp). The height of the peak above the circle correlates to the fitness of a given scale at a given position. Peaks corresponding to genes involved in the chorismate superpathway are numbered. (B) List of genes in the chorismate superpathway identified in (A). (C) The fitness of each gene identified in (A) is color coded and identified in the chorismate superpathway.

[0021] FIG. **5**A This figure depicts the natural metabolic pathways utilize by *E. coli* during bio-production which results in the natural products lactate, formate and acetate FIG. **5**B. The proposed metabolic pathway to produce 3-HP as a bio-production product. Arrows represent enzymatic activities. The non natural enzymatic function to be evolved in this Phase I project is colored in red. Enzyme activities are as follows [i] glucokinase, [ii] phosphoglucose isomerase, [iii] 6-phosphofructose kinase, [iv] fructose bisphosphate aldolase, [v] triose-phosphate isomerase, [vi] glyceraldehydes 3-phosphate dehydrogenase, [vii] phosphoglycerate kinase, [viii] phosphoglycerate mutase, [ix] enolase, [xi] pyruvate kinase, [xi] lactate dehydrogenase, [xii]pyruvate oxidase, [xiii]pyruvate-formate lyase, [xiv] phosphate acetyltransferase, [xv]acetate kinase, [xvi] phosphoenolpyruvate carboxykinase [xvii] the proposed oxaloacetate alpha-oxo decarboxylase, [xviii] 3-hydropxypropionate dehydrogenase and [xix] malonate semialdehyde dehydrogenase

[0022] FIG. **6** This figure depicts the chemical reaction performed by 2-oxo acid decarboxylases. R can be any group.

[0023] FIG. **7**A This figure depicts the chemical reaction performed by alpha-ketoglutarate decarboxylase encoded by the kgd gene from *M. tuberculosis*. FIG. **7**B depicts the proposed reaction performed by the newly evolved enzyme, oxaloacetate alpha-oxo-decarboxylase. The proposed enzyme will be encoded by the oad-2 gene which will be evolved by mutation from the kgd gene.

[0024] FIG. **8** This figure depicts an overview of the methods to select a diverse library of 2-oxo acid decarboxylases for oxaloacetate alpha-oxo-decarboxylase activity. [i] A natural2-oxo acid decarboxylase is mutated to create a variant library, [ii] this library is introduced into a microbial host that will not survive in a given environment without the presence of the product of the alpha-oxo-decarboxylase, malonate semialdehyde. [iii]. Positive mutants are identified by growth under selective conditions.

[0025] FIG. **9**A depicts the proposed selection of the metabolism of *E. coli* strain NZN111 is shown in the left box. The pflB gene is disrupted blocking the formation of acetyl-coA in anaerobic conditions. The lack of acetyl-coA formation severely inhibits growth. The proposed additional enzymatic path to acetyl-coA is outlined in the right box. The characterized mmsA gene can supply acetyl-coA under anaerobic conditions if it is supplied with malonate semialdehyde by an oxaloacetate alpha-oxo decarboxylase. Kgd mutants with this activity will allow the strain to grow under anaerobic conditions.

[0026] FIG. **9**B depicts the proposed selection of the relevant metabolism of *E. coli* strain AB354 is summarized in the left box. The panD gene is mutated blocking the synthesis of beta-alanine, an essential precursor for pantothenate (coA). The lack of pantothenate formation abolishes growth on minimal media. The proposed additional enzymatic path to beta-alanine is outlined in the right box. The characterized *R. norvegicus* beta-alanine aminotransferase gene (gab T) can supply beta-alanine if it is supplied with malonate semialdehyde as a substrate. An active oxaloacetate alpha-decarboxylase will supply this substrate and enable growth on minimal media. Kgd mutants with this activity will allow the strain to grow on minimal media.

[0027] FIG. **10**A depicts the anticipated Selection Results of mutant colonies expressing the desired oxaloacetate alpha-oxo-decarboxylase will grow under anaerobic conditions when expressed in *E. coli* NZN111 expressing mmsA. No growth will be observed under these conditions in the *E. coli* NZN111, *E. coli* NZN111+mmsA controls. Or in mutants not expressing the desired activity.

[0028] FIG. **10**B depicts the anticipated Selection Results of mutant colonies expressing the desired oxaloacetate alpha-decarboxylase will grow on minimal media when expressed in *E. coli* AB354 expressing gabT. No growth will be observed under these conditions in the *E. coli* AB354, *E. coli* AB354+gabT controls, or in kgd mutants not expressing the desired activity.

[0029] FIG. **11** depicts the screening Protocol. Purified enzyme will be mixed in vitro with the appropriate substrate and reagents. A) The control reaction for the native alpha-ketoglutarate decarboxylase. B) Predicted results for the native alpha-ketoglutarate decarboxylase with oxaloacetate as a substrate. C) Predicted results for kgd mutants, both positive and negative, for oxaloacetate alpha-decarboxylase activity.

[0030] FIG. **12** Expression and Purification results of pKK223-Cterm-5×His-kgd ('5×His' disclosed as SEQ ID NO: 24). Lane 1=marker; lane 2=uninduced culture; lane 3=induced culture; lane 4=native lysate; lane 5=flowthrough; lane 6=first wash (wash 1); lane 7=last wash (wash 3); lane 8=first elution; lane 9=second elution, purified kg; lane 10=pelleted cell debris. The arrow points to the band comprising purified alpha-ketoglutarate decarboxylase.

DETAILED DESCRIPTION OF EMBODIMENTS OF THE INVENTION

[0031] Generally the invention is directed to compositions and methods for production of target chemical compounds in an organism. Various aspects of the invention are directed to providing altered/modified proteins having different enzymatic activity/function as compared to the unaltered protein. Further aspects, of the invention are directed to recombinant organisms comprising altered/modified pathways which are enhanced for production of a target compound (e.g., 3-HP). In some embodiments, a recombinant organism of the invention is a microorganism or algae. In further embodiments, a recombinant organism is a bacterium (e.g., *E. coli*). In one aspect of the invention, an organism is modified to include one or more genes encoding a protein involved in biosynthesis to enhance production of a target chemical compound (e.g., 3-HP). In further embodiments, such one or more genes encode one or more proteins which enhance the capability of the organism to produce a target chemical compound in culture. In one embodiment, such a chemical compound is 3-HP. In yet a further embodiment, the organism comprises at least one recombinant gene resulting in pyruvate, oxalocetate and acetyl-coA production without committed formate production.

[0032] In another embodiment, the recombinant organism comprises acetyl-coA that is produced via the intermediate malonate semialdehyde. In yet another embodiment, acetyl-coA is produced via the intermediate pyruvate through pyruvate synthase.

[0033] Another aspect of the invention is directed to a method for producing 3-HP comprising growing a recombinant organism of the invention, where the organism comprises an enzyme which converts oxaloacetate to malonate semialdehyde. In further embodiments, the recombinant organism is engineered to delete or substantially reduce activity of one or more genes, where the gene(s) include but are not limited to pfkA, pfkB, ldhA, pta, poxB, pflB or a combination thereof. In yet a further embodiment, the recombinant organism is modified to enhance the activity (such as by increasing expression or improving the relevant functioning) of one or more enzymes including but not limited to pck, mmsA, mmsB, oad-2, homologs thereof, or any combination thereof.

[0034] In one embodiment, a method is provided for producing 3-HP comprising growing an organism under a condition which enhance said 3-HP production, wherein said condition is selected from acetyl-coA production via malonate semialdehyde, acetyal-coA production via pyruvate by pyrvuate synthase, without committed production of formate, homologs thereof and any combination thereof.

[0035] In a further aspect of the invention a recombinant microorganism is provided capable of producing 3-HP at quantities greater than about 10, 15, 20, 30, 40, 50, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140 145 or 150 g/L. In one embodiment, the recombinant organism is capable of producing 3-HP from about 30 to about 100 g/L of biomass/culture.

[0036] In a further aspect of the invention, a bio-production mixture is provided for producing 3HP, said mixture comprising a recombinant microorganism; one or more products selected from a group consisting tyrosine, phenylalanine, para-aminobenzoate, para-hydroxy-benzoate, 2,3,-dihydrobenzoate and shikimate.

[0037] In further embodiments, the mixture comprises a microorganism which is engineered to produce pck, mmsA, mmsB, oad-2, homologs thereof, or a combination thereof. In further embodiments, the microorganism does not produce enzymes selected from a group consisting of pfkA, pfkB, ldhA, pta, poxB, pflB, homologs thereof and a combination thereof. In various embodiments, the microorganism is *E. coli.*

[0038] In one aspect of the invention, an isolated polypeptide is provided possessing oxaloacetate alpha oxo-decarboxylase activity, converting oxaloacetate to malonate semialdehyde. Furthermore, a nucleic acid encoding the polypeptide is provided. In yet a further embodiments, a functional variant for the polypeptide or nucleic acid sequence is provided which is homologous to the reference polypeptide and/or nucleic acid and functions as an oxaloacetate alpha oxo-decarboxylase. In some embodiments such a functional variant has at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96, 97, 98, or 99% identity with alpha-ketoglutarate decarboxylase.

[0039] Accordingly, in various aspects of the invention improved methods for biomass production of 3-HP at higher concentrations are disclosed. With this development, it is feasible to construct *E. coli* strains that are highly tolerant to 3-HP and that will maintain robust metabolic activity in the presence of higher concentrations of 3-HP. In various embodiments, metabolic pathways which support the bio-production of 3-HP are manipulated to increase 3-HP production. In some embodiments, such metabolic pathways do not rely upon or are not affected by metabolic processes that are themselves inhibited by 3-HP.

[0040] Utilizing processes for identification of a 3-HP insensitive bio-production pathway (infra, under "Metabolic Toxicity of 3-HP"), in various embodiments of the invention a bio-production pathway is characterized for the synthesis of 3-HP in *E. coli*. In a further embodiment, an altered 2-oxo acid decarboxylase is utilized in a bio-production pathway to produce 3-HP. In yet further embodiments, a bio-production pathway is utilized incorporating previously characterized and sequenced enzymes that have been reported in the literature, as discussed below under "Previously Characterized Enzymes".

[0041] In various embodiments, a bio-production pathway (shown in FIG. **6**) relies directly or indirectly on the metabolite oxaloacetate through the intermediate malonate semialdehyde. The desired enzymatic activity carries out the conversion of oxaloacetate to malonate semialdehyde. This can be accomplished via a decarboxylation reaction not previously reported by a particular enzyme. More specifically, the decarboxylation of 2-oxo acids, such as oxaloacetate, is accomplished by a well understood set of thiamine pyrophosphate dependant decarboxylases, including pyruvate decarboxylases and branched chain 2-oxo acid decarboxylases. A more recently characterized enzyme from *M. tuberculosis*, alpha ketoglutarate decarboxylase, coded by the kgd gene, possesses catalytic activity with a primary substrate very similar to oxaloacetate, decarboxylating the metabolite alpha-ketoglutarate to succinate semialdehyde. As described in greater detail below, an alpha-ketoglutarate decarboxylase from *M. tuberculosis* is modified into an oxaloacetate alpha-oxo-decarboxylase or a functional variant thereof. In various embodiments, any 2-oxo acid decarboxylase including but not limited to pyruvate decarboxylases form various sources or branched chain 2-oxo acid decarboxylases are modified into an oxaloacetate alpha-oxo decarboxylase or a functional variant thereof. In various embodiments, a "functional variant" is a protein encoded by a sequence having about 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 96, 97, 98, or 99 percent identity with the nucleic acid sequence encoding the modified/altered oxaloacetate alpha-oxo-decarboxylase. In further embodiments, sequence identity can be on the amino acid sequence level, where a functional variant has a sequence identity with the reference sequence of about 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 96, 97, 98, or 99. For example, a functional variant can have sequence identity that is 90 percent, or 95 percent, but where the enzyme still functions as an oxaloacetate alpha-oxo-decarboxylase when expressed in an organism (e.g., microorganism, algae, plant), such as *E. coli.*

[0042] In other embodiments, a microorganism or algae is engineered to follow a preferred 3-HP bio-production pathway and also to enhance tolerance to 3-HP production at commercially viable levels. In some such embodiments, the microorganism is a bacterium, such as *E. coli*. Thus, as one example an *E. coli* strain is constructed and optimized for a desired pathway as discussed herein, wherein enhanced tolerance to 3-HP also is established so as to produce commercially viable titers of product. Accordingly, it is within the conception of the present invention that its teachings, methods and compositions may be combined with other teachings, methods and compositions more specifically directed to 3-HP tolerance improvement, including co-owned and/or licensed inventions.

Metabolic Toxicity of 3-HP

[0043] Severe growth inhibition has been observed for extracellular 3-HP levels as low as 10 g/L in minimal media (pH 7.0), which limits the economic feasibility of 3-HP production as a platform chemical. FIG. **2** demonstrates the toxic affects of 3-HP on *E. coli* when grown in minimal media. These toxic effects have been observed to be far greater when the strains are grown in minimal media as compared to growth in rich media (containing a mixture of all nutrients, amino acids and vitamins). However toxicity at levels below required titers (100 g/L) are still observed in rich media. These data alone indicate that 3-HP may be exerting toxic effects by suppressing central metabolic pathways essential to amino acid metabolism.

Diagnosis of 3-HP Toxicity Mechanisms

[0044] To better understand the toxic effects of 3-HP on *E. coli*, a genome-wide technology is used (multi-Scale Analysis of Library Enrichments (SCALEs)), such as disclosed in U.S. Patent Application Publication No. 20060084098, with related inventions described in U.S. Patent Application Publication Nos. 20080103060 and 200702185333 (the latter entitled "Enhanced Alcohol Tolerant Microorganism and Methods of Use Thereof,) published Sep. 20, 2007), which are incorporated by reference herein in their entirety for their respective teachings of methods that provide important information which may be analyzed to make a discovery of previously unappreciated metabolic relationships. An overview of the SCALES approach as well as sample data are depicted in FIG. **3**.

[0045] This genome-wide approach allows identification of numerous genetic changes that can reduce the toxic effects of 3-HP. The results of our studies (shown in FIG. **4**) identified hundreds of genes and other genetic elements that when at increased copy confer varying levels of tolerance to the presence of 3-HP in *E. coli*. When applied alone, these genetic changes may allow for small increases in tolerance; but when applied together they allow for insight into the 3-HP toxicity mechanisms. By grouping genetic elements that confer tolerance by their metabolic roles key metabolic pathways that are inhibited by 3-HP were identified.

[0046] The data shown in FIG. **4** depict identification of the chorismate superpathway as a target of 3-HP toxicity. In some embodiments, toxicity is alleviated by several processes. For example, the addition of the downstream products of branches of the chorismate superpathway, tyrosine, phenylalanine, para-aminobenzoate (a tetrahydrofolate precursor), para-hydroxy-benzoate (a precursor of ubiquinone) and 2,3-dihydroxybenzoate (an enterobactin precursor) all alleviate toxicity to a degree.

A 3-HP bio-Production Pathway

[0047] The genetic modifications conferring a 3-HP tolerant phenotype can enhance a 3-HP bio-production process utilizing *E. coli*. In addition, the mechanisms identified indicate that several current pathways under consideration for the production of 3-HP may not be viable routes at high levels of production

[0048] In various embodiments, a bio-production pathway is utilized which uses one or more metabolic pathways not negatively affected by 3-HP. Therefore, in some embodiments one or more traditional fermentation pathways in *E. coli* as well as pathways involving amino acid intermediates that are currently being explored by others [9,10] are bypassed in order to enhance production. In certain embodiments, a pathway to produce 3-HP is that depicted in FIG. **5**. Also, one or more gene deletions in *E. coli* are effectuated as well as the expression of several enzymatic functions new to *E. coli*. In some embodiments, the one or more gene deletions are selected genes including but not limited to gene(s) encoding pyruvate kinase (pfkA and pfkB), lactate dehydrogenase (ldhA), phosphate acetyltransferase (pta), pyruvate oxidase (poxB) and pyruvate-formate lyase (pflB) enzymes. In further embodiments, any of the one or more deletions in the preceding are combined with one or more enzyme modifications, where the enzymes include but are not limited to phosphoenolpyruvate carboxykinase (pck), malonate semialdehyde dehydrogenase A (mmsA), malonate semialdehyde dehydrogenase B (mmsB) and oxaloacetate alpha-oxo-decarboxylase (oad-2) enzymes are expressed. It should be understood that the term "deletion" in this context does not necessarily require an entire gene deletion, but rather, a modification sufficient to knock out or effectively reduce function.

[0049] The enzymatic activity (oxaloacetate alpha-oxo-decarboxylase) utilized in the proposed pathway has not been reported in the known scientific literature. The enzyme oxaloacetate alpha-oxo-decarboxylase enhances 3-HP production.

[0050] In various embodiments, a pathway having features valuable for bio-production of organic acids in general and can be viewed as a metabolic starting point for numerous other products and in various different organisms (e.g., bacteria, yeast, algae). In various embodiments, such a pathway enhancer allows intracellular production of the key intermediate acetyl-coA without the committed production of the fermentative byproduct formate normally produced in microorganisms (e.g., *E. coli*) with acetyl-coA under fermentative conditions.

Previously Characterized Enzymes

[0051] In various embodiments, an engineered pathway of the invention comprises several genetic modifications to wild type microorganisms (e.g., *E. coli*), in addition to the expression of the oxaloacetate alpha-oxo decarboxylase. For example, one or more mutations in a microorganism (e.g., *E. coli*) can include but not limited to genes: pykA, pykF, ldhA, pflB, pta and poxB genes. Standard methodologies can be used to generate these gene deletions and such methods are routine in the art (See, for example, Sambrook and Russell, Molecular Cloning: A Laboratory Manual, Third Edition 2001 (volumes 1-3), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., hereinafter "Sambrook and Russell").

[0052] In addition to these one or more genetic deletions, the following enzymatic activities can be expressed to enhance 3-HP production (e.g., in *E. coli*): phosphoenolpyruvate carboxykinase and malonate semialdehyde dehydrogenase. In a further embodiment, the mmsA gene is expressed (e.g., mmsA from *Rattus norvegicus* which has been shown to possess malonate semialdehyde dehydrogenase activity and converts malonate semialdehyde to acetyl-coA). In yet a further embodiment, the mmsB gene is expressed (e.g., mmsB gene from *Pseudomonas aeruginosa* which has been shown to have 3-hydropxypropionate dehydrogenase activity). In another further embodiment, a GDP dependant phosphoenolpyruvate carboxykinase is expressed (e.g., gene from *Alcaligenes eutrophus* which has been characterized with kinetics favoring the desired direction producing oxaloacetate). Any genes disclosed herein can be readily synthesized using standard methodologies.

2-Oxo Acid Decarboxylases.

[0053] Several 2-oxo decarboxylases (also referred to as 2-keto acid decarboxylases, alpha-oxo decarboxylases, or alpha-keto acid decarboxylases) with a broad substrate range have been previously characterized, including several pyruvate and branched chain 2-keto-acid decarboxylases. In various embodiments, enzymes from this class of decarboxylases are utilized. The reaction carried out by these enzymes is depicted in FIG. **6**. Of additional interest is that a convenient colorimetric method has been developed to assay this enzymatic activity by detection of the products of this enzyme class which are all aldehydes. In one embodiment, a previously characterized enzyme, alpha ketoglutarate decarboxylase, encoded by the kgd gene from *Mycobacterium tuberculosis* is used. The enzymatic reaction performed by this enzyme is depicted in FIG. **7**A, which is very similar to the desired enzymatic activity, the decarboxylation of oxaloacetate to malonate semialdehyde depicted in FIG. **7**B.

Altered Enzyme Activity

[0054] In one embodiment, clones comprising enhanced oxaloacetate alpha-oxo-decarboxylase activity are obtained by mutation of a gene encoding an enzyme having a similar catalytic activity, namely 2-oxo acid decarboxylases. For example, mutant libraries of a 2-oxo acid decarboxylase gene are constructed. Oxaloacetate alpha-oxo-decarboxylase activity is selected from a mutant library of a 2-oxo acid decarboxylase genes and in one embodiment from a mutant library of the kgd gene encoding an alpha-ketoglutarate decarboxylase. In further embodiments, mutant genes encoding enzymes that modulate or enhance the desired activity are identified.

Overview

[0055] To obtain the desired altered enzyme, a mutant library of a 2-oxo acid decarboxylase gene is constructed, which will be used for selections. In various embodiments, various 2-oxo acid decarboxylase genes are cloned into an appropriate expression system for *E. coli*. Several 2-keto acid decarboxylases with a broad substrate range have been previously characterized (Pohl, M., Sprenger, G. A., Muller, M., A new perspective on thiamine catalysis. Current Opinion in Biotechnology, 15(4), 335-342 (2004)). Of particular interest is an enzyme from *M. tuberculosis*, alpha-ketoglutarate decarboxylase, kgd, which has been purified and characterized (Tian, J Bryk, R. ltoh, M., Suematsu, M., and Carl Nathan, C. Variant tricarboxylic acid cycle in *Mycobacterium tuberculosis*: Identification of alpha-ketoglutarate decarboxylase. PNAS. Jul. 26, 2005 vol. 102(30): 10670-10677; Stephanopoulos, G., Challenges in engineering microbes forbiofuels production. Science, 2007. 315(5813):801-804). Numerous 2-oxo acid decarboxylase genes are known in the art, including but limited to pyruvate decarboxylases from several sources, branched-chain 2-keto acid decarboxylases from various sources, benzylformate decarboxylases from various sources and phenylpyruvate decarboxylases from several sources (refer to www.metacyc.org for a more complete list). In one embodiment, the kgd gene, encoding and alpha-ketoglutarate decarboxylase from *M. tuberculosis* is cloned into an appropriate expression system for *E. coli*. Subsequently, this expression clone is mutated to create a library of mutant clones.

Cloning an 2-Oxo Acid Decarboxylase Gene

[0056] Cloning and expression of any 2 oxo-acid decarboxylase gene including but limited to the kgd gene is performed via gene synthesis supplied from a commercial supplier using standard or conventional techniques. Therefore, no culturing or manipulating of *M. tuberculosis* is required in the case of kgd. In addition, gene synthesis allows for codon optimization for a particular host. Once obtained using standard methodology, the gene is cloned into an expression system using standard techniques.

Construction of a 2-Oxo Acid Decarboxylase Gene Library

[0057] The plasmid containing the cloned 2-oxo acid decarboxylase gene, including but limited to the kgd gene is mutated by standard methods resulting in a large library of mutants. Generally, any of a number of well-known standard methods may be used (See, for example, chapters 1-19 of Directed Evolution Library Creation Methods and Protocols, F. H. Arnold & G. Georgiou, Eds., Methods in Molecular Biology, Vol. 231, Humana Press (2003)). The mutant sequences are introduced into a new host cell line, generating a final library for subsequent selection.

Selection of Altered Activity

[0058] A selection based approach such as described herein can result in the rapid identification of a t-oxo acid decarboxylase mutant with oxaloacetate alpha-oxo-decarboxylase activity. In one example, an available strain of *E. coli*, strain NZN111 is utilized as a host for the selection. This *E. coli* strain has deletions in both the ldhA and pflB genes resulting in severely limited growth (–10 hr doubling time) under anaerobic conditions (See right side of FIG. **5**). This growth limitation is due in part to the inability to produce the necessary metabolite acetyl-coA under these conditions. (See FIG. **9**A below.) A strain of *E. coli* NZN111 expressing mmsA (*E. coli* NZN111+mmsA) in addition to a mutant 2-oxo acid decarboxylase gene, including but limited to the kgd gene, having oxaloacetate alpha-oxo-decarboxylase activity is capable of producing the metabolite acetyl-coA from the metabolic intermediate malonate semialdehyde in media supplemented with tartrate (tartate can be used as a supplement and is readily converted to oxaloacetate in *E. coli*.). This proposed strain has increased growth under anaerobic conditions when compared to both *E. coli* NZN111 and *E. coli* NZN11+mmsA, controls. For example, such a selection is depicted in FIG. **10**A. In one embodiment, *E. coli* NZN111 is constructed to express an acetylating malonate semialdehyde dehydrogenase.

[0059] Similar to the 2-oxo acid decarboxylase gene, an acetylating malonate semialdehyde dehydrogenase gene, including but not limited to mmsA, from *Pseudomonas aeruginosa PAOJ*, is obtained via gene synthesis from the commercial provider. It is subsequently be cloned into an expression plasmid.

[0060] In another example, an available strain of *E. coli*, strain AB354 is utilized as a host for the selection. This *E. coli* strain has a mutation in the panD genes resulting in severely limited growth in minimal media conditions, without the supplementation of beta-alanine (See right side of FIG. **5**). This growth limitation is due to the inability to produce beta-alanine under these conditions. (See FIG. **9**B below.) A strain of *E. coli* AB354 expressing a beta alanine aminotransferase (*E. coli* AB354+beta alanine aminotransferase) in addition to a mutant 2-oxo acid decarboxylase gene, including but limited to the kgd gene, having oxaloacetate alpha-oxo-decarboxylase activity is capable of producing the metabolite beta-alanine from the metabolic intermediate malonate semialdehyde in minimal media. This proposed strain has a recovered ability to grow in minimal media with supplementation of beta-alanine. For example, such a selection is depicted in FIG. **10**B. In one embodiment, *E. coli* AB354 is constructed to express a beta-alanine pyruvate aminotransferase.

[0061] Similar to the 2-oxo acid decarboxylase gene, a beta-alanine pyruvate aminotransferase gene, including but not limited to PA0132 from *Pseudomonas aeruginosa* PAOJ, is obtained via gene synthesis from the commercial provider. It is subsequently be cloned into an expression plasmid.

