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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
13 October 2016 (13.10.2016)

WIPO | PCT

(10) International Publication Number
WO 2016/164339 A2

(51) International Patent Classification: Not classified

(21) International Application Number:
PCT/US2016/025991

(22) International Filing Date:
5 April 2016 (05.04.2016)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
14/680,195 7 April 2015 (07.04.2015) US

(71) Applicant: **SYNATA BIO, INC.** [US/US]; 4575 Weaver Parkway, Suite 100, Warrenville, IL 60555 (US).

(72) Inventors: **REEVES, Andrew**; 226 North Clinton Street, Unit 316, Chicago, IL 60661 (US). **TOBEY, Richard**; 42W485 Sylvan Lane, St. Charles, IL 60175 (US). **ENZIEN, Mike**; 6528 Royal Glen Court, Lisle, IL 60532 (US).

(74) Agent: **FENDRICK, Sarah, E.**; McDonnell Boehnen Hulbert, & Berghoff LLP, 300 South Wacker Drive, Chicago, IL 60606 (US).

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

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

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

— *with sequence listing part of description (Rule 5.2(a))*

WO 2016/164339 A2

(54) Title: COMPOSITIONS AND METHODS FOR THE CONVERSION OF SHORT-CHAINED CARBOXYLIC ACIDS TO ALCOHOLS USING CLOSTRIDIAL ENZYMES

(57) Abstract: The invention relates to the fields of bacterial metabolism and the utilization or consumption of short-chain carboxylic acids to reduced products. Specifically, it relates to syngas fermentations using monocultures of syngas-utilizing homoacetogenic bacteria for the production of alcohols using native alcohol dehydrogenase.

COMPOSITIONS AND METHODS FOR THE CONVERSION OF SHORT-CHAINED CARBOXYLIC ACIDS TO ALCOHOLS USING CLOSTRIDIAL ENZYMES

FIELD OF THE INVENTION

[0001] The invention relates to the fields of bacterial metabolism and the utilization or consumption of short-chain carboxylic acids to reduced products. Specifically, it relates to syngas fermentations using monocultures of syngas-utilizing homoacetogenic bacteria for the production of alcohols using native alcohol dehydrogenase. This invention also relates to co-cultures of microorganisms or consortia of microorganisms that comprise a syngas-utilizing homoacetogen and a butyrogenic organism that when grown together produce short-chain carboxylic acids from C_2-C_6 in length and syntrophically convert them efficiently to the commercially useful C_2-C_6 alcohols.

BACKGROUND OF THE INVENTION

[0002] Butanol is an important industrial chemical with a wide range of applications. It can be used as a motor fuel particularly in combination with gasoline to which it can be added in all proportions. Isobutanol can also be used as a precursor to Methyl Tertiary Butyl Ether (MTBE). Currently the world production of n-Butanol is 3.5 million tons /yr. (7.7 billion lb /yr). Furthermore, conversion of alcohols to long-chain linear hydrocarbons that would be suitable for jet fuel use are being developed and demonstrated, which could further increase the demand for n-butanol (The Naval Air Warfare Center-Weapons Division, (2012) Cobalt and Abermarle). Fermentation of carbohydrates to acetone, butanol and ethanol (ABE) is well known and was commercially practiced worldwide from around 1915 to 1955 (Beesch, S.C. (1953). A Microbiological Process Report – Applied Microbiology, 1, 85-95). With the advent of petrochemical processes and low-cost petrochemical feedstocks the carbohydrate-based processes became unattractive and were discontinued.

[0003] Attempts have been made to improve the alcohol yield of bacteria that ferment a variety of sugars to acetate and butyrate. The art has sought to employ recombinant techniques to transform bacterium such as *C. acetobutylicum* (Green et al., 1996 Genetic manipulation of acid formation pathways *Microbiology*, 142, 2079-2086) and *C. tyrobutyricum* (X. Liu et al., 2006 Construction and Characterization of an *ack* Deleted Mutant of *Clostridium tyrobutyricum*, *Biotechnology Progress*, 22, 1265-1275). However, such techniques have only resulted in transformation occurring at low frequencies.