Selection of Oxaloacetate Alpha-Oxo-Decarboxylase Activity

[0062] The mutant library of kgd genes is introduced into *E. coli* strain NZN111 expressing the mmsA gene. This population is grown under anaerobic conditions in media supplemented with oxaloacetate. Individual mutants expressing the desired oxaloacetate alpha-oxo-decarboxylase activity show increased growth rates compared to the control strains. These clones are isolated and the mutant protein they express subsequently screened for oxaloacetate alpha-oxo-decarboxylase activity as described above.

Colorimetric Confirmation of Decarboxylase Activity

[0063] A colorimetric approach is taken from current standard methodologies. This approach necessitates the expression and purification of the mutant enzymes and reaction with the purified enzyme, its cofactor (thiamin pyrophosphate) and the appropriate substrate. Protein expression and purification are performed with standard methodologies.

[0064] The above description of an approach using NZN111 is meant to be exemplary and not limiting. Its teachings may be applied to other microorganism systems to achieve the desired results. For example, and also not meant to be limiting, use of metabolic features of another *E. coli* strain, AB354, is explained in some of the examples below.

Examples Section

[0065] The following examples disclose specific methods for providing an *E. coli* cell with heterologous nucleic acid sequences that encode for enzymes or other polypeptides that confer increased tolerance to 3-HP. Where there is a method to achieve a certain result that is commonly practiced in two or more specific examples (or for other reasons), that method may be provided in a separate Common Methods section that follows the examples. Each such common method is incorporated by reference into the respective specific example that so refers to it. Also, where supplier information is not complete in a particular example, additional manufacturer information may be found in a separate Summary of Suppliers section that may also include product code, catalog number, or other information. This information is intended to be incorporated in respective specific examples that refer to such supplier and/or product. In the following examples, efforts have been made to ensure accuracy with respect to numbers used (e.g., amounts, temperatures, etc.), but some experimental error and deviation should be accounted for. Unless indicated otherwise, temperature is in degrees Celsius and pressure is at or near atmospheric pressure at approximately 5340 feet (1628 meters) above sea level. It is noted that work done at external analytical and synthetic facilities was not conducted at or near atmospheric pressure at approximately 5340 feet (1628 meters) above sea level. All reagents, unless otherwise indicated, were obtained commercially.

[0066] The meaning of abbreviations is as follows: "C" means Celsius or degrees Celsius, as is clear from its usage, "s" means second(s), "min" means minute(s), "h," "hr," or "hrs" means hour(s), "psi" means pounds per square inch, "nm" means nanometers, "d" means day(s), "μ.LL" or "uL" or "ul" means microliter(s), "mL" means milliliter(s), "L" means liter(s), "mm" means millimeter(s), "nm" means nanometers, "mM" means millimolar, "μ.LM" or "uM" means micromolar, "M" means molar, "mmol" means millimole(s), "flmol" or "uMol" means micromole(s)", "g" means gram(s), "μ.Lg" or "ug" means microgram(s) and "ng" means nanogram(s), "PCR" means polymerase chain reaction, "OD" means optical density, "OD_{600}" means the optical density measured at a wavelength of 600 nm, "kDa" means kilodaltons, "g" means the gravitation constant, "bp" means base pair(s), "kbp" means kilobase pair(s), "% w/v" means weight/volume percent, % v/v" means volume/volume percent, "IPTG" means is opropyl-μ.L-D-thiogalactopyranoiside, "RBS" means ribosome binding site, "rpm" means revolutions per minute, "HPLC" means high performance liquid chromatography, and "GC" means gas chromatography. Also, 10^5 and the like are taken to mean 10^5 and the like.

Example 1

Development of a Plasmid Comprising Kgd

[0067] The nucleic acid sequence for the alpha-ketoglutarate decarboxylase (kgd) from *M. tuberculosis* was codon optimized for *E. coli* according to a service from DNA 2.0 (Menlo Park, Calif. USA), a commercial DNA gene synthesis provider. The nucleic acid sequence was synthesized with an eight amino acid N-terminal tag to enable affinity based protein purification. This nucleic acid sequence incorporated an Nco1 restriction site overlapping the gene start codon and was followed by a HindIII restriction site. In addition a Shine Delgarno sequence (i.e., a ribosomal binding site) was placed in front of the start codon preceded by an Ecor1 restriction site. This codon optimized kgd nucleic acid sequence construct (SEQ ID NO: 1), which is designed to encode for the native kgd protein (SEQ ID NO: 2) was synthesized by DNA 2.0 and then provided in a pJ206 vector backbone (SEQ ID NO: 3).

[0068] A circular plasmid based cloning vector termed pKK223-kgd for expression of the alpha-ketoglutarate decarboxylase in *E. coli* was constructed as follows. The kgd gene in the pJ206 vector was amplified via a polymerase chain reaction with the forward primer being TTTTTTTGTATAC-CATGGATCGTAAATTTCGTGATGATC (SEQ ID NO: 4) containing a Nco1 site that incorporates the start methionine for the protein sequence, and the reverse primer being CCCG-GTGAGATCTAGATCCGAACGCTTCGTC-CAAGATTTCTT (SEQ ID NO: 5) containing a Xba1 site and a Bg111 site that replaces the stop codon of the kgd gene with an in-30 frame protein linker sequence SRS. Also, these primers effectively removed the eight amino acid N-terminal tag. The amplified kgd nucleic acid sequence was subjected to enzymatic restriction digestion with the enzymes Nco1 and Bg111 obtained from New England BioLabs (Ipswich, Mass. USA) according to manufacturer's instructions. The digestion mixture was separated by agarose gel electrophoresis, and visualized under UV transillumination as described in Subsection II of the Common Methods Section. An agarose gel slice containing a DNA piece corresponding to the amplified kgd nucleic acid sequence was cut from the gel and the

DNA recovered with a standard gel extraction protocol and components from Qiagen according to manufacturer's instructions.

[0069] An *E. coli* cloning strain bearing pKK223-3 was grown by standard methodologies and plasmid DNA was prepared by a commercial miniprep column from Qiagen.

[0070] A new DNA vector was created by amplifying a pKK223-3 template by polymerase chain reaction with a forward primer being CGGATCTAGATCTCACCATCACCACCATTAGTCGACCTGCAGCCAAG(SEQ ID NO: 6) and a reverse primer being TGAGATCTAGATCCGTTATGTCCCATGGTTCTGTTTCCTGTGTG (SEQ ID NO: 7). The product was prepared by a commercial PCR-purification column from Qiagen. Both primers contain Xba1 restriction sites that allowed for the linear polymerase chain reaction product to be circularized after restriction digestion with Xba1 with enzymes obtained from New England BioLabs (Ipswich, Mass. USA) according to manufacturer's instructions, and subsequent self-ligation. The new vector, named pKK223-ct-his (SEQ ID NO: 8), contained a multiple cloning region containing the a protein coding cassette under control of a IPTG-inducible promoter with an Nco1 site that incorporate the start methionine and with a Xba1 site and a Bglii site that code for the in-frame protein sequence SRSHHHHH (SEQ ID NO: 9), a multi-histidine tag that allows for metal-affinity protein purification of the expressed protein.

[0071] To insert the gene of interest, kgd, this vector was prepared by restriction digestion with the enzymes Nco1 and Bg111 obtained from New England BioLabs (Ipswich, Mass. USA) according to manufacturer's instructions. The digestion mixture was separated by agarose gel electrophoresis, and visualized under UV transillumination as described under Subsection II of the Common Methods Section. An agarose gel slice containing a DNA piece corresponding to the amplified kgd gene product was cut from the gel and the DNA recovered with a standard gel extraction protocol and components from Qiagen according to manufacturer's instructions. Pieces of purified DNA corresponding to the amplified kgd gene product and the pKK223-cterm-5×his vector backbone ('5×his' disclosed as SEQ ID NO: 24) were ligated and the ligation product was transformed and electroporated according to manufacturer's instructions. The sequence of the resulting vector, termed pKK223-cterm-5×his-kgd (SEQ ID NO: 10, and simply pKK223-kgd such as in the electronic sequence listing, '5×his' disclosed as SEQ ID NO: 24), was confirmed by routine sequencing performed by the commercial service provided by Macrogen (USA). pKK223-cterm-5×his-kgd ('5×his' disclosed as SEQ ID NO: 24) confers resistance to beta-lactamase and contains the kgd gene of *M. tuberculosis* under control of a ptac promoter inducible in *E. coli* hosts by IPTG.

Example 2

Development of a Plasmid Comprising Mcr (Partial Prophetic)

[0072] The nucleic acid sequence for the malonyl-coA reductase gene (mcr) from *Chloroflexus auranticus* was codon optimized for *E. coli* according to a service from DNA 2.0 (Menlo Park, Calif. USA), a commercial DNA gene synthesis provider. Attached and extending beyond the ends of this codon optimized mcr nucleic acid sequence (SEQ ID NO: 11) were an EcoRI restriction site before the start codon and a Hind111 restriction site. In addition a Shine Delgarno sequence (i.e., a ribosomal binding site) was placed in front of the start codon preceded by an EcoRI restriction site. This gene construct was synthesized by DNA 2.0 and provided in a pJ206 vector backbone.

[0073] A circular plasmid based cloning vector termed pKK223-mcr for expression of the malonyl-CoA reductase in *E. coli* was constructed as follows. The mcr gene in the pJ206 vector was amplified via a polymerase chain reaction with the forward primer being TCGTACCAACCATGGCCGGTACGGGTCGTTTGGCTGGTAAAATTG (SEQ ID NO: 12) containing a Nco1 site that incorporates the start methionine for the protein sequence, and the reverse primer being CGGTGTGAGATCTAGATCCGACGGTAATCGCACGACCGCGGT ID NO: 13) containing a Xba1 site and a Bg111 site that replaces the stop codon of the mcr gene with an in-frame protein linker sequence SRS. The amplified mcr nucleic acid sequence was subjected to enzymatic restriction digestion with the enzymes Nco1 and Xba1 obtained from New England BioLabs (Ipswich, Mass. USA) according to manufacturer's instructions. The digestion mixture was separated by agarose gel electrophoresis, and visualized under UV transillumination as described under Subsection II of the Common Methods Section. An agarose gel slice containing a DNA piece corresponding to the amplified mcr nucleic acid sequence was cut from the gel and the DNA recovered with a standard gel extraction protocol and components from Qiagen according to manufacturer's instructions.

[0074] An *E. coli* cloning strain bearing pKK223-3 was grown by standard methodologies and plasmid DNA was prepared by a commercial miniprep column from Qiagen.

[0075] As described in Example 1 above, a new DNA vector was created by amplifying a pKK223-3 template by polymerase chain reaction with a forward primer being CGGATCTAGATCTCACCATCACCACCATTAGTCGACCTGCAGCCAAG (SEQ JD NO: 6) and a reverse primer being TGAGATCTAGATCCGTTATGTCCCATGGTTCTGTTTCCTGTGTG (SEQ ID NO: 7). The product was prepared by a commercial PCR-purification column from Qiagen. Both primers contain Xba1 restriction sites that allowed for the linear polymerase chain reaction product to be circularized after restriction digestion with Xba1 and subsequent self-ligation with enzymes obtained from New England BioLabs (Ipswich, Mass. USA) according to manufacturer's instructions. The vector, named pKK223-ct-his (SEQ ID NO: 8), contained a multiple cloning region containing the a protein coding cassette under control of a IPTG-inducible promoter with an Nco1 site that incorporates the start methionine and with a Xba1 site and a Bg111 site that codes for the in-frame protein sequence SRSHHHHH (SEQ ID NO: 9). The latter multi-histidine sequence allows for metal-affinity protein purification of the expressed protein.

[0076] To insert the gene of interest, mcr, this vector was prepared by restriction digestion with the enzymes Nco1 and Xba1 obtained from New England BioLabs (Ipswich, Mass. USA) according to manufacturer's instructions. The digestion mixture was separated by agarose gel electrophoresis, and visualized under UV transillumination as described under Subsection II of the Common Methods Section.

[0077] Pieces of purified DNA corresponding to the amplified codon optimized mcr nucleic acid sequence and the pKK223-ct-his vector backbone were ligated and the ligation product was transformed and electroporated according to manufacturer's instructions. The sequence of the resulting vector termed pKK223-mcr (SEQ ID NO: 14) is confirmed by routine sequencing performed by the commercial service provided by Macrogen(USA). pKK223-mcr confers resistance to beta-lactamase and contains the mcr gene of *M. tuberculosis* under control of a ptac promoter inducible in *E. coli* hosts by IPTG.

Example 3

Development of a Plasmid Comprising a Beta Alanine-Pyruvate Aminotransferase Gene (Prophetic)

[0078] Introduction of a gene, such as the beta alanine pyruvate aminotransferase gene, into bacterial cells requires the addition of transcriptional (promoters) and translational (ribosome binding site) elements for controlled expression and production of proteins encoded by the gene. A nucleic acid sequence for a gene, whether obtained by gene synthesis or by amplification by polymerase chain reaction from genomic sources, can be ligated to nucleic acid sequences defining these transcriptional and translational elements. The present example discloses the addition of an *E. coli* minimal promoter and ribosome binding site properly oriented in the nucleic acid sequence before a gene of interest.

[0079] The beta alanine pyruvate aminotransferase gene from *Pseudomonas aeruginosa* PAO1 (locus_tag="PA0132") is amplified by polymerase chain reaction from a genomic DNA template with the forward primer being GGGTTTCCATGGACCAGCCGCTCAACGTGG (SEQ ID NO: 15) and the reverse primer being GGGTTTTCAGGCGATGCCGTTGAGCGCTTCGCC (SEQ ID NO: 16). The forward primer incorporates an Nco1 restriction site at the start methionine codon of the gene and the reverse primer includes a stop codon for the gene. The amplified nucleic acid sequence is subjected to enzymatic restriction digestion with the restriction enzyme Nco1 from New England BioLabs (Ipswich, Mass. USA) according to manufacturer's instructions. The digestion mixture is separated by agarose gel electrophoresis, and is visualized under UV transillumination as described under Subsection II of the Common Methods Section. An agarose gel slice containing a DNA piece corresponding to the restricted nucleic acid sequence is cut from the gel and the DNA is recovered with a standard gel extraction protocol and components from Qiagen according to manufacturer's instructions. An *E. coli* tpiA promoter and ribosome binding site is produced by polymerase chain reaction using a forward primer GGGAACGGCGGGGAAAAACAAACGTT (SEQ ID NO: 17) and a reverse primer GGTCCATGGTAATTCTCCACGCTTATAAGC (SEQ ID NO: 18). Using genomic *E. Coli* K12 DNA as the template, a PCR reaction was conducted using these primers.

[0080] The forward primer is complimentary to the nucleic acid sequence upstream of the minimal tpiA promoter region (SEQ ID NO: 19), which is the minimal promoter sequence of the *E. coli* K12 tpi gene. The reverse primer is located just downstream of the minimal promoter region and includes an NcoI restriction site at the location of the start methionine and also includes a ribosome binding site. The PCR-amplified nucleic acid sequence is subjected to enzymatic restriction digestion with the restriction enzyme Nco1 from New England BioLabs (Ipswich, Mass. USA) according to manufacturer's instructions. The digestion mixture is separated by agarose gel electrophoresis, and is visualized under UV transillumination as described in Subsection II of the Common Methods Section. An agarose gel slice containing a DNA piece corresponding to the restricted nucleic acid sequence is cut from the gel and the DNA is recovered with a standard gel extraction protocol and components from Qiagen according to manufacturer's instructions. The restricted, purified nucleic acid piece containing the transcriptional and translational elements is ligated to the recovered DNA containing the gene of interest. The ligation product is used as a template for a subsequent polymerase chain reaction using the forward primer GGGAACGGCGGGGAAAAACAAACGTT (SEQ ID NO: 20). Alternatively, any other forward primer may be use so long as it includes sufficient nucleic acid sequences upstream of the minimal tpiA promoter sequence (SEQ ID NO: 19). In the present specific example, the reverse primer is GGGTTTTCAGGCGATGCCGTTGAGCGCTTCGCC (SEQ ID NO: 21). The amplified nucleic acid product is separated by agarose gel electrophoresis, and is visualized under UV transillumination as described in Subsection II of the Common Methods Section. An agarose gel slice containing a DNA piece corresponding to the restricted nucleic acid sequence is cut from the gel and the DNA is recovered with a standard gel extraction protocol and components from Qiagen according to manufacturer's instructions.

[0081] The resulting nucleic acid piece then is ligated into a suitable plasmid or other vector or transposon or other system, for example pSMART (Lucigen Corp, Middleton, Wis., USA), StrataClone (Stratagene, La Jolla, Calif., USA) or pCR2.1-TOPO TA (Invitrogen Corp, Carlsbad, Calif., USA) according to manufacturer's instructions. These methods also are described in the Subsection II of the Common Methods Section. Accordingly, the resulting nucleic acid piece can be restriction digested and purified andre-ligated into any other vector as is standard in the art. A similar method can be used to combine any gene with any transcriptional and translational elements with variation of restriction sites and primers.

[0082] The resulting nucleic acid is cloned using standard methodologies into the multiple cloning site of plasmid pBT-3, resulting in pBT-3-BAAT. This plasmid expresses the beta-alanine aminotransferase has a replicon compatible with pKK223 based vectors and confers chloramphenicol resistance.

Example 4

Development of a Plasmid Comprising an Acetylating Malonate Semialdehyde Dehydrogenase (Prophetic)

[0083] Introduction of a gene, such as an acetylating malonate semialdehyde dehydrogenase gene, into bacterial cells requires the addition of transcriptional (promoters) and translational (ribosome binding site) elements for controlled expression and production of proteins encoded by the gene. A nucleic acid sequence for a gene, whether obtained by gene synthesis or by amplification by polymerase chain reaction from genomic sources, can be ligated to nucleic acid sequences defining these transcriptional and translational elements. The present example discloses the addition of an *E. coli* minimal promoter and ribosome binding site properly oriented in the nucleic acid sequence before a gene of interest.

[0084] The acetylating malonate semialdehyde dehydrogenase gene, such as is readily available from several sources (e.g., http://ca.expasy.org/cgi-bin/nicezyme.pl?1.2.1.18) is amplified by polymerase chain reaction from a genomic DNA template by standard PCR methodology. The forward primer incorporates an NcoI restriction site at the start methionine codon of the gene and the reverse primer includes a stop codon for the gene. The amplified nucleic acid sequence is subjected to enzymatic restriction digestion with the restriction enzyme Nco1 from New England BioLabs (Ipswich, Mass. USA) according to manufacturer's instructions. The digestion mixture is separated by agarose gel electrophoresis, and is visualized under UV transillumination as described under Subsection II of the Common Methods Section. An agarose gel slice containing a DNA piece corresponding to the restricted nucleic acid sequence is cut from the gel and the DNA is recovered with a standard gel extraction protocol and components from Qiagen according to manufacturer's instructions.

[0085] An *E. coli* tpiA promoter and ribosome binding site is produced by polymerase chain reaction using a forward primer GGGAACGGCGGGGAAAAACAAACGTT (SEQ ID NO: 17) and a reverse primer GGTCCATGGTAATTCTCCACGCTTATAAGC (SEQ ID NO: 18). Using genomic *E. Coli* K12 DNA as the template, a PCR reaction was conducted using these primers. The forward primer is complimentary to the nucleic acid sequence upstream of the minimal tpiA promoter region (SEQ ID NO: 19). The reverse primer is located just downstream of the minimal promoter region and includes an NcoI restriction site at the location of the start methionine and also includes a ribosome binding site. The PCR-amplified nucleic acid sequence is subjected to enzymatic restriction digestion with the restriction enzyme Nco1 from New England BioLabs (Ipswich, Mass. USA) according to manufacturer's instructions. The digestion mixture is separated by agarose gel electrophoresis, and is visualized under UV transillumination as described in Subsection II of the Common Methods Section. An agarose gel slice containing a DNA piece corresponding to the restricted nucleic acid sequence is cut from the gel and the DNA is recovered with a standard gel extraction protocol and components from Qiagen according to manufacturer's instructions.

[0086] The restricted, purified nucleic acid piece containing the transcriptional and translational elements is ligated to the recovered DNA containing the gene of interest. The ligation product is used as a template for a subsequent polymerase chain reaction using the forward primer GGGAACGGCGGGGAAAAACAAACGTT (SEQ ID NO: 17). Alternatively, any other forward primer may be use so long as it includes sufficient nucleic acid sequences upstream of the minimal tpiA promoter sequence (SEQ ID NO: 19). In the present specific example, the reverse primer is GGGTTTTCAGGCGATGCCGTTGAGCGCTTCGCC (SEQ ID NO: 21). The amplified nucleic acid product is separated by agarose gel electrophoresis, and is visualized under UV transillumination as described in Subsection II of the Common Methods Section. An agarose gel slice containing a DNA piece corresponding to the restricted nucleic acid sequence is cut from the gel and the DNA is recovered with a standard gel extraction protocol and components from Qiagen according to manufacturer's instructions.

[0087] The resulting nucleic acid piece then is ligated into a suitable plasmid or other vector or transposon or other system, for example pSMART (Lucigen Corp, Middleton, Wis., USA), StrataClone (Stratagene, La Jolla, Calif., USA) or pCR2.1-TOPO TA (Invitrogen Corp, Carlsbad, Calif., USA) according to manufacturer's instructions. These methods also are described in the Subsection II of the Common Methods Section. Accordingly, the resulting nucleic acid piece can be restriction digested and purified andre-ligated into any other vector as is standard in the art. A similar method can be used to combine any gene with any transcriptional and translational elements with variation of restriction sites and primers.

[0088] The resulting nucleic acid is cloned using standard methodologies into the multiple cloning site of plasmid pBT-3, resulting in pBT-3-mmsA. This plasmid expresses an acetylating malonate semialdehyde dehydrogenase has a replicon compatible with pKK223 based vectors and confers chloramphenicol resistance.

Example 5

Development of a Plasmid Comprising a Pyruvate Decarboxylase. Evolution of Pyruvate Decarboxylase Enzymes for the Enzymatic Conversion of Oxaloacetate to Malonate Semialdehyde (Prophetic)

[0089] Similarly to alpha-ketoglutarate dehydrogenase from *Mycobacterium tuberculosis*, the pyruvate decarboxylase from *Zymomonas mobilis* can be evolved to perform the conversion of oxaloacetate to malonate semialdehyde. The pyruvate decarboxylase enzyme is a thiamine diphosphate-dependent enzyme that decarboxylates 2-keto acids and has been shown to prefer short aliphatic substrates (Siegert P et al. (2005). Exchanging the substrate specificities of pyruvate decarboxylase from *Zymomonas mobilis* and benzoylformate decarboxylase from *Pseudomonas putida*. Protein Eng Des Sel 18, 345-357). Additionally, this enzyme does not require substrate activation by pyruvamide (Hoppner, T. C. & Doelle, H. W. (19S3). Purification and kinetic characteristics of pyruvate decarboxylase and ethanol dehydrogenase from *Zymomonas mobilis* in relation to ethanol production. Eur J Appl Microbial Biotechnol 17, 152-157), and a structure of the protein characterized by x-ray crystallography shows the residues responsible for formation of the substrate and cofactor binding pockets (Dobritzsch D et al. (199S). High resolution crystal structure of pyruvate decarboxylase from *Zymomonas mobilis*. Implications for substrate activation in pyruvate decarboxylases. J Biol Chem 273, 20196-20204). Furthermore, alteration of the substrate specificity of this enzyme by specific amino acid changes have previously been reported (Siegert P etal. (2005). Exchanging the substrate specificities of pyruvate decarboxylase from *Zymomonas mobilis* and benzoylformate decarboxylase from *Pseudomonas putida*. Protein Eng Des SellS, 345-357). An example of a process for randomly mutating specific amino acid regions of this protein follows.

[0090] To evolve the binding pocket of the protein for performing the oxaloacetate to malonate semialdehyde conversion, specific regions of the nucleic acid sequence comprising regions of the protein's amino acid sequence will be mutated. Identification of specific amino acid regions within the protein that are involved in the binding pocket interactions is performed by examining the previously determined crystal structure and also by comparing the protein sequence of the *Zymomonas mobilis* pyruvate decarboxylase with pyruvate decarboxylase from other species showing strong sequence similarity. Using this information, the nucleotide sequence of the gene is examined in order to place restriction sites within the nucleotide sequence at the boundaries of the corresponding amino acid regions identified previously. Form this nucleotide sequence, the *Zymomonas mobilis* pyruvate decarboxylase gene with these restrictions sites is codon optimized for *E. coli* according to a service from DNA 2.0 (Menlo Park, Calif. USA), a commercial DNA gene synthesis provider (SEQ ID NO: 22). This gene construct is synthesized by DNA 2.0 and provided in a pJ206 vector backbone. Additionally, the protein sequence includes the addition of a hepta-histidine purification tag (SEQ ID NO: 25), which can be easily removed by restriction digestion of the plasmid with HindIII followed by self-ligation. The protein for which SEQ ID NO: 22 encodes is provided as SEQ ID NO: 23.

[0091] To specifically mutate amino acids in the pyruvate decarboxylase protein, the plasmid containing the codon-optimized sequence is cut at regions of interest via the incorporated restriction sites. Nucleotide sequences is synthesized or produced by polymerase chain reaction with oligonucleotides designed to incorporate specific or random changes at these regions of interest. These nucleotide sequences will incorporate restriction sites or overhanging ends complimentary to the restriction sites used to cut the plasmid such that the new sequences are ligated into the plasmid to create the desired changes in the protein. These changes can be performed singly or multiply. If these changes are performed multiply, the resulting plasmids are transformed into a panD deleted *E. coli* strain and screened in a manner such as depicted in FIGS. **10**A and **10**B. Additionally, the protein produced by these changes may be assayed in a manner such as depicted in FIG. **11**.

Example 6

Development of a Nucleic Acid Sequence Encoding a Protein Sequence Demonstrating Elevated Oxaloacetate Alpha-Decarboxylase Activity (Partial Prophetic)

[0092] Oxaloacetate alpha-decarboxylase activity is selected from a pool of alpha-ketoglutarate decarboxylase (kgd) mutants by selection in an *E. coli* AB354 host expressing a beta-alanine pyruvate aminotransferase. pKK223-cterm-5×his-kgd ('5×his' disclosed as SEQ ID NO: 24) encoding the kgd gene was constructed as described above. Confirmation of alpha-ketoglutarate decarboxylase protein expression and enzymatic activity with appropriate controls were as follows. *E. Cloni* 10GF' electrocompetent cells (Lucigen, Cat.#60061-1) were transformed with the pKK223-Cterm-5×His-kgd ('5×His' disclosed as SEQ ID NO: 24), plasmid containing sequence for 5×HIS-tagged kgd protein ('5×HIS' disclosed as SEQ ID NO: 24) behind a pTAC promoter. Trans formants were confirmed using restriction digest and DNA sequencing (Macrogen, Korea). Expression and purification of his-tagged-kgd was performed as described in Subsection III of the Common Methods Section. SDS-PAGE results of expression and purification are show in FIG. **12**.

[0093] *E. coli* AB354 (Δ panD) was transformed with the vector controls, pKK223, pKK223-Cterm-5×his ('5×His' disclosed as SEQ ID NO: 24), as well as the test vectors pKK223-mcr and pKK223-Cterm-5×His-kgd ('5×His' disclosed as SEQ ID NO: 24), according to standard methods described below. Each of the strains were grown overnight in LB rich media supplemented with 200 mg/L ampicillin (according to standard protocols). Following overnight growth, cells twice were harvested by centrifugation and washed by resuspension in M9 minimal media (standard protocol), diluted 1:10,000 and plated on M9 minimal media plates with 0.05 g/L threonine, 0.1 g/L leucine, 0.067 g/L thiamine, with the additional appropriate supplements, where indicated at the following concentrations (10 g/L beta-alanine (Sigma Aldrich, St. Louis, Mo.), 1 mM Isopropyl β-D-1-thiogalactopyranoside (Thermo Fisher Scientific, Fairlawn, N.J.), 0.2 g/L putrescine (MPBiomedicals, Santa Ana, Calif.), 200 mg/L ampicillin (Research Products International Corp., Mt. Prospect, Ill.) After plating, agarose plates were incubated at 37 C overnight by standard methods. Table 1 depicts the results of these selection controls. A plus (+) indicates growth on a plate, minus (−) indicates no growth. These data confirm the absence of growth in the selection hosts. Putrescine is known to induce the expression of gamma-aminobutyrate transaminase in *E. coli*. This enzyme has been shown in some species including *Rattus norvegicus* to also have beta-alanine aminotransferase activity. The mer gene encoding the malonyl-coA reductase, has been shown to produce malonate semialdehyde. The lack of growth on the strain expressing malonyl-coA reductase in the presence of putrescine indicates the need for the co-expression of a beta-alanine aminotransferase in *E. coli* AB354 for the selection.

TABLE 1

	Supplements						
Strain	None	β-alanine	Amp	Amp + β-alanine	IPTG + Amp	AMP + Put	Amp + IPTG + Put
K12	+	+	−	−	−	−	−
AB354 (ΔpanD)	−	+	−	−	−	−	−
AB354 (ΔpanD) + pKK223	−	+	−	+	−	−	−
AB354 (ΔpanD) + pKK223-mer	−	+	−	+	−	−	−
AB354 (ΔpanD) + pKK223-kgd	−	+	−	+	−	−	−

[0094] Mutant libraries of pKK223-cterm-5×his-kgd ('5×his' disclosed as SEQ ID NO: 24) were constructed as follows. Plasmid DNA of pKK223-cterm-5×his-kgd ('5×his' disclosed as SEQ ID NO: 24) was purified by standard methods and transformed in the mutator strain *E. coli* XL1-Red (Stratagene, La Jolla, Calif.) according to manufacturer's protocols. Cells were harvested according to manufacturer's protocols and mutated plasmid DNA purified by standard methods.

[0095] Mutant pKK223-cterm-5×his-kgd ('5×his' disclosed as SEQ ID NO: 24) DNA is used to transform an *E. coli* host, AB354+pBT-3-BAAT, described above. Greater than

10^5 transformants are collected from LB ampicillin (200 g/L), Chloramphenicol (40 g/L) agarose plates. Cells are washed in M9 minimal media, diluted 1:10,000 and plated on M9 minimal media plates with 0.05 g/L threonine, 0.1 g/L leucine,

[0096] 0.067 g/L thiamine, with 1 mM Isopropyl-D-1-thiogalactopyranoside (Thermo Fisher Scientific, Fairlawn, N.J.), 200 g/L ampicillin and 40 g/L chloramphenicol. Plates are incubated at 37 C for several days. Colonies that grow are individually collected as positives clones bearing oxaloacetate alpha-decarboxylase activity.

Example 7

Development of a Nucleic Acid Sequence Encoding a Protein Sequence Demonstrating Elevated Oxaloacetate Alpha-Decarboxylase Activity (Prophetic)

[0097] Oxaloacetate alpha-decarboxylase activity is selected from a pool of pyruvate decarboxylase (pdc) mutants by selection in an *E. coli* AB354 host expressing a beta-alanine pyruvate aminotransferase. pKK223-cterm-5×his-pdc ('5×his' disclosed as SEQ ID NO: 24) encoding the pdc gene is constructed as described above. Confirmation of pyruvate decarboxylase protein expression and enzymatic activity with appropriate controls are as follows. *E. Cloni* 10GF' electrocompetent cells (Lucigen, Cat.#60061-1) are transformed with the pKK223-Cterm-5×His-pdc ('5×His' disclosed as SEQ ID NO: 24), plasmid containing sequence for 5×HIS-tagged pdc protein ('5×His' disclosed as SEQ ID NO: 24) behind a pTAC promoter. Transformants are confirmed using restriction digest and DNA sequencing (Macrogen, Korea). Expression and purification of his-tagged-pdc are performed as described in Subsection III of the Common Methods Section.

[0098] *E. coli* AB354 (panD) is transformed with the vector controls, pKK223, pKK223-Cterm-5×his ('5×His' disclosed as SEQ ID NO: 24), as well as the test vectors pKK223-mcr and pKK223-Cterm-5×His-pdc ('5×His' disclosed as SEQ ID NO: 24), according to standard methods described below. Each of the strains is grown overnight in LB rich media supplemented with 200 mg/L ampicillin (according to standard protocols). Following overnight growth, cells twice are harvested by centrifugation and washed by resuspension in M9 minimal media (standard protocol), diluted 1:10,000 and plated on M9 minimal media plates with 0.05 g/L threonine, 0.1 g/L leucine, 0.067 g/L thiamine, with the additional appropriate supplements, where indicated at the following concentrations (10 g/L beta-alanine (Sigma Aldrich, St. Louis, Mo.), 1 mM Isopropyl β-D-1-thiogalactopyranoside (Thermo Fisher Scientific, Fairlawn, N.J.), 0.2 g/L putrescine (MPBiomedicals, Santa Ana, Calif.), 200 mg/L ampicillin (Research Products International Corp., Mt. Prospect, Ill.) After plating, agarose plates were incubated at 37 C overnight by standard methods. Putrescine is known to induce the expression of gamma-aminobutyrate transaminase in *E. coli*. This enzyme has been shown in some species including *Rattus norvegicus* to also have beta-alanine aminotransferase activity. The mer gene encoding the malonyl-coA reductase, has been shown to produce malonate semialdehyde. The lack of growth on the strain expressing malonyl-coA reductase in the presence of putrescine indicates the need for the co-expression of a beta-alanine aminotransferase in *E. coli* AB354 for the selection.

[0099] Mutant libraries of pKK223-cterm-5×his-pdc ('5× his' disclosed as SEQ ID NO: 24) are constructed as follows. Plasmid DNA of pKK223-cterm-5×his-pdc ('5×his' disclosed as SEQ ID NO: 24) are purified by standard methods and transformed in the mutator strain *E. coli* XL1-Red (Stratagene, La Jolla, Calif.) according to manufacturer's protocols. Cells are harvested according to manufacturer's protocols and mutated plasmid DNA purified by standard methods.

[0100] Mutant pKK223-cterm-5×his-pdc ('5×his' disclosed as SEQ ID NO: 24) DNA is used to transform an *E. coli* host, AB354+pBT-3-BAAT, described above. Greater than 10^5 transformants are collected from LB ampicillin (200 g/L), Chloramphenicol (40 g/L) agarose plates. Cells are washed in M9 minimal media, diluted 1:10,000 and plated on M9 minimal media plates with 0.05 g/L threonine, 0.1 g/L leucine,

[0101] 0.067 g/L thiamine, with 1 mM Isopropyl β-D-1-thiogalactopyranoside (Thermo Fisher Scientific, Fairlawn, N.J.), 200 g/L ampicillin and 40 g/L chloramphenicol. Plates are incubated at 37 C for several days. Colonies that grow are individually collected as positives clones bearing oxaloacetate alpha-decarboxylase activity.

Example 8

Development of a Nucleic Acid Sequence Encoding a Protein Sequence Demonstrating Elevated Oxaloacetate Alpha-Decarboxylase Activity (Partial Prophetic)

[0102] Oxaloacetate alpha-decarboxylase activity is selected from a pool of alpha-ketoglutarate decarboxylase (kgd) mutants by selection in an *E. coli* NZN111 host expressing an acetylating malonate semialdehyde dehydrogenase. pKK223-cterm-5×his-kgd ('5×his' disclosed as SEQ ID NO: 24) encoding the kgd gene was constructed as described above. Confirmation of alpha-ketoglutarate decarboxylase protein expression and enzymatic activity with appropriate controls were as follows. *E. Cloni* 10GF' electrocompetent cells (Lucigen, Cat.#60061-1) were transformed with the pKK223-Cterm-5×His-kgd ('5×His' disclosed as SEQ ID NO: 24), plasmid containing sequence for 5×HIS-tagged kgd protein ('5×HIS' disclosed as SEQ ID NO: 24) behind a pTAC promoter. Transformants were confirmed using restriction digest and DNA sequencing (Macrogen, Korea). Expression and purification of his-tagged-kgd were performed as described in Subsection III of the Common Methods Section.

[0103] *E. coli* NZN111 is transformed with the vector controls, pKK223, pKK223-Cterm-5×his ('5×His' disclosed as SEQ ID NO: 24), as well as the test vectors pKK223-mcr and pKK223-Cterm-5×His-kgd ('5×His' disclosed as SEQ ID NO: 24), according to standard methods described below. Each of the strains is grown overnight in LB rich media supplemented with 200 mg/L ampicillin (according to standard protocols). Following overnight growth, cells twice are harvested by centrifugation and washed by resuspension in LB media (standard protocol), diluted 1:10,000 and plated on LB media plates with the additional appropriate supplements, where indicated at the following concentrations 1 mM Isopropyl β-D-1-thiogalactopyranoside (Thermo Fisher Scientific, Fairlawn, N.J.), 200 mg/L ampicillin (Research Products International Corp., Mt. Prospect, Ill.) After plating, agarose plates are incubated at 37 C overnight anaerobically in BD type A Bio-Bags according to manufacturer's instructions (BD Biosciences, Franklin Lakes, N.J., Catalog #261214). The mer gene encoding the malonyl-coA reductase, has been shown to produce malonate semialdehyde. The presence of growth of the strain expressing malonyl-coA reductase in the presence of the co expressed acetylating malonate semialdehyde dehydrogenase in *E. coli* NZN111 serves as a positive control for the selection.

[0104] Mutant libraries of pKK223-cterm-5×his-kgd ('5× his' disclosed as SEQ ID NO: 24) were constructed as follows. Plasmid DNA of pKK223-cterm-5×his-kgd ('5×his' disclosed as SEQ ID NO: 24) were purified by standard methods and transformed into the mutator strain *E. coli* XL1-Red (Stratagene, La Jolla, Calif.) according to manufacturer's protocols. Cells were harvested according to manufacturer's protocols and mutated plasmid DNA purified by standard methods.

[0105] Mutant pKK223-cterm-5×his-kgd ('5×his' disclosed as SEQ ID NO: 24) DNA is used to transform an *E. coli* host, NZN111+pBT-3-mmsA, described above. Greater than 10"5 transformants are collected from LB ampicillin (200 g/L), Chloramphenicol (40 g/L) agarose plates. Cells are washed in LB media, diluted 1:10,000 and plated on LB media plates with 1 mM Isopropyl β-D-1-thiogalactopyranoside (Thermo Fisher Scientific, Fairlawn, N.J.), 200 g/L ampicillin and 40 g/L chloramphenicol. Plates are incubated at 37 C for several days anaerobically in BD type A Bio-Bags according to manufacturer's instructions (BD Biosciences, Franklin Lakes, N.J., Catalog #261214). Colonies that grow are individually collected as positives clones bearing oxaloacetate alpha-decarboxylase activity.

Example 9

Development of a Nucleic Acid Sequence Encoding a Protein Sequence Demonstrating Elevated Oxaloacetate Alpha-Decarboxylase Activity (Prophetic)

[0106] Oxaloacetate alpha-decarboxylase activity is selected from a pool of pyruvate decarboxylase (pdc) mutants by selection in an *E. coli* NZN111 host expressing an acetylating malonate semialdehyde dehydrogenase. pKK223-cterm-5×his-pdc ('5×his' disclosed as SEQ ID NO: 24) encoding the pdc gene is constructed as described above. Confirmation of pyruvate decarboxylase protein expression and enzymatic activity with appropriate controls are as follows. *E. Cloni* 10GF' electrocompetent cells (Lucigen, Cat. #60061-1) are transformed with the pKK223-Cterm-5×His-pdc ('5×His' disclosed as SEQ ID NO: 24), plasmid containing sequence for 5×HIS-tagged pdc protein ('5×His' disclosed as SEQ ID NO: 24) behind a pTAC promoter. Transformants are confirmed using restriction digest and DNA sequencing (Macrogen, Korea). Expression and purification of his-tagged-pdc are performed as described in Subsection III of the Common Methods Section. *E. coli* NZN111 and *E. coli* NZN111+pBT3-mmsA is transformed with the vector controls, pKK223, pKK223-Cterm-5×his ('5×His' disclosed as SEQ ID NO: 24), as well as the test vectors pKK223-mcr and pKK223-Cterm-5×His-pdc ('5×His' disclosed as SEQ ID NO: 24), according to standard methods described below. Each of the strains is grown overnight in LB rich media supplemented with 200 mg/L ampicillin (according to standard protocols). Following overnight growth, cells twice are harvested by centrifugation and washed by resuspension in LB media (standard protocol), diluted 1:10,000 and plated on LB media plates with the additional appropriate supplements, where indicated at the following concentrations 1 mM Isopropyl β-D-1-thiogalactopyranoside (Thermo Fisher Scientific, Fairlawn, N.J.), 200 mg/L ampicillin (Research Products International Corp., Mt. Prospect, Ill.) After plating, agarose plates were incubated at 37 C overnight anaerobically in BD type A Bio-Bags according to manufacturer's instructions (BD Biosciences, Franklin Lakes, N.J., Catalog #261214). The mcr gene encoding the malonyl-coA reductase, has been shown to produce malonate semialdehyde. The presence of growth of the strain expressing malonyl-coA reductase in the presence of the co-expressed acetylating malonate semialdehyde in *E. coli* NZN111 serves as a positive control for the selection.

[0107] Mutant libraries of pKK223-cterm-5×his-pdc ('5× his' disclosed as SEQ ID NO: 24) are constructed as follows. Plasmid DNA of pKK223-cterm-5×his-pdc ('5×his' disclosed as SEQ ID NO: 24) are purified by standard methods and transformed into the mutator strain *E. coli XL*1-Red (Stratagene, La Jolla, Calif.) according to manufacturer's protocols. Cells are harvested according to manufacturer's protocols and mutated plasmid DNA purified by standard methods. Mutant pKK223-cterm-5×his-pdc ('5×his' disclosed as SEQ ID NO: 24) DNA is used to transform an *E. coli* host, NZN111+pBT-3-mmsA, described above. Greater than 10^5 transformants are collected from LB ampicillin (200 g/L), chloramphenicol (40 g/L) agarose plates. Cells are washed in LB media, diluted 1:10,000 and plated on LB media plates with 1 mM Isopropyl β-D-1-thiogalactopyranoside (Thermo Fisher Scientific, Fairlawn, N.J.), 200 g/L ampicillin and 40 g/L chloramphenicol. Plates are incubated at 37 C for several days anaerobically in BD type A Bio-Bags according to manufacturer's instructions (BD Biosciences, Franklin Lakes, N.J., Catalog #261214). Colonies that grow are individually collected as positives clones bearing oxaloacetate alpha-decarboxylase activity.

Example 10

Confirmation of Oxaloacetate Alpha-Decarboxylase Activity (Partial Prophetic)

[0108] The colorimetric to confirm enzymatic decarboxylation of 2-oxo-acid substrates is adapted from current standard methodologies and is illustrated below in FIG. **11**. This approach necessitates the expression and purification of the mutant enzymes and reaction with the purified enzyme, its cofactor (thiamin pyrophosphate) and the appropriate substrate. Protein expression and purification are performed with standard methodologies. This colorimetric screening method will be used both to conduct broad screening for positive oxaloacetate alpha-decarboxylase mutants, and also to conduct confirmatory testing of the positive clones identified in a selection method described above.

[0109] Transformants containing a gene cloned into the pKK223-Cterm-5×his expression vector ('5×his' disclosed as SEQ ID NO: 24) are grown overnight in LB+0.2% glucose+200 ug/mL Ampicillin, diluted 1:20 and grown (LB+0.2% glucose+200 ug/mL Ampicillin) to OD600 of 0.4. IPTG is added at 1 mM final concentration to induce protein expression. Cultures are then allowed to grow at 37 degrees C. for four hours. Cells were harvested by centrifugation at 4 degrees C. for 10 minutes at 4000 rpm. Pellets are resuspended and concentrated 50× (e.g. 500 mL culture resuspended in 10 mL buffer) in Qiagen Ni-NTA Lysis Buffer (50 mM Na2HP04, 300 mM NaCl, 10 mM imidazole, pH 8.0)+1 mM PMSF. Lysozyme is added to a final concentration of 1 mg/mL; cells are incubated on ice for 30 minutes. Cells are lysed using a French Press (cell pressure=2000 psi) three times. Lysates are cleared by centrifugation at 4 degrees C. for 20 minutes, applied to Qiagen Ni-NTA columns, washed and eluted as specified by Qiagen (cat.#31314). Samples are analyzed by SDS-PAGE by routine protocols.

[0110] 100 uL reaction mixtures contain 50 mM Potassium phosphate (pH 7.0), 0.2 mM TPP, 1 mM MgCh, 10 mM of the appropriate substrate. 300 pg of purified enzyme is added to the reaction and incubated 16 hours at 37 degrees C. After 16 hours at 37 degrees C., 100 uL of Purpald colorimetric indicator (as per Sigma-Aldrich, cat.#162892) is added to each well in order to detect formation of corresponding aldehyde product. After addition of the Purpald, reactions are incubated at room temperature for 1 hour and read at a wavelength of 540 nm in a Thermomax Microplate Reader (Molecular Devices) using SOFTMax Pro Microplate Reader software, Ver. 4.0. Absorbances greater than control reactions without substrate are used to determine the presence of decarboxylation.

Common Methods Section

[0111] All methods in this Section are provided for incorporation into the above methods where so referenced therein and/or below.

[0112] Subsection I. Bacterial Growth Methods:

[0113] Bacterial growth culture methods, and associated materials and conditions, are disclosed for respective species that may be utilized as needed, as follows:

[0114] *Escherichia coli* K12 is a gift from the Gill lab (University of Colorado at Boulder) and is obtained as an actively growing culture. Serial dilutions of the actively growing *E. coli* K12 culture are made into Luria Broth (RPI Corp, Mt. Prospect, Ill., USA) and are allowed to grow for aerobically for 24 hours at 37° C. at 250 rpm until saturated.

[0115] *Pseudomonas aeruginosa* genomic DNA is a gift from the Gill lab (University of Colorado at Boulder).

[0116] Subsection II: Gel Preparation, DNA Separation, Extraction, Ligation, and Transformation Methods:

[0117] Molecular biology grade agarose (RPI Corp, Mt. Prospect, Ill., USA) is added to 1×TAE to make a 1% Agarose: TAE solution. To obtain 50×TAE add the following to 900 mL of distilled water: add the following to 900 ml distilled H_2O: 242 g Tris base (RPI Corp, Mt. Prospect, Ill., USA), 57.1 ml Glacial Acetic Acid (Sigma-Aldrich, St. Louis, Mo., USA) and 18.6 g EDTA (Fisher Scientific, Pittsburgh, Pa. USA) and adjust volume to 1 L with additional distilled water. To obtain 1×TAE, add 20 mL of 50×TAE to 980 mL of distilled water. The agarose-TAE solution is then heated until boiling occurred and the agarose is fully dissolved. The solution is allowed to cool to 50° C. before 10 mg/mL ethidium bromide (Acros Organics, Morris Plains, N.J., USA) is added at a concentration of 5 ul per 100 mL of 1% agarose solution. Once the ethidium bromide is added, the solution is briefly mixed and poured into a gel casting tray with the appropriate number of combs (Idea Scientific Co., Minneapolis, Minn., USA) per sample analysis. DNA samples are then mixed accordingly with 5×TAE loading buffer. 5×TAE loading buffer consists of 5×TAE(diluted from SOXTAE as described above), 20% glycerol (Acros Organics, Morris Plains, N.J., USA), 0.125% Bromophenol Blue (Alfa Aesar, Ward Hill, Mass., USA), and adjust volume to 50 mL with distilled water. Loaded gels are then run in gel rigs (Idea Scientific Co., Minneapolis, Minn., USA) filled with 1×TAE at a constant voltage of 125 volts for 25-30 minutes. At this point, the gels are removed from the gel boxes with voltage and visualized under a UV transilluminator (FOTODYNE Inc., Hartland, Wis., USA).

[0118] The DNA isolated through gel extraction is then extracted using the QIAquick Gel Extraction Kit following manufacturer's instructions (Qiagen (Valencia Calif. USA)). Similar methods are known to those skilled in the art.

[0119] The thus-extracted DNA then may be ligated into pSMART (Lucigen Corp, Middleton, Wis., USA), StrataClone (Stratagene, La Jolla, Calif., USA) or pCR2.1-TOPO TA (Invitrogen Corp, Carlsbad, Calif., USA) according to manufacturer's instructions. These methods are described in the next subsection of Common Methods.

Ligation Methods:

[0120] For Ligations into pSMART Vectors:

[0121] Gel extracted DNA is blunted using PCRTerminator (Lucigen Corp, Middleton, Wis., USA) according to manufacturer's instructions. Then 500 ng of DNA is added to 2.5 uL 4× CloneSmart vector premix, 1 ul CloneSmart DNA ligase (Lucigen Corp, Middleton, Wis., USA) and distilled water is added for a total volume of 10 ul. The reaction is then allowed to sit at room temperature for 30 minutes and then heat inactivated at 70° C. for 15 minutes and then placed on ice. *E. cloni* 10G Chemically Competent cells (Lucigen Corp, Middleton, Wis., USA) are thawed for 20 minutes on ice. 40 ul of chemically competent cells are placed into a microcentrifuge tube and 1 ul of heat inactivated CloneSmart Ligation is added to the tube. The whole reaction is stirred briefly with a pipette tip. The ligation and cells are incubated on ice for 30 minutes and then the cells are heat shocked for 45 seconds at 42° C. and then put back onto ice for 2 minutes. 960 ul of room temperature Recovery media (Lucigen Corp, Middleton, Wis., USA) and places into microcentrifuge tubes. Shake tubes at 250 rpm for 1 hour at 37° C. Plate 100 ul of transformed cells on Luria Broth plates (RPI Corp, Mt. Prospect, Ill., USA) plus appropriate antibiotics depending on the pSMART vector used. Incubate plates overnight at 37° C.

For Ligations into StrataClone:

[0122] Gel extracted DNA is blunted using PCRTerminator (Lucigen Corp, Middleton, Wis., USA) according to manufacturer's instructions. Then 2 ul of DNA is added to 3 ul StrataClone Blunt Cloning buffer and 1 ul StrataClone Blunt vector mix amp/kan (Stratagene, La Jolla, Calif., USA) for a total of 6 ul. Mix the reaction by gently pipeting up at down and incubate the reaction at room temperature for 30 minutes then place onto ice. Thaw a tube of StrataClone chemically competent cells (Stratagene, La Jolla, Calif., USA) on ice for 20 minutes. Add 1 ul of the cloning reaction to the tube of chemically competent cells and gently mix with a pipette tip and incubate on ice for 20 minutes. Heat shock the transformation at 42° C. for 45 seconds then put on ice for 2 minutes. Add 250 ul pre-warmed Luria Broth (RPI Corp, Mt. Prospect, Ill., USA) and shake at 250 rpm for 37° C. for 2 hour. Plate 100 ul of the transformation mixture onto Luria Broth plates (RPI Corp, Mt. Prospect, Ill., USA) plus appropriate antibiotics. Incubate plates overnight at 37° C.