[0004] Anaerobic acetogenic microorganisms offer a viable route to convert waste gases, such as syngas, to useful products, such as ethanol, via a fermentation process. Such bacteria catalyze the conversion of H_2 and CO_2 and/or CO to acids and/or alcohols with higher specificity, higher yields and lower energy costs than can be attained by traditional production processes. While many of the anaerobic microorganisms utilized in the fermentation of ethanol also produce butanol as a secondary product, to date, no single anaerobic microorganism has been described that can utilize the syngas fermentation process to produce high yields of butanol.

[0005] Thus, there remains a need in the art to produce a biocatalyst bacterium or co-culture of bacteria that produces useful commercial products such as C_2-C_6 alcohols (including ethanol, propanol, butanol, pentanol and hexanol and any of their isomers).

SUMMARY OF THE INVENTION

[0006] Provided herein are methods for the production of an alcohol, comprising: culturing a recombinant C1-fixing homoacetogen microorganism in a culture medium in the presence of syngas, wherein the C1-fixing homoacetogen microorganism comprises a recombinant gene encoding an aldehyde ferredoxin oxidoreductase polypeptide modulated by a promoter, wherein the recombinant C1-fixing microorganism produces an alcohol.

[0007] In particular embodiments the aldehyde ferredoxin oxidoreductase gene encodes a *Clostridium* aldehyde ferredoxin oxidoreductase polypeptide. In particular embodiments the aldehyde oxidoreductase gene is at least 95 % identical to SEQ ID NO: 1 or SEQ ID NO: 3.

[0008] In particular embodiments the promoter is an inducible promoter or a constitutive promoter.

[0009] In particular embodiments the methods provided herein produce ethanol. In other particular embodiments, the C1-fixing homoacetogen microorganism *is C. ljungdahlii, C. ragsdalei, C. autoethanogenum* or *C. coskatii.*

[0010] In other aspects, provided herein are methods for the production of butanol, comprising: culturing a co-culture comprising a recombinant C1-fixing homoacetogen microorganism and a C4-producing butyrate microorganism in a culture medium in the presence of syngas, wherein the C1-fixing homoacetogen microorganism comprises a recombinant gene encoding a aldehyde ferredoxin oxidoreductase polypeptide modulated by a promoter.

[0011] In particular embodiments the C4-producing butyrate microorganism comprises a gene encoding a Butyryl-CoA acetate transferase polypeptide or a Butyrate kinase polypeptide. In other particular embodiments, the aldehyde ferredoxin oxidoreductase gene encodes a *Clostridium* aldehyde ferredoxin oxidoreductase polypeptide. In yet other embodiments, the aldehyde ferredoxin oxidoreductase gene is at least 95% identical to SEQ ID NO: 1 or SEQ ID NO: 3.

[0012] In particular embodiments the promoter is an inducible promoter or a constitutive promoter. In other particular embodiments the C1-fixing homoacetogen microorganism *is C. ljungdahlii, C. ragsdalei, C. autoethanogenum* or *C. coskatii.* In other particular embodiments

the C4-producing butyrate microorganism is *Clostridium kluyveri*, *Clostridium carboxidivorans*, *Butyribacterium methylotrophicum*, or *Clostridium pharus*.

[0013] In other aspects provided herein are microorganism co-cultures for the conversion of syngas to butanol comprising: a recombinant C1-fixing homoacetogen microorganism comprising a recombinant gene encoding a aldehyde ferredoxin oxidoreductase polypeptide modulated by a promoter and a C4-producing butyrate microorganism.

[0014] In particular embodiments the C4-producing butyrate microorganism comprises a gene encoding a Butyryl-CoA acetate transferase polypeptide or a Butyrate kinase polypeptide. In other particular embodiments, the aldehyde ferredoxin oxidoreductase gene encodes a *Clostridium* aldehyde ferredoxin oxidoreductase polypeptide. In yet other embodiments, the aldehyde ferredoxin oxidoreductase gene is at least 95 % identical to SEQ ID NO: 1 or SEQ ID NO: 3.

[0015] In particular embodiments, the promoter is an inducible promoter or a constitutive promoter. In yet other embodiments, the C1-fixing homoacetogen microorganism *is C. ljungdahlii, C. ragsdalei, C. autoethanogenum* or *C. coskatii.* In yet other embodiments, the C4-producing butyrate microorganism is *Clostridium kluyveri*, *Clostridium carboxidivorans*, *Butyribacterium methylotrophicum*, or *Clostridium pharus*.