For Ligations into pCR2.1-TOPO TA:

[0123] Add 1 ul TOPO vector, 1 ul Salt Solution (Invitrogen Corp, Carlsbad, Calif., USA) and 3 ul gel extracted DNA into a microcentrifuge tube. Allow the tube to incubate at room temperature for 30 minutes then place the reaction on ice. Thaw one tube of TOPlOF' chemically competent cells (Invitrogen Corp, Carlsbad, Calif., USA) per reaction. Add 1 ul of reaction mixture into the thawed TOPlOF' cells and mix gently by swirling the cells with a pipette tip and incubate on ice for 20 minutes. Heat shock the transformation at 42° C. for 45 seconds then put on ice for 2 minutes. Add 250 ul pre-warmed SOC media (Invitrogen Corp, Carlsbad, Calif., USA) and shake at 250 rpm for 37° C. for 1 hour. Plate 100 ul of the transformation mixture onto Luria Broth plates (RPI Corp, Mt. Prospect, Ill., USA) plus appropriate antibiotics. Incubate plates overnight at 37° C.

General Transformation and Related Culture Methodologies:

[0124] Chemically competent transformation protocols are carried out according to the manufactures instructions or according to the literature contained in Molecular Cloning (Sambrook and Russell). Generally, plasmid DNA or ligation products are chilled on ice for 5 to 30 min. in solution with chemically competent cells. Chemically competent cells are a widely used product in the field of biotechnology and are available from multiple vendors, such as those indicated above in this Subsection. Following the chilling period cells generally are heat-shocked for 30 seconds at 42° C. without shaking, re-chilled and combined with 250 microliters of rich media, such as S.O.C. Cells are then incubated at 37° C. while shaking at 250 rpm for 1 hour. Finally, the cells are screened for successful transformations by plating on media containing the appropriate antibiotics.

[0125] The choice of an *E. coli* host strain for plasmid transformation is determined by considering factors such as plasmid stability, plasmid compatibility, plasmid screening methods and protein expression. Strain backgrounds can be changed by simply purifying plasmid DNA as described above and transforming the plasmid into a desired or otherwise appropriate *E. coli* host strain such as determined by experimental necessities, such as any commonly used cloning strain (e.g., DH5a, Top lOP', *E. doni* 10G, etc.).

To Make 1 L M9 Minimal Media:

[0126] M9 minimal media was made by combining 5×M9 salts, 1M $MgS0_4$, 20% glucose, 1M $CaCl_2$ and sterile deionized water. The 5×M9 salts are made by dissolving the following salts in deionized water to a final volume of 1 L: 64 g $Na_2HP0_{4.\ 7}H_20$, 15 g KH_2P0_4, 2.5 g NaCl, 5.0 g NH_4Cl. The salt solution was divided into 200 mL aliquots and sterilized by autoclaving for 15 minutes at 15 psi on the liquid cycle. A 1M solution of $MgSO_4$ and 1M $CaCl_2$ were made separately, then sterilized by autoclaving. The glucose was filter sterilized by passing it thought a 0.22 μm filter. All of the components are combined as follows to make 1 L of M9: 750 mL sterile water, 200 mL 5×M9 salts, 2 mL of 1M $MgS0_4$, 20 mL 20% glucose, 0.1 mL $CaCl_2$, Q.S. to a final volume of 1 L.

To Make EZ Rich Media:

[0127] All media components were obtained from TEKnova (Hollister Calif. USA) and combined in the following volumes. 100 mL 10×MOPS mixture, 10 mL 0.132M K_2HP0_4, 100 mL 10×ACGU, 200 mL 5× Supplement EZ, 10 mL 20% glucose, 580 mL sterile water.

Subsection III. Additional Methods Related to Enzyme Evaluation Expression and Purification of Proteins Expressed in pKK223-Cterm-5×his £'5×his' Disclosed as SEQ ID NO: 24) by Expression Plasmids

[0128] Transformants containing a gene cloned into the pKK223-Cterm-5×his ('5×his' disclosed as SEQ ID NO: 24) expression vector were grown overnight in LB+0.2% glucose+200 ug/mL Ampicillin, diluted 1:20 and grown (LB+0.2% glucose+200 ug/mL Ampicillin) to OD600 of 0.4. IPTG was added at 1 mM final concentration to induce protein expression. Cultures were then allowed to grow at 37 degrees C. for four hours. Cells were harvested by centrifugation at 4 degrees C. for 10 minutes at 4000 rpm. Pellets were resuspended and concentrated 50× (e.g. pellet from 500 mL culture resuspended in 10 mL buffer) in Qiagen Ni-NTA Lysis Buffer (50 mM Na_2HP0_4, 300 mM NaCl, 10 mM imidazole, pH 8.0)+1 mM PMSF. Lysozyme was added to a final concentration of 1 mg/mL; cells were incubated on ice for 30 minutes. Cells were lysed using a French Press (cell pressure=2000 psi) three times. Lysates were cleared by centrifugation at 4 degrees C. for 20 minutes, applied to Qiagen Ni-NTA columns, washed and eluted as specified by Qiagen (cat.#31314). Samples were analyzed by SDS-PAGE by routine protocols.

Decarboxylation Enzyme Reactions:

[0129] 100 uL reaction mixtures were added to microwells. Each 100 uL of reaction mixture contained 50 mM Potassium Phosphate (pH 7.0), 0.2 mM TPP, 1 mM $MgCl_2$, and 10 mM of the appropriate substrate. 300 pg of purified enzyme was added to a respective microwell and incubated 16 hours at 37 degrees C. After 16 hours at 37 degrees C., 100 uL of Purpald® colorimetric indicator (Sigma-Aldrich, cat.#162892), prepared per manufacturer's instructions, was added to each microwell in order to detect formation of corresponding aldehyde product. After addition of the Purpald®, the microwells were incubated at room temperature for 1 hour and read at a wavelength of 540 nm in a Thermomax Microplate Reader (Molecular Devices) using SOFTMax Pro Microplate Reader software, Ver. 4.0.

Summary of Suppliers Section

[0130] The names and city addresses of major suppliers are provided in the methods above. In addition, as to Qiagen products, the DNeasy® Blood and Tissue Kit, Cat. No. 69506, is used in the methods for genomic DNA preparation; the QIAprep® Spin ("mini prep"), Cat. No. 27106, is used for plasmid DNA purification, and the QIAquick® Gel Extraction Kit, Cat. No. 28706, is used for gel extractions as described above.

Bio-Production Media

[0131] Bio-production media, which is used in the present invention with recombinant microorganisms having a biosynthetic pathway for 3-HP (and optionally products further downstream of 3-HP), must contain suitable carbon substrates. Suitable substrates may include, but are not limited to, monosaccharides such as glucose and fructose, oligosaccharides such as lactose or sucrose, polysaccharides such as starch or cellulose or mixtures thereof and unpurified mixtures from renewable feed stocks such as cheese whey permeate, comsteep liquor, sugar beet molasses, and barley malt. Additionally the carbon substrate may also be one-carbon substrates such as carbon dioxide, or methanol for which metabolic conversion into key biochemical intermediates has been demonstrated. In addition to one and two carbon substrates methylotrophic organisms are also known to utilize a number of other carbon containing compounds such as methylamine, glucosamine and a variety of amino acids for metabolic activity. For example, methylotrophic yeast are known to utilize the carbon from methylamine to form trehalose or glycerol (Bellion et al., Microb. Growth Cl Compd., [Int. Symp.], 7th (1993), 415-32. Editor(s): Murrell, J. Collin; Kelly, Don P. Publisher: Intercept, Andover, UK). Similarly, various species of *Candida* will metabolize alanine or oleic acid (Suiter et al., Arch. Microbial. 153:485-489 (1990)). Hence it is contemplated that the source of carbon utilized in the present invention may encompass a wide variety of carbon containing substrates and will only be limited by the choice of organism. Although it is contemplated that all of the above mentioned carbon substrates and mixtures thereof are suitable in the present invention as a carbon source, common carbon substrates used as carbon sources are glucose, fructose, and sucrose, as well as mixtures of any of these sugars. Sucrose may be obtained from feed stocks such as sugar cane, sugar beets, cassava, and sweet sorghum. Glucose and dextrose may be obtained through saccharification of starch based feed stocks including grains such as corn, wheat, rye, barley, and oats.

[0132] In addition, sugars may be obtained from cellulosic and lignocellulosic biomass through processes of pretreatment and saccharification, as described, for example, in US patent application US20070031918A1, which is herein incorporated by reference. Biomass refers to any cellulosic or lignocellulosic material and includes materials comprising cellulose, and optionally further comprising hemicellulose, lignin, starch, oligosaccharides and/or monosaccharides. Biomass may also comprise additional components, such as protein and/or lipid. Biomass may be derived from a single source, or biomass can comprise a mixture derived from more than one source; for example, biomass could comprise a mixture of corn cobs and corn stover, or a mixture of grass and leaves. Biomass includes, but is not limited to, bioenergy crops, agricultural residues, municipal solid waste, industrial solid waste, sludge from paper manufacture, yard waste, wood and forestry waste. Examples of biomass include, but are not limited to, corn grain, corn cobs, crop residues such as corn husks, corn stover, grasses, wheat, wheat straw, barley, barley straw, hay, rice straw, switchgrass, waste paper, sugar cane bagasse, sorghum, soy, components obtained from milling of grains, trees, branches, roots, leaves, wood chips, sawdust, shrubs and bushes, vegetables, fruits, flowers and animal manure. Any such biomass may be used in a bio-production method or system to provide a carbon source. In addition to an appropriate carbon source, such as selected from one of the above-disclosed types, bio-production media must contain 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 3-HP (and optionally products further downstream of 3-HP) production.

Culture Conditions

[0133] Typically cells are grown at a temperature in the range of about 25° C. to about 40° C. in an appropriate medium. Suitable growth media in the present invention are common commercially prepared media such as Luria Bertani (LB) broth, M9 minimal media, Sabouraud Dextrose (SD) broth, Yeast medium (YM) broth or (Ymin) yeast synthetic minimal media. Other defined or synthetic growth media may also be used, and the appropriate medium for growth of the particular microorganism will be known by one skilled in the art of microbiology or bio-production science.

[0134] Suitable pH ranges for the bio-production are between pH 5.0 to pH 9.0, where pH 6.0 to pH 8.0 is a typical pH range for the initial condition.

[0135] Bio-productions may be performed under aerobic, microaerobic, or anaerobic conditions, with or without agitation.

[0136] The amount of 3-HP (and optionally products further downstream of 3-HP) produced in a bio-production media generally can be determined using a number of methods known in the art, for example, high performance liquid chromatography (HPLC) or gas chromatography (GC). Specific HPLC methods for the specific examples are provided herein.

Bio-Production Reactors and Systems:

[0137] Any of the recombinant microorganisms as described and/or referred to above may be introduced into an industrial bio-production system where the microorganisms convert a carbon source into 3-HP (and optionally products further downstream of 3-HP) in a commercially viable operation. The bio-production system includes the introduction of such a recombinant microorganism into a bioreactor vessel, with a carbon source substrate and bio-production media suitable for growing the recombinant microorganism, and maintaining the bio-production system within a suitable temperature range (and dissolved oxygen concentration range if the reaction is aerobic or microaerobic) for a suitable time to obtain a desired conversion of a portion of the substrate molecules to 3-HP (and optionally products further downstream of 3-HP). Industrial bio-production systems and their operation are well-known to those skilled in the arts of chemical engineering and bioprocess engineering. The following paragraphs provide an overview of the methods and aspects of industrial systems that may be used for the bio-production of 3-HP (and optionally products further downstream of 3-HP).

[0138] In various embodiments, any of a wide range of sugars, including, but not limited to sucrose, glucose, xylose, cellulose or hemicellulose, are provided to a microorganism, such as in an industrial system comprising a reactor vessel in which a defined media (such as a minimal salts media including but not limited to M9 minimal media, potassium sulfate minimal media, yeast synthetic minimal media and many others or variations of these), an inoculum of a microorganism providing one or more of the 3-HP (and optionally products further downstream of 3-HP) biosynthetic pathway alternatives, and the a carbon source may be combined. The carbon source enters the cell and is catabolized by well-known and common metabolic pathways to yield common metabolic intermediates, including phosphoenolpyruvate (PEP). (See Molecular Biology of the Cell, 3rd Ed., B. Alberts et al. Garland Publishing, New York, 1994, pp. 42-45, 66-74, incorporated by reference for the teachings of basic metabolic catabolic pathways for sugars; Principles of Biochemistry, 3rd Ed., D. L. Nelson & M. M. Cox, Worth Publishers, New York, 2000, pp 527-658, incorporated by reference for the teachings of major metabolic pathways; and Biochemistry, 4th Ed., L. Stryer, W. H. Freeman and Co., New York, 1995, pp. 463-650, also incorporated by reference for the teachings of major metabolic pathways.). Further to types of industrial bio-production, various embodiments of the present invention may employ a batch type of industrial bioreactor. A classical batch bioreactor system is considered "closed" meaning that the composition of the medium is established at the beginning of a respective bio-production event and not subject to artificial alterations and additions during the time period ending substantially with the end of the bio-production event. Thus, at the beginning of the bio-production event the medium is inoculated with the desired organism or organisms, and bio-production is permitted to occur without adding anything to the system. Typically, however, a "batch" type of bio-production event 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 bio-production event 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 a desired end product or intermediate.

[0139] A variation on the standard batch system is the Fed-Batch system. Fed-Batch bio-production processes are also suitable in the present invention and comprise a typical batch system with the exception that the substrate is added in increments as the bio-production 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 may be measured directly, such as by sample analysis at different times, or 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 approaches are common and well known in the art and examples may be found in Thomas D. Brock in Biotechnology: A Textbook of Industrial Microbiology, Second Edition (1989) Sinauer Associates, Inc., Sunderland, Mass., Deshpande, Mukund V., Appl. Biochem. Biotechnol., 36:227, (1992), and Biochemical Engineering Fundamentals, 2nd Ed. J. E. Bailey and D. F. Ollis, McGraw Hill, New York, 1986, herein incorporated by reference for general instruction on bio-production, which as used herein may be aerobic, microaerobic, or anaerobic. Although the present invention may be performed in batch mode, as provided in Example 8, or in fed-batch mode, it is contemplated that the method would be adaptable to continuous bio-production methods. Continuous bio-production is considered an "open" system where a defined bio-production medium is added continuously to a bioreactor and an equal amount of conditioned media is removed simultaneously for processing. Continuous bio-production generally maintains the cultures within a controlled density range where cells are primarily in log phase growth. Two types of continuous bioreactor operation include: 1) Chemostat—where fresh media is fed to the vessel while simultaneously removing an equal rate of the vessel contents. The limitation of this approach is that cells are lost and high cell density generally is not achievable. In fact, typically one can obtain much higher cell density with a fed-batch process. 2) Perfusion culture, which is similar to the chemostat approach except that the stream that is removed from the vessel is subjected to a separation technique which recycles viable cells back to the vessel. This type of continuous bioreactor operation has been shown to yield significantly higher cell densities than fed-batch and can be operated continuously. Continuous bio-production is particularly advantageous for industrial operations because it has less down time associated with draining, cleaning and preparing the equipment for the next bio-production event. Furthermore, it is typically more economical to continuously operate downstream unit operations, such as distillation, than to run them in batch mode.

[0140] Continuous bio-production 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 media turbidity, is kept constant. Continuous systems strive to maintain steady state growth conditions and thus the cell loss due to the medium being drawn off must be balanced against the cell growth rate in the bio-production. Methods of modulating nutrients and growth factors for continuous bio-production 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.

[0141] It is contemplated that embodiments of the present invention may be practiced using either batch, fed-batch or continuous processes and that any known mode of bio-production would be suitable. Additionally, it is contemplated that cells may be immobilized on an inert scaffold as whole cell catalysts and subjected to suitable bio-production conditions for 3-HP (and optionally products further downstream of 3-HP) production.

[0142] The following published resources are incorporated by reference herein for their respective teachings to indicate the level of skill in these relevant arts, and as needed to support a disclosure that teaches how to make and use methods of industrial bio-production of 3-HP (and optionally products further downstream of 3-HP) from sugar sources, and also industrial systems that may be used to achieve such conversion with any of the recombinant microorganisms of the present invention (Biochemical Engineering Fundamentals, 2nd Ed. J. E. Bailey and D. F. 011 is, McGraw Hill, New York, 1986, entire book for purposes indicated and Chapter 9, pages 533-657 in particular for biological reactor design; Unit Operations of Chemical Engineering, 5th Ed., W. L. McCabe et al., McGraw Hill, New York 1993, entire book for purposes indicated, and particularly for process and separation technologies analyses; Equilibrium Staged Separations, P. C. Wankat, Prentice Hall, Englewood Cliffs, N.J. USA, 1988, entire book for separation technologies teachings).

[0143] The scope of the present invention is not meant to be limited to the exact sequences provided herein. It is appreciated that a range of modifications to nucleic acid and to amino acid sequences may be made and still provide a desired functionality. The following discussion is provided to more clearly define ranges of variation that may be practiced and still remain within the scope of the present invention.

[0144] It is recognized in the art that some amino acid sequences of the present invention can be varied without significant effect of the structure or function of the proteins disclosed herein. Variants included can constitute deletions, insertions, inversions, repeats, and type substitutions so long as the indicated enzyme activity is not significantly affected.

Guidance concerning which amino acid changes are likely to be phenotypically silent can be found in Bowie, J. U., et Al., "Deciphering the Message in Protein Sequences: Tolerance to Amino Acid Substitutions," Science 247:1306-1310 (1990).

[0145] In various embodiments polypeptides obtained by the expression of the polynucleotide molecules of the present invention may have at least approximately 80%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to one or more amino acid sequences encoded by the genes and/or nucleic acid sequences described herein for the 3-HP (and optionally products further downstream of 3-HP) biosynthesis pathways. A truncated respective polypeptide has at least about 90% of the full length of a polypeptide encoded by a nucleic acid sequence encoding the respective native enzyme, and more particularly at least 95% of the full length of a polypeptide encoded by a nucleic acid sequence encoding the respective native enzyme. By a polypeptide having an amino acid sequence at least, for example, 95% "identical" to a reference amino acid sequence of a polypeptide is intended that the amino acid sequence of the claimed polypeptide is identical to the reference sequence except that the claimed polypeptide sequence can include up to five amino acid alterations per each 100 amino acids of the reference amino acid of the polypeptide. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a reference amino acid sequence, up to 5% of the amino acid residues in the reference sequence can be deleted or substituted with another amino acid, or a number of amino acids up to 5% of the total amino acid residues in the reference sequence can be inserted into the reference sequence. These alterations of the reference sequence can occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

[0146] As a practical matter, whether any particular polypeptide is at least 80%, 85%, 90%, 92%, 95%, 96%, 97%, 98% or 99% identical to any reference amino acid sequence of any polypeptide described herein (which may correspond with a particular nucleic acid sequence described herein), such particular polypeptide sequence can be determined conventionally using known computer programs such the Bestfit program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, Wis. 53711). When using Bestfit or any other sequence alignment program to determine whether a particular sequence is, for instance, 95% identical to a reference sequence according to the present invention, the parameters are set, of course, such that the percentage of identity is calculated over the full length of the reference amino acid sequence and that gaps in homology of up to 5% of the total number of amino acid residues in the reference sequence are allowed. For example, in a specific embodiment the identity between a reference sequence (query sequence, a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, may be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. 6:237-245 (1990)). Preferred parameters used in a FASTDB amino acid alignment are: Matrix=PAM 0, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=O, Cutoff Score=1, Window Size=sequence length, Gap Penalty=5, Gap Size Penalty=0.05, Window Size=500 or the length of the subject amino acid sequence, whichever is shorter. According to this embodiment, if the subject sequence is shorter than the query sequence due to N- or C-terminal deletions, not because of internal deletions, a manual correction is made to the results to take into consideration the fact that the FASTDB program does not account for N- and C-terminal truncations of the subject sequence when calculating global percent identity. For subject sequences truncated at the N- and C-termini, relative to the query sequence, the percent identity is corrected by calculating the number of residues of the query sequence that are N- and C-terminal of the subject sequence, which are not matched/aligned with a corresponding subject residue, as a percent of the total bases of the query sequence. A determination of whether a residue is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This final percent identity score is what is used for the purposes of this embodiment. Only residues to the N- and C-termini of the subject sequence, which are not matched/aligned with the query sequence, are considered for the purposes of manually adjusting the percent identity score. That is, only query residue positions outside the farthest N- and C-terminal residues of the subject sequence. For example, a 90 amino acid residue subject sequence is aligned with a 100 residue query sequence to determine percent identity. The deletion occurs at the N-terminus of the subject sequence and therefore, the FASTDB alignment does not show a matching/alignment of the first 10 residues at the N-terminus. The 10 unpaired residues represent 10% of the sequence (number of residues at the N- and C-termini not matched/total number of residues in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 residues were perfectly matched the final percent identity would be 90%. In another example, a 90 residue subject sequence is compared with a 100 residue query sequence. This time the deletions are internal deletions so there are no residues at the N- or C-termini of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only residue positions outside the N- and C-terminal ends of the subject sequence, as displayed in the FASTDB alignment, which are not matched/aligned with the query sequence are manually corrected for.

[0147] The above descriptions and methods for sequence homology are intended to be exemplary and it is recognized that this concept is well-understood in the art. Further, it is appreciated that nucleic acid sequences may be varied and still provide a functional enzyme, and such variations are within the scope of the present invention. Nucleic acid sequences that encode polypeptides that provide the indicated functions for 3-HP (and optionally products further downstream of 3-HP) that increase tolerance or production are considered within the scope of the present invention. These may be further defined by the stringency of hybridization, described below, but this is not meant to be limiting when a function of an encoded polypeptide matches a specified 3-HP (and optionally products further downstream of 3-HP) tolerance-related or biosynthesis pathway enzyme activity.

[0148] Further to nucleic acid sequences, "hybridization" refers to the process in which two single-stranded polynucleotides bind non-covalently to form a stable double-stranded polynucleotide. The term "hybridization" may also refer to triple-stranded hybridization. The resulting (usually) double-stranded polynucleotide is a "hybrid" or "duplex." "Hybridization conditions" will typically include salt concentrations of less than about 1M, more usually less than about 500 mM and less than about 200 mM. Hybridization temperatures can be as low as 5° C., but are typically greater than 22° C., more typically greater than about 30° C., and often are in excess of about 37° C. Hybridizations are usually performed under stringent conditions, i.e. conditions under which a probe will hybridize to its target subsequence. Stringent conditions are sequence-dependent and are different in different circumstances. Longer fragments may require higher hybridization temperatures for specific hybridization. As other factors may affect the stringency of hybridization, including base composition and length of the complementary strands, presence of organic solvents and extent of base mismatching, the combination of parameters is more important than the absolute measure of any one alone. Generally, stringent conditions are selected to be about 5° C. lower than the T. for the specific sequence at a defined ionic strength and pH. Exemplary stringent conditions include salt concentration of at least 0.01 M to no more than 1M Na ion concentration (or other salts) at a pH 7.0 to 8.3 and a temperature of at least 25° C. For example, conditions of 5×SSPE (750 mM NaCl, 50 mM Na Phosphate, 5 mM EDTA, pH 7.4) and a temperature of 25-30° C. are suitable for allele-specific probe hybridizations. For stringent conditions, see for example, Sambrook and Russell and Anderson "Nucleic Acid Hybridization" 1st Ed., BIOS Scientific Publishers Limited (1999), which are hereby incorporated by reference for hybridization protocols. "Hybridizing specifically to" or "specifically hybridizing to" or like expressions refer to the binding, duplexing, or hybridizing of a molecule substantially to or only to a particular nucleotide sequence or sequences under stringent conditions when that sequence is present in a complex mixture (e.g., total cellular) DNA or RNA.

[0149] Having so described the present invention and provided examples, and further discussion, and in view of the above paragraphs, it is appreciated that various non-limiting aspects of the present invention may include:

[0150] A genetically modified (recombinant) microorganism comprising a nucleic acid sequence that encodes a polypeptide with at least 85% amino acid sequence identity to any of the enzymes of any of 3-HP tolerance-related or biosynthetic pathways, wherein the polypeptide has enzymatic activity effective to perform the enzymatic reaction of the respective 3-HP biosynthetic pathway enzyme, and the recombinant microorganism exhibits greater 3-H tolerance and/or 3-HP bio-production.

[0151] A genetically modified (recombinant) microorganism comprising a nucleic acid sequence that encodes a polypeptide with at least 90% amino acid sequence identity to any of the enzymes of any of 3-HP tolerance-related or biosynthetic pathways, wherein the polypeptide has enzymatic activity effective to perform the enzymatic reaction of the respective 3-HP tolerance-related or biosynthetic pathway enzyme, and the recombinant microorganism exhibits greater 3-HP tolerance and/or 3-HP bio-production.

[0152] A genetically modified (recombinant) microorganism comprising a nucleic acid sequence that encodes a polypeptide with at least 95% amino acid sequence identity to any of the enzymes of any of 3-HP tolerance-related or biosynthetic pathways, wherein the polypeptide has enzymatic activity effective to perform the enzymatic reaction of the respective 3-HP tolerance-related or biosynthetic pathway enzyme, and the recombinant microorganism exhibits greater 3-HP tolerance and/or 3-HP bio-production.