BRIEF DESCRIPTION OF THE DRAWINGS

[0016] Fig. 1. Schematic illustration of the Wood-Ljungdahl (C_1) and C_2 biosynthesis pathways from syngas showing among others the enzymatic reactions catalyzed by native homoacetogen aldehyde ferredoxin oxidoreductase (AOR).

[0017] Figs. 2A and 2B. Physical maps of AOR expression vectors containing one of the cloned aldehyde ferredoxin oxidoreductase genes identified in *C. autoethanogenum*. 2(A) Expression vector pCKAR192, containing one of the Clostridial aldehyde ferredoxin

oxidoreductase genes expressed from a constitutive Clostridial promoter Pcoos. This promoter has been shown to be highly expressed in *E. coli* DH10B cells. 2(B) *E.coli-Clostridium* shuttle vector pCKAR192T2, the same as pCKAR192 but containing a clostridial replicon and a chloramphenicol resistance gene known to express in homacetogenic Clostridia.

[0018] Figs. 3A and 3B. Physical maps of AOR expression vectors containing cloned aldehyde ferredoxin oxidoreductase genes. 3(A) Expression vector pCKAR214, containing one of the Clostridial aldehyde ferredoxin oxidoreductase genes expressed from a constitutive Clostridial promoter Ppta-ack. This promoter has been shown to be highly expressed in *E. coli* DH10B cells. 3(B) *E.coli-Clostridium* shuttle vector pCKAR214T2, the same as pCKAR214 but containing a clostridial replicon and a chloramphenicol resistance gene known to express in homoacetogenic Clostridia.

[0019] Figs. 4A and 4B. DNA and amino acid sequence alignments of the two native AORs identified in *C. autoethanogenum*. 4(A) alignment of the two AOR gene sequences AOR1 and AOR2 and 4(B) amino acid alignment of AOR1 and AOR2.

[0020] Figs. 5A and 5B. Amino acid alignment of the two native AORs identified in *C. autoethanogenum* aligned with the best homology matches of the AOR amino acid sequences identified in *C. ljungdahlii* PETC (AOR1: SEQ ID NO: 15; AOR2: SEQ ID NO:17) and *C. ragsdalei* ATCC BAA-624 (AOR1: SEQ ID NO: 16; AOR2: SEQ ID NO:18). 5(A) alignment of AOR1 amino acid sequences and 5(B) alignment of AOR2 amino acid sequences.

[0021] Fig. 6. Schematic pathways showing co-culture metabolism and butanol production. Homoacetogen reactions generate ethanol and acetate and butanol, from butyrate, the butyrogen product.

[0022] Fig. 7. Schematic pathways showing acetate production by homoacetogens coupled with its consumption (along with hydrogen gas) by the butyrogen to generate butyrate. This is followed by conversion of the butyrate to butanol by the homoacetogen organism. Butyrate is generated by native phosphotransbutyrylase and butyrate kinase activities.

[0023] Figs. 8A-8D. Agarose gel showing PCR products produced with AOR-specific primers. Figs. 8A and 8B show electrophoresis gels of amplicons using primers AOR1A (Fig. 8A) and BCoAAT (Fig. 8B) designed to target *C. autoethanogenum* AOR1 and butyrogen butyryl-CoA acetate transferase, respectively; Figs 8C and 8D show electrophoresis gels of amplicons using primers AOR2A (Fig. 8C) and Buk (Fig. 8D) designed to target *C. autoethanogenum* AOR2 and butyrogen butyrate kinase.

[0024] Fig. 9. Graph illustrating AOR1and AOR2 gene expression and ethanol and acetate production in *C. autoethanogenum* during syngas co-fermentation. For gene expression analysis, using RT-qPCR (presented as fold-change), primer set 467/9 targeted AOR2-A (SEQ ID NO:14), primer set 474/478 targeted AOR1-A (SEQ ID NO: 14), and primer set 441 a targeted *C. autoethanogenum* hydrogenase.