[0153] The above paragraphs are meant to indicate modifications in the nucleic acid sequences may be made and a respective polypeptide encoded there from remains functional so as to perform an enzymatic catalysis along one of the 3-HP tolerance-related and/or biosynthetic pathways described above.

[0154] The term "heterologous DNA," "heterologous nucleic acid sequence," and the like as used herein refers to a nucleic acid sequence wherein at least one of the following is true: (a) the sequence of nucleic acids is foreign to (i.e., not naturally found in) a given host microorganism; (b) the sequence may be naturally found in a given host microorganism, but in an unnatural (e.g., greater than expected) amount; or (c) the sequence of nucleic acids comprises two or more subsequences that are not found in the same relationship to each other in nature. For example, regarding instance (c), a heterologous nucleic acid sequence that is recombinantly produced will have two or more sequences from unrelated genes arranged to make a new functional nucleic acid. Embodiments of the present invention may result from introduction of an expression vector into a host microorganism, wherein the expression vector contains a nucleic acid sequence coding for an enzyme that is, or is not, normally found in a host microorganism. With reference to the host microorganism's genome, then, the nucleic acid sequence that codes for the enzyme is heterologous.

[0155] Also, and more generally, in accordance with examples and embodiments herein, there may be employed conventional molecular biology, cellular biology, microbiology, and recombinant DNA techniques within the skill of the art. Such techniques are explained fully in the literature. (See, e.g., Sambrook and Russell, Molecular Cloning: A Laboratory Manual, Third Edition 2001 (volumes 1-3), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.; Animal Cell Culture, R. I. Freshney, ed., 1986). These published resources are incorporated by reference herein for their respective teachings of standard laboratory methods found therein. Further, all patents, patent applications, patent publications, and other publications referenced herein (collectively, "published resource(s)") are hereby incorporated by reference in this application. Such incorporation, at a minimum, is for the specific teaching and/or other purpose that may be noted when citing the reference herein. If a specific teaching and/or other purpose is not so noted, then the published resource is specifically incorporated for the teaching(s) indicated by one or more of the title, abstract, and/or summary of the reference. If no such specifically identified teaching and/or other purpose may be so relevant, then the published resource is incorporated in order to more fully describe the state of the art to which the present invention pertains, and/or to provide such teachings as are generally known to those skilled in the art, as may be applicable. However, it is specifically stated that a citation of a published resource herein shall not be construed as an admission that such is prior art to the present invention.

[0156] Thus, based on the above disclosure, it is appreciated that within the scope of the present invention are methods for selection and identification of mutant polynucleotides comprising nucleic acid sequences that encode mutant polypeptides that demonstrate elevated activity of oxaloacetate alpha-oxo decarboxylase activity (also referred to herein as oxaloacetate alpha-decarboxylase activity). Also within the scope of the present invention may be compositions that comprise such identified mutant polynucleotides and polypeptides. In various embodiments, these methods are directed for the specific purpose of obtaining recombinant microorganisms that have capacity for increased bio-production of 3-HP. Although specific genes, enzymes, plasmids and other constructs are described in the above examples, these are not meant to limit the scope of the invention, particularly in view of the level of skill in the art.

[0157] Thus, while various embodiments of the present invention have been shown and described herein, it will be obvious that such embodiments are provided by way of example only. Numerous variations, changes and substitutions may be made without departing from the invention herein in its various embodiments. Specifically, and for whatever reason, for any grouping of compounds, nucleic acid sequences, polypeptides including specific proteins including functional enzymes, metabolic pathway enzymes or intermediates, elements, or other compositions, or concentrations stated or otherwise presented herein in a list, table, or other grouping (such as metabolic pathway enzymes shown in a figure), unless clearly stated otherwise, it is intended that each such grouping provides the basis for and serves to identify various subset embodiments, the subset embodiments in their broadest scope comprising every subset of such grouping by exclusion of one or more members (or subsets) of the respective stated grouping. Moreover, when any range is described herein, unless clearly stated otherwise, that range includes all values therein and all sub-ranges therein. Accordingly, it is intended that the invention be limited only by the spirit and scope of appended claims, and of later claims, and of either such claims as they may be amended during prosecution of this or a later application claiming priority hereto.

SEQUENCE LISTING

```
<160> NUMBER OF SEQ ID NOS: 25

<210> SEQ ID NO 1
<211> LENGTH: 3736
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      Codon optimized kgd gene sequence with customized ends

<400> SEQUENCE: 1

ttgacggata tcaagcttct attaaccgaa cgcttcgtcc aagatttctt gctgctccac     60

ggcatgaacc ttcgaggaac cgctgctcgg cgcagacatc gcgcgacggg agatacgctt    120

gatgcccgcc aacttgtccg gcagcaactc cggcaactcc aaaccgaaac gcggccaggc    180

gccctggttc gctggttcct cctggaccca aaagaactct ttgacatttt cgtaacgatc    240

cagggtttca cgcagacgac gacgcggcaa cggcgccagt tgctccagac gcacaattgc    300

cagatcatta cggttgtctt tcgccttgcg tgccgccaat tcataataca acttaccgga    360

ggtcaacaga atacggctaa ctttattgcg gtcgccgata ccatcctcgt aggtcggttc    420

ctccaggacg ctacggaatt taatctcggt aaaatcctta atctccgaaa ccgccgcctt    480

gtggcgcagc atggatttcg gggtaaacac gatcaacgga cgttggatgc cgtccaacgc    540

gtggcggcgc agcaagtgaa agtaattgga cggggtgctc ggcatcgcga tcgtcatgct    600

accctcagcc cacagctgca aaaagcgctc aatgcgcgcc gacgtgtggt ccgggccctg    660

accctcgtgg ccgtgtggca gcaacagaac aacgttggac agctgacccc atttcgcttc    720

gcccgagctg atgaactcgt caatgatcga ctgggcacca ttaacaaagt cgccgaattg    780

cgcctcccac agcacaaccg cgtctggatt gccaacggta taaccgtact caaaacccac    840

agccgcgtat tccgacaacg gcgaatcata aaccaggaac ttaccaccgg tcgggctgcc    900

gtcgctgtta gtcgccagca gctgcagcgg ggtgaactcc tcgccggtgt gacggtcgat    960

cagcacagaa tgacgctggc taaaggtgcc acgacgagag tcttgaccgc tcagacggac   1020

cagcttgccc tcagcaacca ggctgcccag cgccagcaac tcaccaaacg cccagtcaat   1080

cttaccctca tacgccatct cacgacgctt ttccagcacc ggctgaacgc gcgggtgtgc   1140

cgtgaaaccg ttaggcagcg ccaggaaggc atcaccgata cgtgccagca gggatttgtc   1200
```

-continued

```
aacagcggtc gccaggcccg ctgggatcat ttggtccgac tcgacgctct ccgacggttg   1260
gacgccgtgc ttctccagtt cgcgcacttc gttgaacaca cgctccagtt ggccctggta   1320
atcgcgcagc gcatcctccg cctctttcat gctgatgtcg ccacgaccga tcagagcctc   1380
ggtgtaggat ttacgggcgc cacgcttggt gtccacgacg tcatagacat acggattggt   1440
catagatgga tcgtcaccct cattatgacc acgacgacga tagcacagca tgtcaataac   1500
aacgtccttt ttaaagcgtt ggcggaagtc aactgccaaa cgcgcaaccc agacacacgc   1560
ttctggatcg tcgccgttca cgtggaagat tggcgcaccg atcatcttcg ccacatccgt   1620
gcagtactcg ctcgaacgag agtattccgg agcggtggtg aagccgattt ggttgttcac   1680
aatgatgtga atcgtaccac ccacacgata gccaggcaga tttgccagat tcagcgtctc   1740
cgcaacaacg ccttgacccg cgaaggccgc atcaccgtgc agcatcaacg gaacgacgga   1800
gaatgcacgt tggccgtcgg agtcaatgga accgtggtcc agcaggtctt gcttcgcacg   1860
caccaaacct tccagcactg gatcgaccgc ttccaaatgg gacggatttg ccgtcaggga   1920
aacctgaata tcgttatcgc caaacatttg cagatacaga cccgtcgcac ccaggtggta   1980
cttgacatcg ccggaaccgt gagcctggga cgggttcaga ttgccttcaa actccgtaaa   2040
gatttgcgaa tacggtttgc ccacgatgtt cgccaggaca ttcaagcgac cacggtgcgg   2100
catgccaatg accacttcat ccaaaccatg ctcggcacat tggtcaatcg ccgcgtccat   2160
cattggaata acagattccg caccctccag gctaaaacgc ttttggccca catacttggt   2220
ttgcaggaag gtttcaaacg cctccgctgc gttcagtttc gacaagatgt acttttgttg   2280
agcaacggtc ggtttgacgt gcttcgtctc gacacgctgc tccagccact ccttttgttc   2340
cgggtccaga atgtgcgcgt actcaacacc gatgtgacgg cagtacgcgt cgcgcagcaa   2400
acccagcacg tcacgcagct tttgtattg agcacccgcg aaaccgtcaa ccttaaagac   2460
gcggtccagg tcccacagag tcaggccatg cgtcaacacc tccaaatccg gatgcgaacg   2520
aaagcgcgcc ttatccaagc gcaacgggtc ggtgtccgcc atcagatggc cgcggttgcg   2580
ataggccgcg atcaggttca tcacacgtgc gttcttgtca acgatcgagt ccggattatc   2640
ggtgctccaa cgcactggca ggtacgggat gctcagctcg cggaagacct catcccagaa   2700
gccatcagac aacagcagtt catgaatggt acgcaggaag tcaccgcttt ccgcaccttg   2760
aatgatacgg tggtcgtagg tagaagtcag ggtaatcagt ttgccaatac ccagctcagc   2820
gatgcgttcc tcggacgcgc cttggaactc cgccggatat tccatcgcac cgacaccgat   2880
gatagcacct tgacctggca tcaggcgtgg cacagagtgc accgtgccaa tcgtgcccgg   2940
attcgtcagc gaaatcgtaa cgccagcgaa gtcctcggtg gtcagtttac catcacgagc   3000
acgacggacg atgtcctcgt acgcggtaac gaactgcgcg aagcgcatcg tctcgcaacg   3060
tttgataccg gccaccacca gagagcgctt gccatcttta ccttgcaggt caatagccag   3120
acccagattg gtgtgcgcag gagtaaccgc cgtcggctta ccgtccacct ccgtgtagtg   3180
gcgattcata ttcgggaact tcttaaccgc ctgaaccaga gcataaccca gcaaatgggt   3240
aaagctgatt ttaccaccac gcgtgcgttt caactgatta ttgatcacga tacgattatc   3300
gatcaacaat ttcgctggca cagcacgcac cgaggtcgcc gtaggcactt ccagcgacgc   3360
gctcatgttc ttcacgacag cagccgccgc accacgcagg acagcaactt catcgccttc   3420
ggctggcgga ggcaccgcgg tcttggcggc cagagccgcc acgacaccgt tgcccgctgc   3480
```

-continued

```
ggccgtgtcg gccggtttcg gcggcgcttg cggagccgcc gctgccgcac gctcagcgac   3540

caaagggctg gtcacacgag tcggctcagc agccggttgg gaagtcggct ccgggctata   3600

gtccaccaga aactcatgcc agcttgggtc aacggaagac ggatcatcac gaaatttacg   3660

atacatacca acacgagacg ggtcctgagt caccatggat atatctcctt cttaaagaat   3720

tcgatatctc agcgac                                                   3736
```

<210> SEQ ID NO 2
<211> LENGTH: 1224
<212> TYPE: PRT
<213> ORGANISM: Mycobacterium tuberculosis

<400> SEQUENCE: 2

```
Met Val Thr Gln Asp Pro Ser Arg Val Gly Met Tyr Arg Lys Phe Arg
1               5                   10                  15

Asp Asp Pro Ser Ser Val Asp Pro Ser Trp His Glu Phe Leu Val Asp
            20                  25                  30

Tyr Ser Pro Glu Pro Thr Ser Gln Pro Ala Ala Glu Pro Thr Arg Val
        35                  40                  45

Thr Ser Pro Leu Val Ala Glu Arg Ala Ala Ala Ala Ala Pro Gln Ala
    50                  55                  60

Pro Pro Lys Pro Ala Asp Thr Ala Ala Ala Gly Asn Gly Val Val Ala
65                  70                  75                  80

Ala Leu Ala Ala Lys Thr Ala Val Pro Pro Pro Ala Glu Gly Asp Glu
                85                  90                  95

Val Ala Val Leu Arg Gly Ala Ala Ala Ala Val Val Lys Asn Met Ser
            100                 105                 110

Ala Ser Leu Glu Val Pro Thr Ala Thr Ser Val Arg Ala Val Pro Ala
        115                 120                 125

Lys Leu Leu Ile Asp Asn Arg Ile Val Ile Asn Asn Gln Leu Lys Arg
    130                 135                 140

Thr Arg Gly Gly Lys Ile Ser Phe Thr His Leu Leu Gly Tyr Ala Leu
145                 150                 155                 160

Val Gln Ala Val Lys Lys Phe Pro Asn Met Asn Arg His Tyr Thr Glu
                165                 170                 175

Val Asp Gly Lys Pro Thr Ala Val Thr Pro Ala His Thr Asn Leu Gly
            180                 185                 190

Leu Ala Ile Asp Leu Gln Gly Lys Asp Gly Lys Arg Ser Leu Val Val
        195                 200                 205

Ala Gly Ile Lys Arg Cys Glu Thr Met Arg Phe Ala Gln Phe Val Thr
    210                 215                 220

Ala Tyr Glu Asp Ile Val Arg Arg Ala Arg Asp Gly Lys Leu Thr Thr
225                 230                 235                 240

Glu Asp Phe Ala Gly Val Thr Ile Ser Leu Thr Asn Pro Gly Thr Ile
                245                 250                 255

Gly Thr Val His Ser Val Pro Arg Leu Met Pro Gly Gln Gly Ala Ile
            260                 265                 270

Ile Gly Val Gly Ala Met Glu Tyr Pro Ala Glu Phe Gln Gly Ala Ser
        275                 280                 285

Glu Glu Arg Ile Ala Glu Leu Gly Ile Gly Lys Leu Ile Thr Leu Thr
    290                 295                 300

Ser Thr Tyr Asp His Arg Ile Ile Gln Gly Ala Glu Ser Gly Asp Phe
305                 310                 315                 320
```

-continued

```
Leu Arg Thr Ile His Glu Leu Leu Leu Ser Asp Gly Phe Trp Asp Glu
                325                 330                 335

Val Phe Arg Glu Leu Ser Ile Pro Tyr Leu Pro Val Arg Trp Ser Thr
            340                 345                 350

Asp Asn Pro Asp Ser Ile Val Asp Lys Asn Ala Arg Val Met Asn Leu
        355                 360                 365

Ile Ala Ala Tyr Arg Asn Arg Gly His Leu Met Ala Asp Thr Asp Pro
    370                 375                 380

Leu Arg Leu Asp Lys Ala Arg Phe Arg Ser His Pro Asp Leu Glu Val
385                 390                 395                 400

Leu Thr His Gly Leu Thr Leu Trp Asp Leu Asp Arg Val Phe Lys Val
                405                 410                 415

Asp Gly Phe Ala Gly Ala Gln Tyr Lys Lys Leu Arg Asp Val Leu Gly
            420                 425                 430

Leu Leu Arg Asp Ala Tyr Cys Arg His Ile Gly Val Glu Tyr Ala His
        435                 440                 445

Ile Leu Asp Pro Glu Gln Lys Glu Trp Leu Glu Gln Arg Val Glu Thr
    450                 455                 460

Lys His Val Lys Pro Thr Val Ala Gln Gln Lys Tyr Ile Leu Ser Lys
465                 470                 475                 480

Leu Asn Ala Ala Glu Ala Phe Glu Thr Phe Leu Gln Thr Lys Tyr Val
                485                 490                 495

Gly Gln Lys Arg Phe Ser Leu Glu Gly Ala Glu Ser Val Ile Pro Met
            500                 505                 510

Met Asp Ala Ala Ile Asp Gln Cys Ala Glu His Gly Leu Asp Glu Val
        515                 520                 525

Val Ile Gly Met Pro His Arg Gly Arg Leu Asn Val Leu Ala Asn Ile
    530                 535                 540

Val Gly Lys Pro Tyr Ser Gln Ile Phe Thr Glu Phe Glu Gly Asn Leu
545                 550                 555                 560

Asn Pro Ser Gln Ala His Gly Ser Gly Asp Val Lys Tyr His Leu Gly
                565                 570                 575

Ala Thr Gly Leu Tyr Leu Gln Met Phe Gly Asp Asn Asp Ile Gln Val
            580                 585                 590

Ser Leu Thr Ala Asn Pro Ser His Leu Glu Ala Val Asp Pro Val Leu
        595                 600                 605

Glu Gly Leu Val Arg Ala Lys Gln Asp Leu Leu Asp His Gly Ser Ile
    610                 615                 620

Asp Ser Asp Gly Gln Arg Ala Phe Ser Val Val Pro Leu Met Leu His
625                 630                 635                 640

Gly Asp Ala Ala Phe Ala Gly Gln Gly Val Val Ala Glu Thr Leu Asn
                645                 650                 655

Leu Ala Asn Leu Pro Gly Tyr Arg Val Gly Gly Thr Ile His Ile Ile
            660                 665                 670

Val Asn Asn Gln Ile Gly Phe Thr Thr Ala Pro Glu Tyr Ser Arg Ser
        675                 680                 685

Ser Glu Tyr Cys Thr Asp Val Ala Lys Met Ile Gly Ala Pro Ile Phe
    690                 695                 700

His Val Asn Gly Asp Asp Pro Glu Ala Cys Val Trp Val Ala Arg Leu
705                 710                 715                 720

Ala Val Asp Phe Arg Gln Arg Phe Lys Lys Asp Val Val Ile Asp Met
```

-continued

```
                725                 730                 735

Leu Cys Tyr Arg Arg Arg Gly His Asn Glu Gly Asp Asp Pro Ser Met
            740                 745                 750

Thr Asn Pro Tyr Val Tyr Asp Val Val Asp Thr Lys Arg Gly Ala Arg
        755                 760                 765

Lys Ser Tyr Thr Glu Ala Leu Ile Gly Arg Gly Asp Ile Ser Met Lys
    770                 775                 780

Glu Ala Glu Asp Ala Leu Arg Asp Tyr Gln Gly Gln Leu Glu Arg Val
785                 790                 795                 800

Phe Asn Glu Val Arg Glu Leu Glu Lys His Gly Val Gln Pro Ser Glu
                805                 810                 815

Ser Val Glu Ser Asp Gln Met Ile Pro Ala Gly Leu Ala Thr Ala Val
            820                 825                 830

Asp Lys Ser Leu Leu Ala Arg Ile Gly Asp Ala Phe Leu Ala Leu Pro
        835                 840                 845

Asn Gly Phe Thr Ala His Pro Arg Val Gln Pro Val Leu Glu Lys Arg
    850                 855                 860

Arg Glu Met Ala Tyr Glu Gly Lys Ile Asp Trp Ala Phe Gly Glu Leu
865                 870                 875                 880

Leu Ala Leu Gly Ser Leu Val Ala Glu Gly Lys Leu Val Arg Leu Ser
                885                 890                 895

Gly Gln Asp Ser Arg Arg Gly Thr Phe Ser Gln Arg His Ser Val Leu
            900                 905                 910

Ile Asp Arg His Thr Gly Glu Glu Phe Thr Pro Leu Gln Leu Leu Ala
        915                 920                 925

Thr Asn Ser Asp Gly Ser Pro Thr Gly Gly Lys Phe Leu Val Tyr Asp
    930                 935                 940

Ser Pro Leu Ser Glu Tyr Ala Ala Val Gly Phe Glu Tyr Gly Tyr Thr
945                 950                 955                 960

Val Gly Asn Pro Asp Ala Val Val Leu Trp Glu Ala Gln Phe Gly Asp
                965                 970                 975

Phe Val Asn Gly Ala Gln Ser Ile Ile Asp Glu Phe Ile Ser Ser Gly
            980                 985                 990

Glu Ala Lys Trp Gly Gln Leu Ser  Asn Val Val Leu Leu  Leu Pro His
        995                 1000                1005

Gly His  Glu Gly Gln Gly Pro  Asp His Thr Ser Ala  Arg Ile Glu
    1010                1015                1020

Arg Phe  Leu Gln Leu Trp Ala  Glu Gly Ser Met Thr  Ile Ala Met
    1025                1030                1035

Pro Ser  Thr Pro Ser Asn Tyr  Phe His Leu Leu Arg  Arg His Ala
    1040                1045                1050

Leu Asp  Gly Ile Gln Arg Pro  Leu Ile Val Phe Thr  Pro Lys Ser
    1055                1060                1065

Met Leu  Arg His Lys Ala Ala  Val Ser Glu Ile Lys  Asp Phe Thr
    1070                1075                1080

Glu Ile  Lys Phe Arg Ser Val  Leu Glu Glu Pro Thr  Tyr Glu Asp
    1085                1090                1095

Gly Ile  Gly Asp Arg Asn Lys  Val Ser Arg Ile Leu  Leu Thr Ser
    1100                1105                1110

Gly Lys  Leu Tyr Tyr Glu Leu  Ala Ala Arg Lys Ala  Lys Asp Asn
    1115                1120                1125
```

-continued

```
Arg Asn  Asp Leu Ala Ile Val  Arg Leu Glu Gln Leu  Ala Pro Leu
    1130                 1135                 1140

Pro Arg  Arg Arg Leu Arg Glu  Thr Leu Asp Arg Tyr  Glu Asn Val
    1145                 1150                 1155

Lys Glu  Phe Phe Trp Val Gln  Glu Glu Pro Ala Asn  Gln Gly Ala
    1160                 1165                 1170

Trp Pro  Arg Phe Gly Leu Glu  Leu Pro Glu Leu Leu  Pro Asp Lys
    1175                 1180                 1185

Leu Ala  Gly Ile Lys Arg Ile  Ser Arg Arg Ala Met  Ser Ala Pro
    1190                 1195                 1200

Ser Ser  Gly Ser Ser Lys Val  His Ala Val Glu Gln  Gln Glu Ile
    1205                 1210                 1215

Leu Asp  Glu Ala Phe Gly
    1220
```

<210> SEQ ID NO 3
<211> LENGTH: 8621
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      p206 plasmid with codon optimized kgd gene nucleic
      acid sequence and softag

<400> SEQUENCE: 3

```
ggtggcggta cttgggtcga tatcaaagtg catcacttct tcccgtatgc ccaactttgt     60

atagagagcc actgcgggat cgtcaccgta atctgcttgc acgtagatca cataagcacc    120

aagcgcgttg gcctcatgct tgaggagatt gatgagcgcg gtggcaatgc cctgcctccg    180

gtgctcgccg gagactgcga gatcatagat atagatctca ctacgcggct gctcaaactt    240

gggcagaacg taagccgcga gagcgccaac aaccgcttct tggtcgaagg cagcaagcgc    300

gatgaatgtc ttactacgga gcaagttccc gaggtaatcg gagtccggct gatgttggga    360

gtaggtggct acgtcaccga actcacgacc gaaaagatca agagcagccc gcatggattt    420

gacttggtca gggccgagcc tacatgtgcg aatgatgccc atacttgagc cacctaactt    480

tgttttaggg cgactgccct gctgcgtaac atcgttgctg ctccataaca tcaaacatcg    540

acccacggcg taacgcgctt gctgcttgga tgcccgaggc atagactgta caaaaaaaca    600

gtcataacaa gccatgaaaa ccgccactgc gccgttacca ccgctgcgtt cggtcaaggt    660

tctggaccag ttgcgtgagc gcattttttt ttcctcctcg gcgtttacgc cccgccctgc    720

cactcatcgc agtactgttg taattcatta agcattctgc cgacatggaa gccatcacag    780

acggcatgat gaacctgaat cgccagcggc atcagcacct tgtcgccttg cgtataatat    840

ttgcccatag tgaaaacggg ggcgaagaag ttgtccatat tggccacgtt taaatcaaaa    900

ctggtgaaac tcacccaggg attggcgctg acgaaaaaca tattctcaat aaacccttta    960

gggaaatagg ccaggttttc accgtaacac gccacatctt gcgaatatat gtgtagaaac   1020

tgccggaaat cgtcgtggta ttcactccag agcgatgaaa acgtttcagt ttgctcatgg   1080

aaaacggtgt aacaagggtg aacactatcc catatcacca gctcaccgtc tttcattgcc   1140

atacggaact ccggatgagc attcatcagg cgggcaagaa tgtgaataaa ggccggataa   1200

aacttgtgct tatttttctt tacggtcttt aaaaaggccg taatatccag ctgaacggtc   1260

tggttatagg tacattgagc aactgactga aatgcctcaa aatgttcttt acgatgccat   1320
```