[0025] Fig. 10. Graph illustrating quantitative gene expression of *C. autoethanogenum* AOR and hydrogenase genes during a butanol co-fermentation run over the course of >900 hours. Transcript copy number at six different time points is shown. For gene expression analysis using RT-qPCR the following primer sets were used to calculate copy number/µg RNA throughout the fermentation: 467/9 targeted AOR2; 474/8 targeted AOR1; 431 targeted a NiFe hydrogenase; 435 targeted an iron-only hydrogenase; 437 targeted a second iron-only hydrogenase; and 441 targeted a third iron-only hydrogenase found in the C. autoethanogenum chromosome.

[0026] Fig. 11. Graph illustrating heterologous AOR1 expression in *E. coli*. Bioconversion of butyrate to butanol in *E. coli* expressing AOR1 on pCKAR192. No butanol production is observed in *E. coli* containing only pCKAR162.

[0027] Figs. 12A and 12B. Diagrams of *E. coli-Clostridium* shuttle vectors pCK32 (Fig. 12A) and pCK32AOR1 (Fig 12B). *E. coli-Clostridium* shuttle vectors used for transforming Clostridium biocatalyst strains. pCK32AOR1 differs from pCK232 only in that it contains the AOR1 gene expressed from the Clostridium promoter Ppta.

[0028] Figs. 13A and 13B. Graph illustrating aldehyde ferredoxin oxidoreductase expression in wild-type and TF18 strains. RT-qPCR fold change analysis for wild-type and TF18 *C. autoethanogenum* cells grown on syngas (Fig. 13A) and fructose (Fig 13B). Expression patterns of the AOR1 gene were compared during early-log (on syngas) and mid-log (on fructose).

[0029] Figs. 14A and 14B. Graph illustrating the specific productivity (grams/OD) of ethanol and acetate for the wild-type and TF18 strains at 1, 2, and 5 days post-inoculation.

[0030] Fig. 15. Selectivity of 5-L fermentation using strain TF18. In the monoculture of strain TF18 only ethanol and acetate were produced and the percentage of products is shown (total C_2 products equals 100%).

DETAILED DESCRIPTION OF THE INVENTION

[0031] All publications, patents and patent applications cited herein are hereby expressly incorporated by reference for all purposes.

[0032] Methods well known to those skilled in the art can be used to construct genetic expression constructs and recombinant microorganisms according to this invention. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, *in vivo* recombination techniques, and PCR techniques. See, for example, techniques as described in Maniatis *et al.*, 1989, MOLECULAR CLONING: A LABORATORY MANUAL, Cold Spring Harbor Laboratory, New York; Ausubel et al., 1989, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, Greene Publishing Associates and Wiley Interscience, New York, and *PCR Protocols: A Guide to Methods and Applications* (Innis *et al.*, 1990, Academic Press, San Diego, CA).

[0033] As used herein, the terms "polynucleotide", "nucleotide", "oligonucleotide", and "nucleic acid" can be used interchangeably to refer to nucleic acid comprising DNA, RNA, derivatives thereof, or combinations thereof.

[0034] Provided herein are methods for increased production of an alcohol, comprising: culturing a recombinant C1-fixing homoacetogen microorganism in a culture medium in the presence of syngas, wherein the C1-fixing homoacetogen microorganism comprises a recombinant gene encoding a aldehyde ferredoxin oxidoreductase (AOR) polypeptide modulated by a promoter, wherein the recombinant C1-fixing microorganism produces an increased amount of alcohol compared to a wild-type C1-fixing homoacetogen microorganism.

[0035] AOR proteins are involved in the conversion of short-chain carboxylic acids such as acetate and butyrate to their corresponding aldehydes via a native activity in syngas-utilizing homoacetogens. The reduction of carboxylic acids to aldehydes by the AOR proteins is thermodynamically uphill, but the net reaction sequence to alcohols is favorable and easily generates enough free energy to drive it. Additionally, the reaction is driven by a low-redox ferredoxin which interacts with tungsten in the catalysis (White et al. (1989), *Eur. J. Biochem*, **184**: 89-96). This reaction has been documented to occur in several purified

enzyme preparations from Clostridia and importantly unlike aerobic carboxylic acid reductase converts the acids to alcohols via a non-activated intermediate (White et al. (1991), *Biol. Chem. Hoppe-Seyler* **372**:999-1005). In a C2 fermentation, conversion of acetate to aldehyde is important for keeping the free acid concentration from accumulating to detrimental or toxic levels. The reverse reaction (back to Acetyl-CoA) would require using an ATP to rephosphorylate acetate, which would not be energetically favorable, thus making the AOR activity very important in controlling acid levels and cell stress. Recombinant Clostridium microorganisms containing constitutively expressed AOR activity allows for expression to continue unabated even when the down regulation signals would decrease native AOR expression.