-continued

```
tgggatatat caacggtggt atatccagtg attttttttct ccattttttt ttcctccttt   1380
agaaaaactc atcgagcatc aaatgaaact gcaatttatt catatcagga ttatcaatac   1440
catatttttg aaaaagccgt ttctgtaatg aaggagaaaa ctcaccgagg cagttccata   1500
ggatggcaag atcctggtat cggtctgcga ttccgactcg tccaacatca atacaaccta   1560
ttaatttccc ctcgtcaaaa ataaggttat caagtgagaa atcaccatga gtgacgactg   1620
aatccggtga gaatggcaaa agtttatgca tttctttcca gacttgttca acaggccagc   1680
cattacgctc gtcatcaaaa tcactcgcat caaccaaacc gttattcatt cgtgattgcg   1740
cctgagcgag gcgaaatacg cgatcgctgt taaaaggaca attacaaaca ggaatcgagt   1800
gcaaccggcg caggaacact gccagcgcat caacaatatt ttcacctgaa tcaggatatt   1860
cttctaatac ctggaacgct gtttttccgg ggatcgcagt ggtgagtaac catgcatcat   1920
caggagtacg gataaaatgc ttgatggtcg gaagtggcat aaattccgtc agccagttta   1980
gtctgaccat ctcatctgta acatcattgg caacgctacc tttgccatgt ttcagaaaca   2040
actctggcgc atcgggcttc ccatacaagc gatagattgt cgcacctgat tgcccgacat   2100
tatcgcgagc ccatttatac ccatataaat cagcatccat gttggaattt aatcgcggcc   2160
tcgacgtttc ccgttgaata tggctcattt tttttcctc ctttaccaat gcttaatcag   2220
tgaggcacct atctcagcga tctgtctatt tcgttcatcc atagttgcct gactccccgt   2280
cgtgtagata actacgatac gggagggctt accatctggc cccagcgctg cgatgatacc   2340
gcgagaacca cgctcaccgg ctccggattt atcagcaata aaccagccag ccggaagggc   2400
cgagcgcaga agtggtcctg caactttatc cgcctccatc cagtctatta attgttgccg   2460
ggaagctaga gtaagtagtt cgccagttaa tagtttgcgc aacgttgttg ccatcgctac   2520
aggcatcgtg gtgtcacgct cgtcgtttgg tatggcttca ttcagctccg gttcccaacg   2580
atcaaggcga gttacatgat cccccatgtt gtgcaaaaaa gcggttagct ccttcggtcc   2640
tccgatcgtt gtcagaagta agttggccgc agtgttatca ctcatggtta tggcagcact   2700
gcataattct cttactgtca tgccatccgt aagatgcttt tctgtgactg gtgagtactc   2760
aaccaagtca ttctgagaat agtgtatgcg gcgaccgagt tgctcttgcc cggcgtcaat   2820
acgggataat accgcgccac atagcagaac tttaaaagtg ctcatcattg gaaaacgttc   2880
ttcggggcga aaactctcaa ggatcttacc gctgttgaga tccagttcga tgtaacccac   2940
tcgtgcaccc aactgatctt cagcatcttt tactttcacc agcgtttctg ggtgagcaaa   3000
aacaggaagg caaaatgccg caaaaaaggg aataagggcg acacggaaat gttgaatact   3060
catattcttc ctttttcaat attattgaag catttatcag ggttattgtc tcatgagcgg   3120
atacatattt gaatgtattt agaaaaataa acaaataggg gtcagtgtta caaccaatta   3180
accaattctg aacattatcg cgagcccatt tatacctgaa tatggctcat aacacccctt   3240
gtttgcctgg cggcagtagc gcggtggtcc cacctgaccc catgccgaac tcagaagtga   3300
aacgccgtag cgccgatggt agtgtgggga ctccccatgc gagagtaggg aactgccagg   3360
catcaaataa aacgaaaggc tcagtcgaaa gactgggcct ttcgcccggg ctaattgagg   3420
ggtgtcgccc ttttgacgga tatcaagctt ctattaaccg aacgcttcgt ccaagatttc   3480
ttgctgctcc acggcatgaa ccttcgagga accgctgctc ggcgcagaca tcgcgcgacg   3540
ggagatacgc ttgatgcccg ccaacttgtc cggcagcaac tccggcaact ccaaaccgaa   3600
acgcggccag gcgccctggt tcgctggttc ctcctggacc caaaagaact ctttgacatt   3660
```

-continued

```
ttcgtaacga tccagggttt cacgcagacg acgacgcggc aacggcgcca gttgctccag   3720
acgcacaatt gccagatcat tacggttgtc tttcgccttg cgtgccgcca attcataata   3780
caacttaccg gaggtcaaca gaatacggct aactttattg cggtcgccga taccatcctc   3840
gtaggtcggt tcctccagga cgctacggaa tttaatctcg gtaaaatcct taatctccga   3900
aaccgccgcc ttgtggcgca gcatggattt cggggtaaac acgatcaacg gacgttggat   3960
gccgtccaac gcgtggcggc gcagcaagtg aaagtaattg gacggggtgc tcggcatcgc   4020
gatcgtcatg ctaccctcag cccacagctg caaaaagcgc tcaatgcgcg ccgacgtgtg   4080
gtccgggccc tgaccctcgt ggccgtgtgg cagcaacaga acaacgttgg acagctgacc   4140
ccatttcgct tcgcccgagc tgatgaactc gtcaatgatc gactgggcac cattaacaaa   4200
gtcgccgaat tgcgcctccc acagcacaac cgcgtctgga ttgccaacgg tataaccgta   4260
ctcaaaaccc acagccgcgt attccgacaa cggcgaatca taaaccagga acttaccacc   4320
ggtcgggctg ccgtcgctgt tagtcgccag cagctgcagc ggggtgaact cctcgccggt   4380
gtgacggtcg atcagcacag aatgacgctg gctaaaggtg ccacgacgag agtcttgacc   4440
gctcagacgg accagcttgc cctcagcaac caggctgccc agcgccagca actcaccaaa   4500
cgcccagtca atcttaccct catacgccat ctcacgacgc ttttccagca ccggctgaac   4560
gcgcgggtgt gccgtgaaac cgttaggcag cgccaggaag gcatcaccga tacgtgccag   4620
cagggatttg tcaacagcgg tcgccaggcc cgctgggatc atttggtccg actcgacgct   4680
ctccgacggt tggacgccgt gcttctccag ttcgcgcact tcgttgaaca cacgctccag   4740
ttggccctgg taatcgcgca gcgcatcctc cgcctctttc atgctgatgt cgccacgacc   4800
gatcagagcc tcggtgtagg atttacgggc gccacgcttg gtgtccacga cgtcatagac   4860
atacggattg gtcatagatg gatcgtcacc ctcattatga ccacgacgac gatagcacag   4920
catgtcaata acaacgtcct ttttaaagcg ttggcggaag tcaactgcca aacgcgcaac   4980
ccagacacac gcttctggat cgtcgccgtt cacgtggaag attggcgcac cgatcatctt   5040
cgccacatcc gtgcagtact cgctcgaacg agagtattcc ggagcggtgg tgaagccgat   5100
ttggttgttc acaatgatgt gaatcgtacc acccacacga tagccaggca gatttgccag   5160
attcagcgtc tccgcaacaa cgccttgacc cgcgaaggcc gcatcaccgt gcagcatcaa   5220
cggaacgacg gagaatgcac gttggccgtc ggagtcaatg gaaccgtggt ccagcaggtc   5280
ttgcttcgca cgcaccaaac cttccagcac tggatcgacc gcttccaaat gggacggatt   5340
tgccgtcagg gaaacctgaa tatcgttatc gccaaacatt tgcagataca gacccgtcgc   5400
acccaggtgg tacttgacat cgccggaacc gtgagcctgg gacgggttca gattgccttc   5460
aaactccgta aagatttgcg aatacggttt gcccacgatg ttcgccagga cattcaagcg   5520
accacggtgc ggcatgccaa tgaccacttc atccaaacca tgctcggcac attggtcaat   5580
cgccgcgtcc atcattggaa taacagattc cgcaccctcc aggctaaaac gcttttggcc   5640
cacatacttg gtttgcagga aggtttcaaa cgcctccgct gcgttcagtt tcgacaagat   5700
gtacttttgt tgagcaacgg tcggtttgac gtgcttcgtc tcgacacgct gctccagcca   5760
ctccttttgt tccgggtcca gaatgtgcgc gtactcaaca ccgatgtgac ggcagtacgc   5820
gtcgcgcagc aaacccagca cgtcacgcag ctttttgtat tgagcacccg cgaaaccgtc   5880
aaccttaaag acgcggtcca ggtcccacag agtcaggcca tgcgtcaaca cctccaaatc   5940
```

-continued

```
cggatgcgaa cgaaagcgcg ccttatccaa gcgcaacggg tcggtgtccg ccatcagatg   6000
gccgcggttg cgataggccg cgatcaggtt catcacacgt gcgttcttgt caacgatcga   6060
gtccggatta tcggtgctcc aacgcactgg caggtacggg atgctcagct cgcggaagac   6120
ctcatcccag aagccatcag acaacagcag ttcatgaatg gtacgcagga agtcaccgct   6180
ttccgcacct tgaatgatac ggtggtcgta ggtagaagtc agggtaatca gtttgccaat   6240
acccagctca gcgatgcgtt cctcggacgc gccttggaac tccgccggat attccatcgc   6300
accgacaccg atgatagcac cttgacctgg catcaggcgt ggcacagagt gcaccgtgcc   6360
aatcgtgccc ggattcgtca gcgaaatcgt aacgccagcg aagtcctcgg tggtcagttt   6420
accatcacga gcacgacgga cgatgtcctc gtacgcggta acgaactgcg cgaagcgcat   6480
cgtctcgcaa cgtttgatac cggccaccac cagagagcgc ttgccatctt taccttgcag   6540
gtcaatagcc agacccagat tggtgtgcgc aggagtaacc gccgtcggct taccgtccac   6600
ctccgtgtag tggcgattca tattcgggaa cttcttaacc gcctgaacca gagcataacc   6660
cagcaaatgg gtaaagctga ttttaccacc acgcgtgcgt ttcaactgat tattgatcac   6720
gatacgatta tcgatcaaca atttcgctgg cacagcacgc accgaggtcg ccgtaggcac   6780
ttccagcgac gcgctcatgt tcttcacgac agcagccgcc gcaccacgca ggacagcaac   6840
ttcatcgcct tcggctggcg gaggcaccgc ggtcttggcg gccagagccg ccacgacacc   6900
gttgcccgct gcggccgtgt cggccggttt cggcggcgct tgcggagccg ccgctgccgc   6960
acgctcagcg accaaagggc tggtcacacg agtcggctca gcagccggtt gggaagtcgg   7020
ctccgggcta tagtccacca gaaactcatg ccagcttggg tcaacggaag acggatcatc   7080
acgaaattta cgatacatac caacacgaga cgggtcctga gtcaccatgg atatatctcc   7140
ttcttaaaga attcgatatc tcagcgacaa gggcgacaca aaatttattc taaatgcata   7200
ataaatactg ataacatctt atagtttgta ttatattttg tattatcgtt gacatgtata   7260
attttgatat caaaaactga ttttcccttt attattttcg agatttattt tcttaattct   7320
ctttaacaaa ctagaaatat tgtatataca aaaaatcata aataatagat gaatagttta   7380
attataggtg ttcatcaatc gaaaaagcaa cgtatcttat ttaaagtgcg ttgctttttt   7440
ctcatttata aggttaaata attctcatat atcaagcaaa gtgacaggcg cccttaaata   7500
ttctgacaaa tgctctttcc ctaaactccc cccataaaaa aacccgccga agcgggtttt   7560
tacgttattt gcggattaac gattactcgt tatcagaacc gcccaggggg cccgagctta   7620
agactggccg tcgttttaca acacagaaag agtttgtaga aacgcaaaaa ggccatccgt   7680
caggggcctt ctgcttagtt tgatgcctgg cagttcccta ctctcgcctt ccgcttcctc   7740
gctcactgac tcgctgcgct cggtcgttcg gctgcggcga gcggtatcag ctcactcaaa   7800
ggcggtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa   7860
aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggct   7920
ccgcccccct gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaacccgac   7980
aggactataa agataccagg cgtttccccc tggaagctcc ctcgtgcgct ctcctgttcc   8040
gaccctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg tggcgctttc   8100
tcatagctca cgctgtaggt atctcagttc ggtgtaggtc gttcgctcca agctgggctg   8160
tgtgcacgaa ccccccgttc agcccgaccg ctgcgcctta tccggtaact atcgtcttga   8220
gtccaacccg gtaagacacg acttatcgcc actggcagca gccactggta acaggattag   8280
```

-continued

```
cagagcgagg tatgtaggcg gtgctacaga gttcttgaag tggtgggcta actacggcta   8340

cactagaaga acagtatttg gtatctgcgc tctgctgaag ccagttacct tcggaaaaag   8400

agttggtagc tcttgatccg gcaaacaaac caccgctggt agcggtggtt tttttgtttg   8460

caagcagcag attacgcgca gaaaaaaagg atctcaagaa gatcctttga tcttttctac   8520

ggggtctgac gctcagtgga acgacgcgcg cgtaactcac gttaagggat tttggtcatg   8580

agcttgcgcc gtcccgtcaa gtcagcgtaa tgctctgctt a                       8621


<210> SEQ ID NO 4
<211> LENGTH: 39
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      Forward primer for PCR amplification of codon optimized kgd
      nucleic acid sequence

<400> SEQUENCE: 4

ttttttgta taccatggat cgtaaatttc gtgatgatc                            39


<210> SEQ ID NO 5
<211> LENGTH: 42
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      Reverse primer for PCR amplification of codon optimized kgd
      nucleic acid sequence

<400> SEQUENCE: 5

cccggtgaga tctagatccg aacgcttcgt ccaagatttc tt                       42


<210> SEQ ID NO 6
<211> LENGTH: 47
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      Forward primer for PCR amplification of pKK223-3 template

<400> SEQUENCE: 6

cggatctaga tctcaccatc accaccatta gtcgacctgc agccaag                  47


<210> SEQ ID NO 7
<211> LENGTH: 44
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      Reverse primer for PCR amplification of pKK223-3 template

<400> SEQUENCE: 7

tgagatctag atccgttatg tcccatggtt ctgtttcctg tgtg                     44


<210> SEQ ID NO 8
<211> LENGTH: 4614
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      pKK223-ct-his vector sequence

<400> SEQUENCE: 8
```

-continued

```
ttgacaatta atcatcggct cgtataatgt gtggaattgt gagcggataa caatttcaca      60
caggaaacag aaccatggga cataacggat ctagatctca ccatcaccac cattagtcga     120
cctgcagcca agcttggctg ttttggcgga tgagagaaga ttttcagcct gatacagatt     180
aaatcagaac gcagaagcgg tctgataaaa cagaatttgc ctggcggcag tagcgcggtg     240
gtcccacctg accccatgcc gaactcagaa gtgaaacgcc gtagcgccga tggtagtgtg     300
gggtctcccc atgcgagagt agggaactgc caggcatcaa ataaaacgaa aggctcagtc     360
gaaagactgg gcctttcgtt ttatctgttg tttgtcggtg aacgctctcc tgagtaggac     420
aaatccgccg ggagcggatt tgaacgttgc gaagcaacgg cccggagggt ggcgggcagg     480
acgcccgcca taaactgcca ggcatcaaat taagcagaag gccatcctga cggatggcct     540
ttttgcgttt ctacaaactc ttttgtttat ttttctaaat acattcaaat atgtatccgc     600
tcatgagaca ataaccctga taaatgcttc aataatattg aaaaaggaag agtatgagta     660
ttcaacattt ccgtgtcgcc cttattccct tttttgcggc attttgcctt cctgtttttg     720
ctcacccaga aacgctggtg aaagtaaaag atgctgaaga tcagttgggt gcacgagtgg     780
gttacatcga actggatctc aacagcggta agatccttga gagttttcgc cccgaagaac     840
gttttccaat gatgagcact tttaaagttc tgctatgtgg cgcggtatta tcccgtgttg     900
acgccgggca agagcaactc ggtcgccgca tacactattc tcagaatgac ttggttgagt     960
actcaccagt cacagaaaag catcttacgg atggcatgac agtaagagaa ttatgcagtg    1020
ctgccataac catgagtgat aacactgcgg ccaacttact tctgacaacg atcggaggac    1080
cgaaggagct aaccgctttt ttgcacaaca tgggggatca tgtaactcgc cttgatcgtt    1140
gggaaccgga gctgaatgaa gccataccaa acgacgagcg tgacaccacg atgctgtagc    1200
aatggcaaca acgttgcgca aactattaac tggcgaacta cttactctag cttcccggca    1260
acaattaata gactggatgg aggcggataa agttgcagga ccacttctgc gctcggccct    1320
tccggctggc tggtttattg ctgataaatc tggagccggt gagcgtgggt ctcgcggtat    1380
cattgcagca ctggggccag atggtaagcc ctcccgtatc gtagttatct acacgacggg    1440
gagtcaggca actatggatg aacgaaatag acagatcgct gagataggtg cctcactgat    1500
taagcattgg taactgtcag accaagttta ctcatatata ctttagattg atttaaaact    1560
tcatttttaa tttaaaagga tctaggtgaa gatccttttt gataatctca tgaccaaaat    1620
cccttaacgt gagttttcgt tccactgagc gtcagacccc gtagaaaaga tcaaaggatc    1680
ttcttgagat cctttttttc tgcgcgtaat ctgctgcttg caaacaaaaa aaccaccgct    1740
accagcggtg gtttgtttgc cggatcaaga gctaccaact ctttttccga aggtaactgg    1800
cttcagcaga gcgcagatac caaatactgt ccttctagtg tagccgtagt taggccacca    1860
cttcaagaac tctgtagcac cgcctacata cctcgctctg ctaatcctgt taccagtggc    1920
tgctgccagt ggcgataagt cgtgtcttac cgggttggac tcaagacgat agttaccgga    1980
taaggcgcag cggtcgggct gaacgggggg ttcgtgcaca cagcccagct tggagcgaac    2040
gacctacacc gaactgagat acctacagcg tgagcattga gaaagcgcca cgcttcccga    2100
agggagaaag gcggacaggt atccggtaag cggcagggtc ggaacaggag agcgcacgag    2160
ggagcttcca gggggaaacg cctggtatct ttatagtcct gtcgggtttc gccacctctg    2220
acttgagcgt cgatttttgt gatgctcgtc aggggggcgg agcctatgga aaaacgccag    2280
```

-continued

```
caacgcggcc tttttacggt tcctggcctt ttgctggcct tttgctcaca tgttctttcc   2340
tgcgttatcc cctgattctg tggataaccg tattaccgcc tttgagtgag ctgataccgc   2400
tcgccgcagc cgaacgaccg agcgcagcga gtcagtgagc gaggaagcgg aagagcgcct   2460
gatgcggtat tttctcctta cgcatctgtg cggtatttca caccgcatat ggtgcactct   2520
cagtacaatc tgctctgatg ccgcatagtt aagccagtat acactccgct atcgctacgt   2580
gactgggtca tggctgcgcc ccgacacccg ccaacacccg ctgacgcgcc ctgacgggct   2640
tgtctgctcc cggcatccgc ttacagacaa gctgtgaccg tctccgggag ctgcatgtgt   2700
cagaggtttt caccgtcatc accgaaacgc gcgaggcagc tgcggtaaag ctcatcagcg   2760
tggtcgtgaa gcgattcaca gatgtctgcc tgttcatccg cgtccagctc gttgagtttc   2820
tccagaagcg ttaatgtctg gcttctgata aagcgggcca tgttaagggc ggtttttcc    2880
tgtttggtca ctgatgcctc cgtgtaaggg ggatttctgt tcatgggggt aatgataccg   2940
atgaaacgag agaggatgct cacgatacgg gttactgatg atgaacatgc ccggttactg   3000
gaacgttgtg agggtaaaca actggcggta tggatgcggc gggaccagag aaaaatcact   3060
cagggtcaat gccagcgctt cgttaataca gatgtaggtg ttccacaggg tagccagcag   3120
catcctgcga tgcagatccg gaacataatg gtgcagggcg ctgacttccg cgtttccaga   3180
ctttacgaaa cacggaaacc gaagaccatt catgttgttg ctcaggtcgc agacgttttg   3240
cagcagcagt cgcttcacgt tcgctcgcgt atcggtgatt cattctgcta accagtaagg   3300
caaccccgcc agcctagccg ggtcctcaac gacaggagca cgatcatgcg cacccgtggc   3360
caggacccaa cgctgcccga gatgcgccgc gtgcggctgc tggagatggc ggacgcgatg   3420
gatatgttct gccaagggtt ggtttgcgca ttcacagttc tccgcaagaa ttgattggct   3480
ccaattcttg gagtggtgaa tccgttagcg aggtgccgcc ggcttccatt caggtcgagg   3540
tggcccggct ccatgcaccg cgacgcaacg cggggaggca gacaaggtat agggcggcgc   3600
ctacaatcca tgccaacccg ttccatgtgc tcgccgaggc ggcataaatc gccgtgacga   3660
tcagcggtcc agtgatcgaa gttaggctgg taagagccgc gagcgatcct tgaagctgtc   3720
cctgatggtc gtcatctacc tgcctggaca gcatggcctg caacgcgggc atcccgatgc   3780
cgccggaagc gagaagaatc ataatgggga aggccatcca gcctcgcgtc gcgaacgcca   3840
gcaagacgta gcccagcgcg tcggccgcca tgccggcgat aatggcctgc ttctcgccga   3900
aacgtttggt ggcgggacca gtgacgaagg cttgagcgag ggcgtgcaag attccgaata   3960
ccgcaagcga caggccgatc atcgtcgcgc tccagcgaaa gcggtcctcg ccgaaaatga   4020
cccagagcgc tgccggcacc tgtcctacga gttgcatgat aaagaagaca gtcataagtg   4080
cggcgacgat agtcatgccc cgcgcccacc ggaaggagct gactgggttg aaggctctca   4140
agggcatcgg tcgacgctct cccttatgcg actcctgcat taggaagcag cccagtagta   4200
ggttgaggcc gttgagcacc gccgccgcaa ggaatggtgc atgcaaggag atggcgccca   4260
acagtccccc ggccacgggg cctgccacca tacccacgcc gaaacaagcg ctcatgagcc   4320
cgaagtggcg agcccgatct tccccatcgg tgatgtcggc gatataggcg ccagcaaccg   4380
cacctgtggc gccggtgatg ccggccacga tgcgtccggc gtagaggatc cgggcttatc   4440
gactgcacgg tgcaccaatg cttctggcgt caggcagcca tcggaagctg tggtatggct   4500
gtgcaggtcg taaatcactg cataattcgt gtcgctcaag gcgcactccc gttctggata   4560
atgtttttg cgccgacatc ataacggttc tggcaaatat tctgaaatga gctg          4614
```

-continued

<210> SEQ ID NO 9
<211> LENGTH: 8
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
Histidine tag peptide sequence

<400> SEQUENCE: 9

Ser Arg Ser His His His His His
1   5

<210> SEQ ID NO 10
<211> LENGTH: 8241
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
pKK223-kgd plasmid with codon optimized kgd gene nucleic acid
sequence