[0036] The C1- fixing microorganisms suitable for use in the methods disclosed herein are also known as homoacetogens. Homoacetogens have the ability, under anaerobic conditions, to produce acetic acid and ethanol from the substrates, $CO + H_2O$, or $H_2 + CO_2$ or $CO + H_2 + CO_2$. The CO and CO_2 provide the carbon source and the H_2 and CO provide the electron source for the reactions producing acetic acid and ethanol.

[0037] C1- fixing microorganisms suitable for use in the inventive methods include, without limitation, homoacetogens such as *Clostridium ljungdahlii, Clostridium autoethanogenum, Clostridium ragsdalei,* and *Clostridium coskatii*. Additional C1 fixing microorganisms that are suitable for use in the disclosed methods include *Alkalibaculum bacchi, Clostridium thermoaceticum*, and *Clostridium aceticum.*

[0038] As used herein, synthesis gas (syngas) is a gas containing carbon monoxide, carbon dioxide and frequently hydrogen. "Syngas" includes streams that contain carbon dioxide in combination with hydrogen and that may include little or no carbon monoxide. "Syngas" may also include carbon monoxide gas streams that may have little or no hydrogen.

[0039] In particular embodiments, plasmid vectors comprising highly active Clostridial promoters operably linked to a nucleotide sequence encoding an AOR1 and AOR2 polynucleotide are employed in the methods disclosed herein.

[0040] In particular embodiments the recombinant aldehyde ferredoxin oxidoreductase gene encodes a *Clostridium* aldehyde ferredoxin oxidoreductase polypeptide.

[0041] Functional homologs of the polypeptides described above are also suitable for use in the disclosed methods. A functional homolog is a polypeptide that has sequence similarity to a reference polypeptide, and that carries out one or more of the biochemical or physiological function(s) of the reference polypeptide. A functional homolog and the reference polypeptide can be natural occurring polypeptides, and the sequence similarity can be due to convergent or divergent evolutionary events. As such, functional homologs are sometimes designated in the literature as homologs, or orthologs, or paralogs. Variants of a naturally occurring functional homolog, such as polypeptides encoded by mutants of a wild type coding sequence, can themselves be functional homologs. Functional homologs can also be created via site-directed mutagenesis of the coding sequence for a polypeptide, or by combining domains from the coding sequences for different naturally-occurring polypeptides ("domain swapping"). Techniques for modifying genes encoding functional AOR polypeptides described herein are known and include, *inter alia*, directed evolution techniques, site-directed mutagenesis techniques and random mutagenesis techniques, and can be useful to increase specific activity of a polypeptide, alter substrate specificity, alter expression levels, alter subcellular location, or modify polypeptide:polypeptide interactions in a desired manner. Such modified polypeptides are considered functional homologs. The term "functional homolog" is sometimes applied to the nucleic acid that encodes a functionally homologous polypeptide.

[0042] Functional homologs can be identified by analysis of nucleotide and polypeptide sequence alignments. For example, performing a query on a database of nucleotide or polypeptide sequences can identify homologs of polypeptides described herein. Sequence analysis can involve BLAST, Reciprocal BLAST, or PSI-BLAST analysis of non-redundant databases using the amino acid sequence of interest as the reference sequence. Amino acid sequence is, in some instances, deduced from the nucleotide sequence. Those polypeptides in the database that have greater than 40% sequence identity are candidates for further evaluation. Amino acid sequence similarity allows for conservative amino acid substitutions, such as substitution of one hydrophobic residue for another or substitution of one polar residue for another. When desired, manual inspection of such candidates can be carried out in order to narrow the number of candidates to be further evaluated. Manual inspection can be performed by selecting those candidates that appear to have conserved functional domains.