<400> SEQUENCE: 10

ttgacaatta atcatcggct cgtataatgt gtggaattgt gagcggataa caatttcaca 60

caggaaacag aaccatggat cgtaaatttc gtgatgatcc gtcttccgtt gacccaagct 120

ggcatgagtt tctggtggac tatagcccgg agccgacttc ccaaccggct gctgagccga 180

ctcgtgtgac cagccctttg gtcgctgagc gtgcggcagc ggcggctccg caagcgccgc 240

cgaaaccggc cgacacggcc gcagcgggca acggtgtcgt ggcggctctg gccgccaaga 300

ccgcggtgcc tccgccagcc gaaggcgatg aagttgctgt cctgcgtggt gcggcggctg 360

ctgtcgtgaa gaacatgagc gcgtcgctgg aagtgcctac ggcgacctcg gtgcgtgctg 420

tgccagcgaa attgttgatc gataatcgta tcgtgatcaa taatcagttg aaacgcacgc 480

gtggtggtaa aatcagcttt acccatttgc tgggttatgc tctggttcag gcggttaaga 540

agttcccgaa tatgaatcgc cactacacgg aggtggacgg taagccgacg gcggttactc 600

ctgcgcacac caatctgggt ctggctattg acctgcaagg taaagatggc aagcgctctc 660

tggtggtggc cggtatcaaa cgttgcgaga cgatgcgctt cgcgcagttc gttaccgcgt 720

acgaggacat cgtccgtcgt gctcgtgatg gtaaactgac caccgaggac ttcgctggcg 780

ttacgatttc gctgacgaat ccgggcacga ttggcacggt gcactctgtg ccacgcctga 840

tgccaggtca aggtgctatc atcggtgtcg gtgcgatgga atatccggcg gagttccaag 900

gcgcgtccga ggaacgcatc gctgagctgg gtattggcaa actgattacc ctgacttcta 960

cctacgacca ccgtatcatt caaggtgcgg aaagcggtga cttcctgcgt accattcatg 1020

aactgctgtt gtctgatggc ttctgggatg aggtcttccg cgagctgagc atcccgtacc 1080

tgccagtgcg ttggagcacc gataatccgg actcgatcgt tgacaagaac gcacgtgtga 1140

tgaacctgat cgcggcctat cgcaaccgcg gccatctgat ggcggacacc gacccgttgc 1200

gcttggataa ggcgcgcttt cgttcgcatc cggatttgga ggtgttgacg catggcctga 1260

ctctgtggga cctggaccgc gtctttaagg ttgacggttt cgcgggtgct caatacaaaa 1320

agctgcgtga cgtgctgggt ttgctgcgcg acgcgtactg ccgtcacatc ggtgttgagt 1380

acgcgcacat tctggacccg gaacaaaagg agtggctgga gcagcgtgtc gagacgaagc 1440

acgtcaaacc gaccgttgct caacaaaagt acatcttgtc gaaactgaac gcagcggagg 1500

-continued

```
cgtttgaaac cttcctgcaa accaagtatg tgggccaaaa gcgttttagc ctggagggtg   1560
cggaatctgt tattccaatg atggacgcgg cgattgacca atgtgccgag catggtttgg   1620
atgaagtggt cattggcatg ccgcaccgtg gtcgcttgaa tgtcctggcg aacatcgtgg   1680
gcaaaccgta ttcgcaaatc tttacggagt ttgaaggcaa tctgaacccg tcccaggctc   1740
acggttccgg cgatgtcaag taccacctgg gtgcgacggg tctgtatctg caaatgtttg   1800
gcgataacga tattcaggtt tccctgacgg caaatccgtc ccatttggaa gcggtcgatc   1860
cagtgctgga aggtttggtg cgtgcgaagc aagacctgct ggaccacggt tccattgact   1920
ccgacggcca acgtgcattc tccgtcgttc cgttgatgct gcacggtgat gcggccttcg   1980
cgggtcaagg cgttgttgcg gagacgctga atctggcaaa tctgcctggc tatcgtgtgg   2040
gtggtacgat tcacatcatt gtgaacaacc aaatcggctt caccaccgct ccggaatact   2100
ctcgttcgag cgagtactgc acggatgtgg cgaagatgat cggtgcgcca atcttccacg   2160
tgaacggcga cgatccagaa gcgtgtgtct gggttgcgcg tttggcagtt gacttccgcc   2220
aacgctttaa aaaggacgtt gttattgaca tgctgtgcta tcgtcgtcgt ggtcataatg   2280
agggtgacga tccatctatg accaatccgt atgtctatga cgtcgtggac accaagcgtg   2340
gcgcccgtaa atcctacacc gaggctctga tcggtcgtgg cgacatcagc atgaaagagg   2400
cggaggatgc gctgcgcgat taccagggcc aactggagcg tgtgttcaac gaagtgcgcg   2460
aactggagaa gcacggcgtc caaccgtcgg agagcgtcga gtcggaccaa atgatcccag   2520
cgggcctggc gaccgctgtt gacaaatccc tgctggcacg tatcggtgat gccttcctgg   2580
cgctgcctaa cggtttcacg gcacacccgc gcgttcagcc ggtgctggaa aagcgtcgtg   2640
agatggcgta tgagggtaag attgactggg cgtttggtga gttgctggcg ctgggcagcc   2700
tggttgctga gggcaagctg gtccgtctga gcggtcaaga ctctcgtcgt ggcaccttta   2760
gccagcgtca ttctgtgctg atcgaccgtc acaccggcga ggagttcacc ccgctgcagc   2820
tgctggcgac taacagcgac ggcagcccga ccggtggtaa gttcctggtt tatgattcgc   2880
cgttgtcgga atacgcggct gtgggttttg agtacggtta taccgttggc aatccagacg   2940
cggttgtgct gtgggaggcg caattcggcg actttgttaa tggtgcccag tcgatcattg   3000
acgagttcat cagctcgggc gaagcgaaat ggggtcagct gtccaacgtt gttctgttgc   3060
tgccacacgg ccacgagggt cagggcccgg accacacgtc ggcgcgcatt gagcgctttt   3120
tgcagctgtg ggctgagggt agcatgacga tcgcgatgcc gagcaccccg tccaattact   3180
ttcacttgct gcgccgccac gcgttggacg gcatccaacg tccgttgatc gtgtttaccc   3240
cgaaatccat gctgcgccac aaggcggcgg tttcggagat taaggatttt accgagatta   3300
aattccgtag cgtcctggag gaaccgacct acgaggatgg tatcggcgac cgcaataaag   3360
ttagccgtat tctgttgacc tccggtaagt tgtattatga attggcggca cgcaaggcga   3420
aagacaaccg taatgatctg gcaattgtgc gtctggagca actggcgccg ttgccgcgtc   3480
gtcgtctgcg tgaaaccctg gatcgttacg aaaatgtcaa agagttcttt tgggtccagg   3540
aggaaccagc gaaccagggc gcctggccgc gtttcggttt ggagttgccg gagttgctgc   3600
cggacaagtt ggcgggcatc aagcgtatct cccgtcgcgc gatgtctgcg ccgagcagcg   3660
gttcctcgaa ggttcatgcc gtggagcagc aagaaatctt ggacgaagcg ttcggatcta   3720
gatctcacca tcaccaccat tagtcgacct gcagccaagc ttggctgttt tggcggatga   3780
gagaagattt tcagcctgat acagattaaa tcagaacgca gaagcggtct gataaaacag   3840
```

-continued

```
aatttgcctg gcggcagtag cgcggtggtc ccacctgacc ccatgccgaa ctcagaagtg   3900
aaacgccgta gcgccgatgg tagtgtgggg tctccccatg cgagagtagg gaactgccag   3960
gcatcaaata aaacgaaagg ctcagtcgaa agactgggcc tttcgtttta tctgttgttt   4020
gtcggtgaac gctctcctga gtaggacaaa tccgccggga gcggatttga acgttgcgaa   4080
gcaacggccc ggagggtggc gggcaggacg cccgccataa actgccaggc atcaaattaa   4140
gcagaaggcc atcctgacgg atggcctttt tgcgtttcta caaactcttt tgtttatttt   4200
tctaaataca ttcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat   4260
aatattgaaa aaggaagagt atgagtattc aacatttccg tgtcgccctt attccctttt   4320
ttgcggcatt ttgccttcct gtttttgctc acccagaaac gctggtgaaa gtaaaagatg   4380
ctgaagatca gttgggtgca cgagtgggtt acatcgaact ggatctcaac agcggtaaga   4440
tccttgagag ttttcgcccc gaagaacgtt ttccaatgat gagcactttt aaagttctgc   4500
tatgtggcgc ggtattatcc cgtgttgacg ccgggcaaga gcaactcggt cgccgcatac   4560
actattctca gaatgacttg gttgagtact caccagtcac agaaaagcat cttacggatg   4620
gcatgacagt aagagaatta tgcagtgctg ccataaccat gagtgataac actgcggcca   4680
acttacttct gacaacgatc ggaggaccga aggagctaac cgcttttttg cacaacatgg   4740
gggatcatgt aactcgcctt gatcgttggg aaccggagct gaatgaagcc ataccaaacg   4800
acgagcgtga caccacgatg ctgtagcaat ggcaacaacg ttgcgcaaac tattaactgg   4860
cgaactactt actctagctt cccggcaaca attaatagac tggatggagg cggataaagt   4920
tgcaggacca cttctgcgct cggcccttcc ggctggctgg tttattgctg ataaatctgg   4980
agccggtgag cgtgggtctc gcggtatcat tgcagcactg gggccagatg gtaagccctc   5040
ccgtatcgta gttatctaca cgacggggag tcaggcaact atggatgaac gaaatagaca   5100
gatcgctgag ataggtgcct cactgattaa gcattggtaa ctgtcagacc aagtttactc   5160
atatatactt tagattgatt taaaacttca tttttaattt aaaaggatct aggtgaagat   5220
cctttttgat aatctcatga ccaaaatccc ttaacgtgag ttttcgttcc actgagcgtc   5280
agaccccgta gaaaagatca aaggatcttc ttgagatcct ttttttctgc gcgtaatctg   5340
ctgcttgcaa acaaaaaaac caccgctacc agcggtggtt tgtttgccgg atcaagagct   5400
accaactctt tttccgaagg taactggctt cagcagagcg cagataccaa atactgtcct   5460
tctagtgtag ccgtagttag gccaccactt caagaactct gtagcaccgc ctacatacct   5520
cgctctgcta atcctgttac cagtggctgc tgccagtggc gataagtcgt gtcttaccgg   5580
gttggactca agacgatagt taccggataa ggcgcagcgg tcgggctgaa cggggggttc   5640
gtgcacacag cccagcttgg agcgaacgac ctacaccgaa ctgagatacc tacagcgtga   5700
gcattgagaa agcgccacgc ttcccgaagg gagaaaggcg gacaggtatc cggtaagcgg   5760
cagggtcgga acaggagagc gcacgaggga gcttccaggg ggaaacgcct ggtatcttta   5820
tagtcctgtc gggtttcgcc acctctgact tgagcgtcga tttttgtgat gctcgtcagg   5880
ggggcggagc ctatggaaaa acgccagcaa cgcggccttt ttacggttcc tggccttttg   5940
ctggcctttt gctcacatgt tctttcctgc gttatcccct gattctgtgg ataaccgtat   6000
taccgccttt gagtgagctg ataccgctcg ccgcagccga acgaccgagc gcagcgagtc   6060
agtgagcgag gaagcggaag agcgcctgat gcggtatttt ctccttacgc atctgtgcgg   6120
```

-continued

```
tatttcacac cgcatatggt gcactctcag tacaatctgc tctgatgccg catagttaag   6180
ccagtataca ctccgctatc gctacgtgac tgggtcatgg ctgcgccccg acacccgcca   6240
acacccgctg acgcgccctg acgggcttgt ctgctcccgg catccgctta cagacaagct   6300
gtgaccgtct ccgggagctg catgtgtcag aggttttcac cgtcatcacc gaaacgcgcg   6360
aggcagctgc ggtaaagctc atcagcgtgg tcgtgaagcg attcacagat gtctgcctgt   6420
tcatccgcgt ccagctcgtt gagtttctcc agaagcgtta atgtctggct tctgataaag   6480
cgggccatgt taagggcggt tttttcctgt ttggtcactg atgcctccgt gtaaggggga   6540
tttctgttca tgggggtaat gataccgatg aaacgagaga ggatgctcac gatacgggtt   6600
actgatgatg aacatgcccg gttactggaa cgttgtgagg gtaaacaact ggcggtatgg   6660
atgcggcggg accagagaaa aatcactcag ggtcaatgcc agcgcttcgt taatacagat   6720
gtaggtgttc cacagggtag ccagcagcat cctgcgatgc agatccggaa cataatggtg   6780
cagggcgctg acttccgcgt ttccagactt tacgaaacac ggaaaccgaa gaccattcat   6840
gttgttgctc aggtcgcaga cgttttgcag cagcagtcgc ttcacgttcg ctcgcgtatc   6900
ggtgattcat tctgctaacc agtaaggcaa ccccgccagc ctagccgggt cctcaacgac   6960
aggagcacga tcatgcgcac ccgtggccag gacccaacgc tgcccgagat gcgccgcgtg   7020
cggctgctgg agatggcgga cgcgatggat atgttctgcc aagggttggt ttgcgcattc   7080
acagttctcc gcaagaattg attggctcca attcttggag tggtgaatcc gttagcgagg   7140
tgccgccggc ttccattcag gtcgaggtgg cccggctcca tgcaccgcga cgcaacgcgg   7200
ggaggcagac aaggtatagg gcggcgccta caatccatgc caacccgttc catgtgctcg   7260
ccgaggcggc ataaatcgcc gtgacgatca gcggtccagt gatcgaagtt aggctggtaa   7320
gagccgcgag cgatccttga agctgtccct gatggtcgtc atctacctgc ctggacagca   7380
tggcctgcaa cgcgggcatc ccgatgccgc cggaagcgag aagaatcata atggggaagg   7440
ccatccagcc tcgcgtcgcg aacgccagca agacgtagcc cagcgcgtcg gccgccatgc   7500
cggcgataat ggcctgcttc tcgccgaaac gtttggtggc gggaccagtg acgaaggctt   7560
gagcgagggc gtgcaagatt ccgaataccg caagcgacag gccgatcatc gtcgcgctcc   7620
agcgaaagcg gtcctcgccg aaaatgaccc agagcgctgc cggcacctgt cctacgagtt   7680
gcatgataaa gaagacagtc ataagtgcgg cgacgatagt catgccccgc gcccaccgga   7740
aggagctgac tgggttgaag gctctcaagg gcatcggtcg acgctctccc ttatgcgact   7800
cctgcattag gaagcagccc agtagtaggt tgaggccgtt gagcaccgcc gccgcaagga   7860
atggtgcatg caaggagatg gcgcccaaca gtcccccggc cacggggcct gccaccatac   7920
ccacgccgaa acaagcgctc atgagcccga agtggcgagc ccgatcttcc ccatcggtga   7980
tgtcggcgat ataggcgcca gcaaccgcac ctgtggcgcc ggtgatgccg gccacgatgc   8040
gtccggcgta gaggatccgg gcttatcgac tgcacggtgc accaatgctt ctggcgtcag   8100
gcagccatcg gaagctgtgg tatggctgtg caggtcgtaa atcactgcat aattcgtgtc   8160
gctcaaggcg cactcccgtt ctggataatg ttttttgcgc cgacatcata acggttctgg   8220
caaatattct gaaatgagct g                                             8241
```

<210> SEQ ID NO 11
<211> LENGTH: 3716
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence

-continued

<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
Codon-optimized mcr nucleic acid sequence

<400> SEQUENCE: 11

```
gatatcgaat tccgctagca ggagctaagg aagctaaaat gtccggtacg ggtcgtttgg      60
ctggtaaaat tgcattgatc accggtggtg ctggtaacat tggttccgag ctgacccgcc     120
gttttctggc cgagggtgcg acggttatta tcagcggccg taaccgtgcg aagctgaccg     180
cgctggccga gcgcatgcaa gccgaggccg gcgtgccggc caagcgcatt gatttggagg     240
tgatggatgg ttccgaccct gtggctgtcc gtgccggtat cgaggcaatc gtcgctcgcc     300
acggtcagat tgacattctg gttaacaacg cgggctccgc cggtgcccaa cgtcgcttgg     360
cggaaattcc gctgacggag gcagaattgg gtccgggtgc ggaggagact ttgcacgctt     420
cgatcgcgaa tctgttgggc atgggttggc acctgatgcg tattgcggct ccgcacatgc     480
cagttggctc cgcagttatc aacgtttcga ctattttctc gcgcgcagag tactatggtc     540
gcattccgta cgttaccccg aaggcagcgc tgaacgcttt gtcccagctg gctgcccgcg     600
agctgggcgc tcgtggcatc cgcgttaaca ctattttccc aggtcctatt gagtccgacc     660
gcatccgtac cgtgtttcaa cgtatggatc aactgaaggg tcgcccggag ggcgacaccg     720
cccatcactt tttgaacacc atgcgcctgt gccgcgcaaa cgaccaaggc gctttggaac     780
gccgctttcc gtccgttggc gatgttgctg atgcggctgt gtttctggct tctgctgaga     840
gcgcggcact gtcgggtgag acgattgagg tcacccacgg tatggaactg ccggcgtgta     900
gcgaaacctc cttgttggcg cgtaccgatc tgcgtaccat cgacgcgagc ggtcgcacta     960
ccctgatttg cgctggcgat caaattgaag aagttatggc cctgacgggc atgctgcgta    1020
cgtgcggtag cgaagtgatt atcggcttcc gttctgcggc tgccctggcg caatttgagc    1080
aggcagtgaa tgaatctcgc cgtctggcag gtgcggattt caccccgccg atcgctttgc    1140
cgttggaccc acgtgacccg gccaccattg atgcggtttt cgattggggc gcaggcgaga    1200
atacgggtgg catccatgcg gcggtcattc tgccggcaac ctcccacgaa ccggctccgt    1260
gcgtgattga agtcgatgac gaacgcgtcc tgaatttcct ggccgatgaa attaccggca    1320
ccatcgttat tgcgagccgt ttggcgcgct attggcaatc ccaacgcctg accccgggtg    1380
cccgtgcccg cggtccgcgt gttatctttc tgagcaacgg tgccgatcaa aatggtaatg    1440
tttacggtcg tattcaatct gcggcgatcg gtcaattgat tcgcgtttgg cgtcacgagg    1500
cggagttgga ctatcaacgt gcatccgccg caggcgatca cgttctgccg ccggtttggg    1560
cgaaccagat tgtccgtttc gctaaccgct ccctggaagg tctggagttc gcgtgcgcgt    1620
ggaccgcaca gctgctgcac agccaacgtc atattaacga aattacgctg aacattccag    1680
ccaatattag cgcgaccacg ggcgcacgtt ccgccagcgt cggctgggcc gagtccttga    1740
ttggtctgca cctgggcaag gtggctctga ttaccggtgg ttcggcgggc atcggtggtc    1800
aaatcggtcg tctgctggcc ttgtctggcg cgcgtgtgat gctggccgct cgcgatcgcc    1860
ataaattgga acagatgcaa gccatgattc aaagcgaatt ggcggaggtt ggttataccg    1920
atgtggagga ccgtgtgcac atcgctccgg gttgcgatgt gagcagcgag gcgcagctgg    1980
cagatctggt ggaacgtacg ctgtccgcat tcggtaccgt ggattatttg attaataacg    2040
ccggtattgc gggcgtggag gagatggtga tcgacatgcc ggtggaaggc tggcgtcaca    2100
ccctgtttgc caacctgatt tcgaattatt cgctgatgcg caagttggcg ccgctgatga    2160
```

-continued

```
agaagcaagg tagcggttac atcctgaacg tttcttccta ttttggcggt gagaaggacg   2220

cggcgattcc ttatccgaac cgcgccgact acgccgtctc caaggctggc caacgcgcga   2280

tggcggaagt gttcgctcgt ttcctgggtc cagagattca gatcaatgct attgccccag   2340

gtccggttga aggcgaccgc ctgcgtggta ccggtgagcg tccgggcctg tttgctcgtc   2400

gcgcccgtct gatcttggag aataaacgcc tgaacgaatt gcacgcggct ttgattgctg   2460

cggcccgcac cgatgagcgc tcgatgcacg agttggttga attgttgctg ccgaacgacg   2520

tggccgcgtt ggagcagaac ccagcggccc ctaccgcgct gcgtgagctg gcacgccgct   2580

tccgtagcga aggtgatccg gcggcaagct cctcgtccgc cttgctgaat cgctccatcg   2640

ctgccaagct gttggctcgc ttgcataacg gtggctatgt gctgccggcg gatatttttg   2700

caaatctgcc taatccgccg gacccgttct ttacccgtgc gcaaattgac cgcgaagctc   2760

gcaaggtgcg tgatggtatt atgggtatgc tgtatctgca gcgtatgcca accgagtttg   2820

acgtcgctat ggcaaccgtg tactatctgg ccgatcgtaa cgtgagcggc gaaactttcc   2880

atccgtctgg tggtttgcgc tacgagcgta ccccgaccgg tggcgagctg ttcggcctgc   2940

catcgccgga acgtctggcg gagctggttg gtagcacggt gtacctgatc ggtgaacacc   3000

tgaccgagca cctgaacctg ctggctcgtg cctatttgga gcgctacggt gcccgtcaag   3060

tggtgatgat tgttgagacg gaaaccggtg cggaaaccat gcgtcgtctg ttgcatgatc   3120

acgtcgaggc aggtcgcctg atgactattg tggcaggtga tcagattgag gcagcgattg   3180

accaagcgat cacgcgctat ggccgtccgg gtccggtggt gtgcactcca ttccgtccac   3240

tgccaaccgt tccgctggtc ggtcgtaaag actccgattg gagcaccgtt ttgagcgagg   3300

cggaatttgc ggaactgtgt gagcatcagc tgacccacca tttccgtgtt gctcgtaaga   3360

tcgccttgtc ggatggcgcg tcgctggcgt tggttacccc ggaaacgact gcgactagca   3420

ccacggagca atttgctctg gcgaacttca tcaagaccac cctgcacgcg ttcaccgcga   3480

ccatcggtgt tgagtcggag cgcaccgcgc aacgtattct gattaaccag gttgatctga   3540

cgcgccgcgc ccgtgcggaa gagccgcgtg acccgcacga gcgtcagcag gaattggaac   3600

gcttcattga agccgttctg ctggttaccg ctccgctgcc tcctgaggca gacacgcgct   3660

acgcaggccg tattcaccgc ggtcgtgcga ttaccgtcta atagaagctt gatatc       3716
```

```
<210> SEQ ID NO 12
<211> LENGTH: 45
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      PCR primer

<400> SEQUENCE: 12

tcgtaccaac catggccggt acgggtcgtt tggctggtaa aattg                   45


<210> SEQ ID NO 13
<211> LENGTH: 42
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      PCR primer

<400> SEQUENCE: 13
```

-continued

```
cggtgtgaga tctagatccg acggtaatcg cacgaccgcg gt                  42


<210> SEQ ID NO 14
<211> LENGTH: 8262
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      pkk223 plasmid comprising codon-optimized mcr nucleic acid
      sequence

<400> SEQUENCE: 14

ttgacaatta atcatcggct cgtataatgt gtggaattgt gagcggataa caatttcaca     60

caggaaacag aaccatggcc ggtacgggtc gtttggctgg taaaattgca ttgatcaccg    120

gtggtgctgg taacattggt tccgagctga cccgccgttt tctggccgag ggtgcgacgg    180

ttattatcag cggccgtaac cgtgcgaagc tgaccgcgct ggccgagcgc atgcaagccg    240

aggccggcgt gccggccaag cgcattgatt tggaggtgat ggatggttcc gaccctgtgg    300

ctgtccgtgc cggtatcgag gcaatcgtcg ctcgccacgg tcagattgac attctggtta    360

acaacgcggg ctccgccggt gcccaacgtc gcttggcgga aattccgctg acggaggcag    420

aattgggtcc gggtgcggag gagactttgc acgcttcgat cgcgaatctg ttgggcatgg    480

gttggcacct gatgcgtatt gcggctccgc acatgccagt tggctccgca gttatcaacg    540

tttcgactat tttctcgcgc gcagagtact atggtcgcat tccgtacgtt accccgaagg    600

cagcgctgaa cgctttgtcc cagctggctg cccgcgagct gggcgctcgt ggcatccgcg    660

ttaacactat tttcccaggt cctattgagt ccgaccgcat ccgtaccgtg tttcaacgta    720

tggatcaact gaagggtcgc ccggagggcg acaccgccca tcactttttg aacaccatgc    780

gcctgtgccg cgcaaacgac caaggcgctt tggaacgccg ctttccgtcc gttggcgatg    840

ttgctgatgc ggctgtgttt ctggcttctg ctgagagcgc ggcactgtcg ggtgagacga    900

ttgaggtcac ccacggtatg gaactgccgg cgtgtagcga aacctccttg ttggcgcgta    960

ccgatctgcg taccatcgac gcgagcggtc gcactaccct gatttgcgct ggcgatcaaa   1020

ttgaagaagt tatggccctg acgggcatgc tgcgtacgtg cggtagcgaa gtgattatcg   1080

gcttccgttc tgcggctgcc ctggcgcaat ttgagcaggc agtgaatgaa tctcgccgtc   1140

tggcaggtgc ggatttcacc ccgccgatcg ctttgccgtt ggacccacgt gacccggcca   1200

ccattgatgc ggttttcgat tggggcgcag gcgagaatac gggtggcatc catgcggcgg   1260

tcattctgcc ggcaacctcc cacgaaccgg ctccgtgcgt gattgaagtc gatgacgaac   1320

gcgtcctgaa tttcctggcc gatgaaatta ccggcaccat cgttattgcg agccgtttgg   1380

cgcgctattg gcaatcccaa cgcctgaccc cgggtgcccg tgcccgcggt ccgcgtgtta   1440

tctttctgag caacggtgcc gatcaaaatg gtaatgttta cggtcgtatt caatctgcgg   1500

cgatcggtca attgattcgc gtttggcgtc acgaggcgga gttggactat caacgtgcat   1560

ccgccgcagg cgatcacgtt ctgccgccgg tttgggcgaa ccagattgtc cgtttcgcta   1620

accgctccct ggaaggtctg gagttcgcgt gcgcgtggac cgcacagctg ctgcacagcc   1680

aacgtcatat taacgaaatt acgctgaaca ttccagccaa tattagcgcg accacgggcg   1740

cacgttccgc cagcgtcggc tgggccgagt ccttgattgg tctgcacctg ggcaaggtgg   1800

ctctgattac cggtggttcg gcgggcatcg gtggtcaaat cggtcgtctg ctggccttgt   1860

ctggcgcgcg tgtgatgctg gccgctcgcg atcgccataa attggaacag atgcaagcca   1920
```

-continued

```
tgattcaaag cgaattggcg gaggttggtt ataccgatgt ggaggaccgt gtgcacatcg   1980
ctccgggttg cgatgtgagc agcgaggcgc agctggcaga tctggtggaa cgtacgctgt   2040
ccgcattcgg taccgtggat tatttgatta ataacgccgg tattgcgggc gtggaggaga   2100
tggtgatcga catgccggtg gaaggctggc gtcacaccct gtttgccaac ctgatttcga   2160
attattcgct gatgcgcaag ttggcgccgc tgatgaagaa gcaaggtagc ggttacatcc   2220
tgaacgtttc ttcctatttt ggcggtgaga aggacgcggc gattccttat ccgaaccgcg   2280
ccgactacgc cgtctccaag gctggccaac gcgcgatggc ggaagtgttc gctcgtttcc   2340
tgggtccaga gattcagatc aatgctattg ccccaggtcc ggttgaaggc gaccgcctgc   2400
gtggtaccgg tgagcgtccg ggcctgtttg ctcgtcgcgc ccgtctgatc ttggagaata   2460
aacgcctgaa cgaattgcac gcggctttga ttgctgcggc ccgcaccgat gagcgctcga   2520
tgcacgagtt ggttgaattg ttgctgccga acgacgtggc cgcgttggag cagaacccag   2580
cggcccctac cgcgctgcgt gagctggcac gccgcttccg tagcgaaggt gatccggcgg   2640
caagctcctc gtccgccttg ctgaatcgct ccatcgctgc caagctgttg gctcgcttgc   2700
ataacggtgg ctatgtgctg ccggcggata tttttgcaaa tctgcctaat ccgccggacc   2760
cgttctttac ccgtgcgcaa attgaccgcg aagctcgcaa ggtgcgtgat ggtattatgg   2820
gtatgctgta tctgcagcgt atgccaaccg agtttgacgt cgctatggca accgtgtact   2880
atctggccga tcgtaacgtg agcggcgaaa ctttccatcc gtctggtggt ttgcgctacg   2940
agcgtacccc gaccggtggc gagctgttcg gcctgccatc gccggaacgt ctggcggagc   3000
tggttggtag cacggtgtac ctgatcggtg aacacctgac cgagcacctg aacctgctgg   3060
ctcgtgccta tttggagcgc tacggtgccc gtcaagtggt gatgattgtt gagacggaaa   3120
ccggtgcgga aaccatgcgt cgtctgttgc atgatcacgt cgaggcaggt cgcctgatga   3180
ctattgtggc aggtgatcag attgaggcag cgattgacca agcgatcacg cgctatggcc   3240
gtccgggtcc ggtggtgtgc actccattcc gtccactgcc aaccgttccg ctggtcggtc   3300
gtaaagactc cgattggagc accgttttga gcgaggcgga atttgcggaa ctgtgtgagc   3360
atcagctgac ccaccatttc cgtgttgctc gtaagatcgc cttgtcggat ggcgcgtcgc   3420
tggcgttggt taccccggaa acgactgcga ctagcaccac ggagcaattt gctctggcga   3480
acttcatcaa gaccaccctg cacgcgttca ccgcgaccat cggtgttgag tcggagcgca   3540
ccgcgcaacg tattctgatt aaccaggttg atctgacgcg ccgcgcccgt gcggaagagc   3600
cgcgtgaccc gcacgagcgt cagcaggaat tggaacgctt cattgaagcc gttctgctgg   3660
ttaccgctcc gctgcctcct gaggcagaca cgcgctacgc aggccgtatt caccgcggtc   3720
gtgcgattac cgtcggatct agatctcacc atcaccacca ttagtcgacc tgcagccaag   3780
cttggctgtt ttggcggatg agagaagatt ttcagcctga tacagattaa atcagaacgc   3840
agaagcggtc tgataaaaca gaatttgcct ggcggcagta gcgcggtggt cccacctgac   3900
cccatgccga actcagaagt gaaacgccgt agcgccgatg gtagtgtggg gtctccccat   3960
gcgagagtag ggaactgcca ggcatcaaat aaaacgaaag gctcagtcga aagactgggc   4020
ctttcgtttt atctgttgtt tgtcggtgaa cgctctcctg agtaggacaa atccgccggg   4080
agcggatttg aacgttgcga agcaacggcc cggagggtgg cgggcaggac gcccgccata   4140
aactgccagg catcaaatta agcagaaggc catcctgacg gatggccttt ttgcgtttct   4200
```

-continued

```
acaaactctt ttgtttattt ttctaaatac attcaaatat gtatccgctc atgagacaat   4260
aaccctgata aatgcttcaa taatattgaa aaaggaagag tatgagtatt caacatttcc   4320
gtgtcgccct tattcccttt tttgcggcat tttgccttcc tgtttttgct cacccagaaa   4380
cgctggtgaa agtaaaagat gctgaagatc agttgggtgc acgagtgggt tacatcgaac   4440
tggatctcaa cagcggtaag atccttgaga gttttcgccc cgaagaacgt tttccaatga   4500
tgagcacttt taaagttctg ctatgtggcg cggtattatc ccgtgttgac gccgggcaag   4560
agcaactcgg tcgccgcata cactattctc agaatgactt ggttgagtac tcaccagtca   4620
cagaaaagca tcttacggat ggcatgacag taagagaatt atgcagtgct gccataacca   4680
tgagtgataa cactgcggcc aacttacttc tgacaacgat cggaggaccg aaggagctaa   4740
ccgctttttt gcacaacatg ggggatcatg taactcgcct tgatcgttgg gaaccggagc   4800
tgaatgaagc cataccaaac gacgagcgtg acaccacgat gctgtagcaa tggcaacaac   4860
gttgcgcaaa ctattaactg gcgaactact tactctagct tcccggcaac aattaataga   4920
ctggatggag gcggataaag ttgcaggacc acttctgcgc tcggcccttc cggctggctg   4980
gtttattgct gataaatctg gagccggtga gcgtgggtct cgcggtatca ttgcagcact   5040
ggggccagat ggtaagccct cccgtatcgt agttatctac acgacgggga gtcaggcaac   5100
tatggatgaa cgaaatagac agatcgctga gataggtgcc tcactgatta agcattggta   5160
actgtcagac caagtttact catatatact ttagattgat ttaaaacttc atttttaatt   5220
taaaaggatc taggtgaaga tcctttttga taatctcatg accaaaatcc cttaacgtga   5280
gttttcgttc cactgagcgt cagaccccgt agaaaagatc aaaggatctt cttgagatcc   5340
tttttttctg cgcgtaatct gctgcttgca aacaaaaaaa ccaccgctac cagcggtggt   5400
ttgtttgccg gatcaagagc taccaactct ttttccgaag gtaactggct tcagcagagc   5460
gcagatacca aatactgtcc ttctagtgta gccgtagtta ggccaccact tcaagaactc   5520
tgtagcaccg cctacatacc tcgctctgct aatcctgtta ccagtggctg ctgccagtgg   5580
cgataagtcg tgtcttaccg ggttggactc aagacgatag ttaccggata aggcgcagcg   5640
gtcgggctga acggggggtt cgtgcacaca gcccagcttg gagcgaacga cctacaccga   5700
actgagatac ctacagcgtg agcattgaga aagcgccacg cttcccgaag ggagaaaggc   5760
ggacaggtat ccggtaagcg gcagggtcgg aacaggagag cgcacgaggg agcttccagg   5820
gggaaacgcc tggtatcttt atagtcctgt cgggtttcgc cacctctgac ttgagcgtcg   5880
atttttgtga tgctcgtcag gggggcggag cctatggaaa aacgccagca acgcggcctt   5940
tttacggttc ctggcctttt gctggccttt tgctcacatg ttctttcctg cgttatcccc   6000
tgattctgtg gataaccgta ttaccgcctt tgagtgagct gataccgctc gccgcagccg   6060
aacgaccgag cgcagcgagt cagtgagcga ggaagcggaa gagcgcctga tgcggtattt   6120
tctccttacg catctgtgcg gtatttcaca ccgcatatgg tgcactctca gtacaatctg   6180
ctctgatgcc gcatagttaa gccagtatac actccgctat cgctacgtga ctgggtcatg   6240
gctgcgcccc gacacccgcc aacacccgct gacgcgccct gacgggcttg tctgctcccg   6300
gcatccgctt acagacaagc tgtgaccgtc tccgggagct gcatgtgtca gaggttttca   6360
ccgtcatcac cgaaacgcgc gaggcagctg cggtaaagct catcagcgtg gtcgtgaagc   6420
gattcacaga tgtctgcctg ttcatccgcg tccagctcgt tgagtttctc cagaagcgtt   6480
aatgtctggc ttctgataaa gcgggccatg ttaagggcgg tttttcctg tttggtcact    6540
```

-continued

```
gatgcctccg tgtaaggggg atttctgttc atgggggtaa tgataccgat gaaacgagag   6600

aggatgctca cgatacgggt tactgatgat gaacatgccc ggttactgga acgttgtgag   6660

ggtaaacaac tggcggtatg gatgcggcgg gaccagagaa aaatcactca gggtcaatgc   6720

cagcgcttcg ttaatacaga tgtaggtgtt ccacagggta gccagcagca tcctgcgatg   6780

cagatccgga acataatggt gcagggcgct gacttccgcg tttccagact ttacgaaaca   6840

cggaaaccga agaccattca tgttgttgct caggtcgcag acgttttgca gcagcagtcg   6900

cttcacgttc gctcgcgtat cggtgattca ttctgctaac cagtaaggca accccgccag   6960

cctagccggg tcctcaacga caggagcacg atcatgcgca cccgtggcca ggacccaacg   7020

ctgcccgaga tgcgccgcgt gcggctgctg gagatggcgg acgcgatgga tatgttctgc   7080

caagggttgg tttgcgcatt cacagttctc cgcaagaatt gattggctcc aattcttgga   7140

gtggtgaatc cgttagcgag gtgccgccgg cttccattca ggtcgaggtg gcccggctcc   7200

atgcaccgcg acgcaacgcg gggaggcaga caaggtatag ggcggcgcct acaatccatg   7260

ccaacccgtt ccatgtgctc gccgaggcgg cataaatcgc cgtgacgatc agcggtccag   7320

tgatcgaagt taggctggta agagccgcga gcgatccttg aagctgtccc tgatggtcgt   7380

catctacctg cctggacagc atggcctgca acgcgggcat cccgatgccg ccggaagcga   7440

gaagaatcat aatggggaag gccatccagc ctcgcgtcgc gaacgccagc aagacgtagc   7500

ccagcgcgtc ggccgccatg ccggcgataa tggcctgctt ctcgccgaaa cgtttggtgg   7560

cgggaccagt gacgaaggct tgagcgaggg cgtgcaagat tccgaatacc gcaagcgaca   7620

ggccgatcat cgtcgcgctc cagcgaaagc ggtcctcgcc gaaaatgacc cagagcgctg   7680

ccggcacctg tcctacgagt tgcatgataa agaagacagt cataagtgcg gcgacgatag   7740

tcatgccccg cgcccaccgg aaggagctga ctgggttgaa ggctctcaag ggcatcggtc   7800

gacgctctcc cttatgcgac tcctgcatta ggaagcagcc cagtagtagg ttgaggccgt   7860

tgagcaccgc cgccgcaagg aatggtgcat gcaaggagat ggcgcccaac agtcccccgg   7920

ccacggggcc tgccaccata cccacgccga aacaagcgct catgagcccg aagtggcgag   7980

cccgatcttc cccatcggtg atgtcggcga tataggcgcc agcaaccgca cctgtggcgc   8040

cggtgatgcc ggccacgatg cgtccggcgt agaggatccg ggcttatcga ctgcacggtg   8100

caccaatgct tctggcgtca ggcagccatc ggaagctgtg gtatggctgt gcaggtcgta   8160

aatcactgca taattcgtgt cgctcaaggc gcactcccgt tctggataat gttttttgcg   8220

ccgacatcat aacggttctg gcaaatattc tgaaatgagc tg                       8262
```

```
<210> SEQ ID NO 15
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      PCR primer

<400> SEQUENCE: 15

gggtttccat ggaccagccg ctcaacgtgg                                       30


<210> SEQ ID NO 16
<211> LENGTH: 33
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
```

-continued

```
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      PCR primer

<400> SEQUENCE: 16

gggttttcag gcgatgccgt tgagcgcttc gcc                                33


<210> SEQ ID NO 17
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      PCR primer

<400> SEQUENCE: 17

gggaacggcg gggaaaaaca aacgtt                                        26


<210> SEQ ID NO 18
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      PCR primer

<400> SEQUENCE: 18

ggtccatggt aattctccac gcttataagc                                    30


<210> SEQ ID NO 19
<211> LENGTH: 85
<212> TYPE: DNA
<213> ORGANISM: Escherichia coli

<400> SEQUENCE: 19

ggtttgaata aatgacaaaa agcaaagcct ttgtgccgat gaatctctat actgtttcac   60

agacctgctg ccctgcgggg cggcc                                         85


<210> SEQ ID NO 20
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      PCR primer

<400> SEQUENCE: 20

gggaacggcg gggaaaaaca aacgtt                                        26


<210> SEQ ID NO 21
<211> LENGTH: 33
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      PCR primer

<400> SEQUENCE: 21

gggttttcag gcgatgccgt tgagcgcttc gcc                                33


<210> SEQ ID NO 22
<211> LENGTH: 1754
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
```

-continued

<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
Codon-optimized pyruvate decarboxylase DNA sequence based on gene
of Zymomonas mobilis

<400> SEQUENCE: 22

```
catatgtcct acactgttgg tacttatctg gctgaacgtc tggttcaaat tggtctgaag     60

catcactttg cggtagcggg cgattacaac ctggtgctgc tggataatct gctgctgaac    120

aaaaacatgg aacaggtcta ctgttgtaac gaactgaact gcggcttctc tgctgaaggt    180

tatgcccggg ccaaaggcgc agctgcggcc gtagtgacct actccgttgg tgctctgtcc    240

gcatttgatg caattggtgg cgcctacgca gaaaacctgc cggtaattct gatctctggc    300

gctccgaaca acaacgatca cgcagctggt cacgtgctgc accatgcgct gggcaaaact    360

gattatcatt accagctgga aatggcgaag aacatcactg ccgcagcaga agctatctat    420

actccagaag aagcgccggc aaaaatcgat catgtaatca aaacggccct gcgtgagaag    480

aaaccggtgt atctggaaat tgcttgtaat atcgcgtcta tgccgtgtgc ggccccaggc    540

ccagcatctg ctctgtttaa cgatgaagct agcgatgaag cctctctgaa cgcagctgtg    600

gaagaaaccc tgaaattcat tgcaaaccgt gacaaagttg cggtactggt tggctctaaa    660

ctgcgtgccg cgggtgcaga agaagcggcg gttaaattcg ctgacgccct gggtggtgct    720

gtggccacca tggctgcggc taaatccttt ttcccggaag aaaatccgca ttacatcggt    780

acttcctggg gcgaggtttc ttacccaggt gtcgagaaaa ccatgaagga agctgacgcg    840

gtgatcgccc tggccccggt tttcaatgac tactccacta ctggttggac cgacatcccg    900

gacccaaaga aactggttct ggcagagccg cgctccgttg ttgttaacgg tattcgcttt    960

ccgtccgtac acctgaagga ttatctgact cgtctggcgc agaaagtgag caagaaaacc   1020

ggcgctctgg atttctttaa atctctgaat gcgggtgagc tgaagaaagc cgcaccggcg   1080

gacccttctg ctccgctggt taacgccgaa attgcgcgcc aggtagaagc gctgctgact   1140

ccgaatacta ccgtaattgc ggagactggc gattcctggt tcaacgcaca acgtatgaag   1200

ctgcctaacg gcgctcgagt tgaatacgaa atgcagtggg gccacatcgg ctggtctgtt   1260

cctgcagcct tcggctacgc cgtaggtgct ccggaacgtc gtaacatcct gatggtcggt   1320

gacggctctt tccaactgac cgcgcaggaa gtagcacaga tggttcgtct gaaactgccg   1380

gtaatcatct tcctgattaa caactacggc tataccattg aggtcatgat tcatgatggt   1440

ccgtataata acatcaaaaa ctgggattat gctggtctga tggaagtttt caacggcaac   1500

ggcggctacg attctggtgc tggtaaaggc ctgaaagcaa agacgggtgg cgagctcgca   1560

gaagcgatca aggttgctct ggctaacacc gatggtccga ctctgatcga atgttttatc   1620

ggtcgtgaag attgcactga ggaactggtg aagtggggta agcgtgtggc tgccgcgaat   1680

tcccgtaaac cggtaaataa gcttctcggc catcaccatc accatcacta gaagcttctc   1740

tagagaacta tttc                                                     1754
```

<210> SEQ ID NO 23
<211> LENGTH: 575
<212> TYPE: PRT
<213> ORGANISM: Zymomonas mobilis

<400> SEQUENCE: 23

```
Met Ser Tyr Thr Val Gly Thr Tyr Leu Ala Glu Arg Leu Val Gln Ile
1               5                   10                  15
```

-continued

```
Gly Leu Lys His His Phe Ala Val Ala Gly Asp Tyr Asn Leu Val Leu
            20                  25                  30

Leu Asp Asn Leu Leu Leu Asn Lys Asn Met Glu Gln Val Tyr Cys Cys
        35                  40                  45

Asn Glu Leu Asn Cys Gly Phe Ser Ala Glu Gly Tyr Ala Arg Ala Lys
    50                  55                  60

Gly Ala Ala Ala Ala Val Val Thr Tyr Ser Val Gly Ala Leu Ser Ala
65                  70                  75                  80

Phe Asp Ala Ile Gly Gly Ala Tyr Ala Glu Asn Leu Pro Val Ile Leu
                85                  90                  95

Ile Ser Gly Ala Pro Asn Asn Asn Asp His Ala Ala Gly His Val Leu
            100                 105                 110

His His Ala Leu Gly Lys Thr Asp Tyr His Tyr Gln Leu Glu Met Ala
        115                 120                 125

Lys Asn Ile Thr Ala Ala Ala Glu Ala Ile Tyr Thr Pro Glu Glu Ala
    130                 135                 140

Pro Ala Lys Ile Asp His Val Ile Lys Thr Ala Leu Arg Glu Lys Lys
145                 150                 155                 160

Pro Val Tyr Leu Glu Ile Ala Cys Asn Ile Ala Ser Met Pro Cys Ala
                165                 170                 175

Ala Pro Gly Pro Ala Ser Ala Leu Phe Asn Asp Glu Ala Ser Asp Glu
            180                 185                 190

Ala Ser Leu Asn Ala Ala Val Glu Glu Thr Leu Lys Phe Ile Ala Asn
        195                 200                 205

Arg Asp Lys Val Ala Val Leu Val Gly Ser Lys Leu Arg Ala Ala Gly
    210                 215                 220

Ala Glu Glu Ala Ala Val Lys Phe Ala Asp Ala Leu Gly Gly Ala Val
225                 230                 235                 240

Ala Thr Met Ala Ala Ala Lys Ser Phe Phe Pro Glu Glu Asn Pro His
                245                 250                 255

Tyr Ile Gly Thr Ser Trp Gly Glu Val Ser Tyr Pro Gly Val Glu Lys
            260                 265                 270

Thr Met Lys Glu Ala Asp Ala Val Ile Ala Leu Ala Pro Val Phe Asn
        275                 280                 285

Asp Tyr Ser Thr Thr Gly Trp Thr Asp Ile Pro Asp Pro Lys Lys Leu
    290                 295                 300

Val Leu Ala Glu Pro Arg Ser Val Val Val Asn Gly Ile Arg Phe Pro
305                 310                 315                 320

Ser Val His Leu Lys Asp Tyr Leu Thr Arg Leu Ala Gln Lys Val Ser
                325                 330                 335

Lys Lys Thr Gly Ala Leu Asp Phe Phe Lys Ser Leu Asn Ala Gly Glu
            340                 345                 350

Leu Lys Lys Ala Ala Pro Ala Asp Pro Ser Ala Pro Leu Val Asn Ala
        355                 360                 365

Glu Ile Ala Arg Gln Val Glu Ala Leu Leu Thr Pro Asn Thr Thr Val
    370                 375                 380

Ile Ala Glu Thr Gly Asp Ser Trp Phe Asn Ala Gln Arg Met Lys Leu
385                 390                 395                 400

Pro Asn Gly Ala Arg Val Glu Tyr Glu Met Gln Trp Gly His Ile Gly
                405                 410                 415

Trp Ser Val Pro Ala Ala Phe Gly Tyr Ala Val Gly Ala Pro Glu Arg
```

-continued

```
            420                 425                 430

Arg Asn Ile Leu Met Val Gly Asp Gly Ser Phe Gln Leu Thr Ala Gln
        435                 440                 445

Glu Val Ala Gln Met Val Arg Leu Lys Leu Pro Val Ile Ile Phe Leu
    450                 455                 460

Ile Asn Asn Tyr Gly Tyr Thr Ile Glu Val Met Ile His Asp Gly Pro
465                 470                 475                 480

Tyr Asn Asn Ile Lys Asn Trp Asp Tyr Ala Gly Leu Met Glu Val Phe
                485                 490                 495

Asn Gly Asn Gly Gly Tyr Asp Ser Gly Ala Gly Lys Gly Leu Lys Ala
            500                 505                 510

Lys Thr Gly Gly Glu Leu Ala Glu Ala Ile Lys Val Ala Leu Ala Asn
        515                 520                 525

Thr Asp Gly Pro Thr Leu Ile Glu Cys Phe Ile Gly Arg Glu Asp Cys
    530                 535                 540

Thr Glu Glu Leu Val Lys Trp Gly Lys Arg Val Ala Ala Ala Asn Ser
545                 550                 555                 560

Arg Lys Pro Val Asn Lys Leu Leu Gly His His His His His His
                565                 570                 575


<210> SEQ ID NO 24
<211> LENGTH: 5
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      Histidine tag peptide sequence

<400> SEQUENCE: 24

His His His His His
1               5


<210> SEQ ID NO 25
<211> LENGTH: 7
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      Histidine tag peptide sequence

<400> SEQUENCE: 25

His His His His His His His
1               5
```

What is claimed is:

1. A method for producing an acrylic acid-based consumer product, said method comprising

i) combining a carbon source and a microorganism cell culture to produce 3-hydroxypropionic acid;

ii) converting said 3-hydroxypropionic acid to acrylic acid; and

iii) processing said acrylic acid into a consumer product.

2. The method of claim **1**, wherein said cell culture comprises a genetically modified microorganism.

3. The method of claim **2**, wherein said microorganism is modified for increased tolerance to 3-hydroxypropionic acid.

4. The method of claim **3**, wherein said modification modulates one or more components of the chorismate superpathway.

5. The method of claim **2**, wherein said microorganism is modified for increased production of 3-hydroxypropionic acid.

6. The method of claim **5**, wherein said modification comprises reduction in activity of one or more enzymes selected from pyruvate kinase, phosphofructokinase, lactate dehydrogenase, phosphate acetyltransferase, pyruvate oxidase, pyruvate-formate lyase, and homologs thereof.

7. The method of claim **6**, wherein said pyruvate kinase is selected from pykA and pykF, said phosphofructokinase is selected from pfkA and pfkB, said lactate dehydrogenase is selected from ldhA, said phosphate acetyltransferase is selected from pta, said pyruvate oxidase is selected from poxB, said pyruvate-formate lyase is selected from pflB, and homologs thereof.

8. The method of claim **5**, wherein said modification comprises increase in activity of one or more enzymes selected from phosphoenolpyruvate carboxykinase, malonyl-CoA reductase, 3-hydroxypropionate dehydrogenase, malonate semialdehyde dehydrogenase A, alpha-ketoglutarate decarboxylase, oxaloacetate alpha-oxo-decarboxylase, and homologs thereof.

9. The method of claim **8**, wherein said phosphoenolpyruvate carboxykinase is selected from pck, said malonyl-CoA reductase is selected from mcr, said malonate semialdehyde dehydrogenase A is selected from mmsA, said 3-hydroxypropionic acid dehydrogenase is selected from mmsB, said alpha-ketoglutarate decarboxylase is selected from kgd, said oxaloacetate alpha-oxo-decarboxylase is selected from oad, and homologs thereof.

10. The method of claim **8**, wherein said modification comprises an increase in activity in a malonyl-CoA reductase enzyme.

11. The method of claim **8**, wherein said modification comprises an increase in activity in an oxaloacetate alpha-oxo-decarboxylase enzyme.

12. The method of claim **2**, wherein said microorganism is modified for increased tolerance to 3-hydroxypropionic acid, and wherein said microorganism is modified for increased production of 3-hydroxypropionic acid.

13. A method for producing an acrylic acid-based consumer product, said method comprising

i) combining a carbon source and a genetically modified microorganism in cell culture to produce 3-hydroxypropionic acid;

ii) converting said 3-hydroxypropionic acid to acrylic acid; and

iii) processing said acrylic acid into a consumer product;

wherein said microorganism is modified for increased production of 3-hydroxypropionic acid via an increase in activity in an oxaloacetate alpha decarboxylase enzyme or homolog thereof.

14. A method for producing acrylic acid, said method comprising

i) combining a carbon source and a microorganism cell culture to produce 3-hydroxypropionic acid in a concentration of at least 10 g/L; and

ii) converting said 3-hydroxypropionic acid to acrylic acid.

15. The method of claim **14**, wherein the combining comprises combining a carbon source and a microorganism cell culture comprising a microorganism genetically modified to increase activity of one or more enzymes selected from phosphoenolpyruvate carboxykinase, malonyl-CoA reductase, malonate semialdehyde dehydrogenase A, 3-HP dehydrogenase, alpha-ketoglutarate decarboxylase, oxaloacetate alpha-oxo-decarboxylase, and homologs thereof.

16. The method of claim **15**, wherein said phosphoenolpyruvate carboxykinase is selected from pck, said malonyl-CoA reductase is selected from mcr, said malonate semialdehyde dehydrogenase A is selected from mmsA, said 3-hydroxypropionic acid dehydrogenase is selected from mmsB, said alpha-ketoglutarate decarboxylase is selected from kgd, said oxaloacetate alpha-oxo-decarboxylase is selected from oad, and homologs thereof.

17. The method of claim **14**, wherein the combining comprises combining a carbon source and a microorganism cell culture comprising a microorganism genetically modified to increase activity of malonyl-CoA reductase enzyme increasing production of 3-hydroxypropionic acid.

18. The method of claim **14**, wherein the combining comprises combining a carbon source and a microorganism cell culture comprising a microorganism genetically modified to increase activity of an oxaloacetate alpha-oxo-decarboxylase enzyme, increasing production of 3-hydroxypropionic acid.

19. Biologically-produced acrylic acid, wherein said acrylic acid is produced according to claim **14**.

20. A consumer product produced with acrylic acid according to claim **19**.

* * * * *