[0043] Conserved regions can be identified by locating a region within the primary amino acid sequence of a polypeptide described herein that is a repeated sequence, forms some secondary structure (*e.g.*, helices and beta sheets), establishes positively or negatively charged domains, or represents a protein motif or domain. *See, e.g.*, the Pfam web site describing consensus sequences for a variety of protein motifs and domains on the World Wide Web at sanger.ac.uk/Software/Pfam/ and pfam.janelia.org/. The information included at the Pfam database is described in Sonnhammer *et al.*, Nucl. Acids Res., 26:320-322 (1998); Sonnhammer *et al.*, Proteins, 28:405-420 (1997); and Bateman *et al.*, Nucl. Acids Res., 27:260-262 (1999). Conserved regions also can be determined by aligning sequences of the same or related polypeptides from closely related species. Closely related species preferably are from the same family. In some embodiments, alignment of sequences from two different species can be adequate.

[0044] Typically, polypeptides that exhibit at least about 40% amino acid sequence identity are useful to identify conserved regions. Conserved regions of related polypeptides exhibit at least 45% amino acid sequence identity (*e.g.*, at least 50%, at least 60%, at least 70%, at least 80%, or at least 90% amino acid sequence identity). In some embodiments, a conserved region exhibits at least 92%, 94%, 96%, 98%, or 99% amino acid sequence identity.

[0045] In particular aspects, the AOR gene exhibits at least 60%, 70%, 80%, 92%, 94%, 96%, 98%, or 99% sequence identity to SEQ ID NO: 1 or SEQ ID NO: 3.

[0046] Further provided herein are methods for the production of butanol, comprising: culturing a co-culture comprising a recombinant C1-fixing homoacetogen microorganism and a C4-producing butyrate microorganism in a culture medium in the presence of syngas, wherein the C1-fixing homoacetogen microorganism comprises a recombinant gene encoding a aldehyde ferredoxin oxidoreductase polypeptide modulated by a promoter.

[0047] In particular embodiments the co-culture of microorganisms contains at least one microorganism having at least one nucleotide sequence that encodes a gene to produce a tungsten-dependent aldehyde ferredoxin oxidoreductase (AOR) and at least one additional microorganism that encodes a gene for producing a Butyryl-CoA acetate transferase (BuCoAAT) or a Butyrate kinase (Buk). The co-culture is exposed to gaseous substrates selected from the group consisting of carbon monoxide, carbon dioxide and hydrogen gas or combinations thereof so that a C1-fixing microorganism containing an AOR gene or genes and a C4-producing microorganism containing at least one of the BuCoAAT or Buk gene under conditions effective for the co-culture to convert the gaseous substrate into butanol or/and into butyric acid so that the microorganism composition can produce butanol.

[0048] As used herein, the term "syntrophic" refers to the association of two or more different types (e.g. organisms, populations, strains, species, genera, families, etc.) of anaerobic microorganisms which form a tightly associated metabolic relationship.

[0049] As used herein, the term "co-culture" of microorganisms refers to joint incubation or incubation together, of the microorganisms. The co-culture does not require cellular population growth during the joint incubation of the microorganisms.

[0050] In one embodiment, illustrated in Figure 6 and Figure 7, two types of anaerobic microorganism are utilized to create the co-cultures for production of butyrate and butanol. The first type of microorganism in the co-culture is a primary C1- fixing homacetogenic microorganism (illustrated metabolically in Figures 6 and 7), which utilizes syngas as the sole carbon and electron source and produces C2 compounds such as ethanol and acetate as the dissimilatory metabolite products. The second type of microorganism in the co-culture is capable of growing on the dissimilatory metabolites of the C1- fixing homacetogenic microorganism (ethanol and acetate) as its sole carbon and energy source to produce a C4-carbon molecule, such as butanol or butyric acid, as its primary product or together with syngas (as additional carbon and/or electron source) convert the metabolites of the C1-carbon fixing microorganism to C4-carbon molecules. This second microorganism is also referred to herein as the C4- producing butyrate microorganism (illustrated metabolically in Figures 6 and 7). Advantageously, the C1-fixing homacetogenic microorganism may also be capable of converting the butyrate produced by the C4-producing microorganism into butanol and more often n-butanol (Figures 6 and 7). The term "butanol" refers to all four isomers of C4 alcohol (e.g. 2-butanol, isobutanol, 1-butanol and tert-butanol) and the term "n-butanol" refers to 1-butanol.

[0051] The homoacetogenic organism typically has the primary Wood-Ljungdahl pathway to convert the CO and $H2/CO_2$ from the syngas feed to ethanol and acetate which

are then utilized by the butyrogens to produce butyrate. The homoacetogens can uptake the butyrate and very efficiently convert it to n-butanol because of favored thermodynamics. Such symbiosis is preferably developed to form a very close association between the C_1-fixing and the C_4-producing microorganisms so that interspecies proton and electron transfer occur very efficiently across very short distances (approximately 1 micron). This combination of microorganism co-culture and substrates vastly improves the n-butanol production over that produced by single culture (mono-) fermentations. This discovery enables high-yield production of butanol directly from syngas and leads to economical and efficient production processes for butanol from a wide range of feedstocks.

[0052] In particular embodiments the C4-producing microorganisms are butyrogens capable of growing on ethanol and/or acetate as their primary carbon source. Butyrogens refers to any microorganism capable of converting syngas intermediates, such as ethanol and acetate, and some hydrogen to primarily n-butyrate. Butyrogens of the invention utilize at least one of two distinct pathways for butyrate production – the Butyryl-CoA Acetate Transferase pathway (shown in Figure 6) and the Butyryl Kinase (Buk) pathway (shown in Figure 7). As can be seen from Figures 6, the Butyryl CoA Acetyl Transferase (BuCoAAT) pathway converts ethanol and acetate to butyrate:

$$\text{Ethanol} + \text{Acetate} \leftrightarrow \text{Butyrate} + H_2O$$

As shown in Figure 7, the BuK pathway converts acetate and hydrogen to Butyrate.

$$2H_2 + 2\text{Acetate} \leftrightarrow \text{Butyrate} + 2H_2O$$

[0053] In particular embodiments a recombinant homoacetogen containing unregulated AOR activity is brought into close contact with a butyrogen to further increase the production of butanol. The co-culture of microorganisms includes a unique set of nucleotide sequences that can produce butanol from the syngas components of CO or CO_2 and H_2 at much higher concentrations than previous methods for anaerobically producing butanol with

microorganisms. In other particular embodiments the homoacetogenic microorganism is cultured in a fermenter until it produces a concentration of ethanol of at least 1 g/L and the butyrogenic microorganism is added to the fermenter to produce the microorganism co-culture.

[0054] In other particular embodiments the homoacetogenic microorganism is cultured in a fermenter until it produces a concentration of ethanol of at least 1 g/L or at least 10 g/L and the butyrogenic microorganism is added to the fermenter to produce the microorganism co-culture.

[0055] In particular embodiments, *C. pharus* and *C. kluyveri* are employed. In other embodiments the co-culture includes one or more homoacetogenic microorganisms selected from the group consisting of *C. ljungdahlii, C. ragsdalei, C. autoethanongenum* and *C. coskatii.* In yet other embodiments the co-culture comprises a mixture of homoacetogenic microorganisms and a butyrogenic microorganism.

[0056] During the co-culture fermentation, the syngas-utilizing homoacetogenic organisms produce ethanol and acetate by the Wood-Ljungdahl pathway and downstream dehydrogenase reactions (Fig. 1). The butyrogenic cells consume the ethanol (by oxidation to acetyl-CoA) and the exogenous acetate produced and convert them in a condensation reaction to acetoacetyl-CoA. This reaction is carried out by thiolase, a selenium-dependent enzyme. The butyrogen converts the acetoacetyl-CoA to 3-hydroxybutyryl-CoA (3-HBCoA) by a 3-hydroxybutyryl-CoA dehydrogenase (*hbd*) using NADH(PH) as cofactor. The 3-HBCoA is dehydrated by crotonase to form crotonyl-CoA. The dehydration reaction is followed by a reduction reaction catalyzed by butyryl-CoA dehydrogenase (*bcd*) and NADH(PH) as cofactor and converts the crotonyl-CoA to butyryl-CoA. This reaction also involves ferredoxin and flavin adenine dinucleotide to generate the cell's PMF. The butyryl-CoA is then converted to butyrate via two possible routes, (i) by butyryl-CoA acetate

transferase (BCoAAT) that transfers a CoA to acetate from butyryl-CoA (Fig. 6) or (ii) by the phosphotransbutyrylase-butyrate kinase (ptb-Buk) route (Fig. 7). Both routes are possible and results of molecular probing (data not shown) using specific primers to amplify the genes encoding butyrate production have shown that both are present in the co-cultures of consortia and only the BCoAAT is present in the two organism co-culture of *C. autoethanogenum* and the butyrogen *C. pharus*.

[0057] The butyrate produced can diffuse out of the butyrogen cells and be taken up by the homoacetogen cells. Upon butyrate uptake the homoacetogen cells can convert it to butanol by either acyl-CoA transferase activity involving NADH(PH) as cofactor, CoA synthetase activity or by direct reduction of the non-activated carboxylic acid by AOR activity. The former reactions are possible but genes encoding acyl-CoA transferase and CoA synthetase activities have not been identified in our homoacetogen genomes. Two AOR genes encoding tungsten-dependent aldehyde ferredoxin oxidoreductase activity have been identified in the *C. autoethanogenum et al.* genomes (Figs. 4-5). Purified AOR activity has been demonstrated in several clostridial and non-clostridial (aerobic) organisms (White (1991), *Biol. Chem. Hoppe-Seyler* **372**:999-1005.). A genome analysis of the two AOR genes (Figs. 4-5) showed good conservation with the two well-characterized AOR enzymes at both the DNA and amino acid sequence levels, suggesting that these enzymes perform similar functions *in vivo*.

[0058] In the BCoAAT pathway ethanol and acetate are converted to butyrate through a Butyryl-CoA intermediate. Similarly, acetate plus reducing equivalents through H_2 oxidation are converted to butyrate through a butyryl-CoA intermediate. The pathways differ in their conversion steps from butyryl-CoA to butyrate. The BuCoAAT pathway converts butyryl-CoA to butyrate through the BCoAAT enzyme while transferring the CoA moiety to acetate to form acetyl-CoA, which can later be used to form more butyrate. At the same time the

Buk pathway converts butyryl-CoA through a phosphotransbutyrylase and Buk enzyme. The tungsten-dependent aldehyde ferredoxin oxidoreductase converts butyrate directly into butyraldehyde in a two electron transfer reaction using reduced ferredoxin and tungsten. The butyraldehyde is then converted to butanol using native homoacetogen butanol dehydrogenase enzymes. Suitable butyrogens for this invention include any microorganism that contains either or both of the BuCoAAT pathway and Buk pathway and can grow on acetate and ethanol or on acetate and hydrogen as typically found in syngas. While many microorganism are known to produce butyrate from various carbohydrate sources (*C. butyricum, C. acetobutylicum, C. tyrobutyricum, C. beijerinckii, C. pasteurianum, C. barkeri, C. thermobutyricum, C. thermopalmarium, Butyrvibrio, Sarcina, Eubacterium, Fusobacterium,* and *Megasphera*), only a few are known to grow exclusively on ethanol, acetate or syngas such as *Clostridium kluyveri*, *Clostridium carboxidivorans*, *Butyribacterium methylotrophicum*, and *Clostridium pharus.*

[0059] Provided herein is a combination of the genes for tungsten-dependent aldehyde ferredoxin oxidoreductase and for the genes of a Butyryl-CoA acetate transferase and/or a Butyrate kinase such that this unique gene combination can make butanol from one or more syngas components. Tungsten-dependent aldehyde ferredoxin oxidoreductase does not occur in the butyrogenic organisms nor do the Butyryl-CoA Acetate transferase or Butyrate kinase occur in the homocetogenic organism. The genetic novelty of these gene combinations was established by identifying key genes in the butyrate production pathway using targeted gene probes. The novelty of the butyryl-CoA transferase genes in the butanologenic consortia appears to be a highly specific transferase reaction and the reduction of carboxylic acids also appears to be highly specific under the culture conditions. Hence, unique combinations of genes exist in these co-cultures that do not occur in other organisms that have been used to produce butanol. Additionally and significantly, provided herein is a recombinant

homoacetogen strain with improved aldehyde ferredoxin oxidoreductase qualities and activities. During periods of AOR down-regulation, where the conversion of carboxylic acids is reduced or inhibited altogether since transcript is reduced or limited, the recombinant strain would provide unregulated gene expression and protein activity to continue the conversion of carboxylic acids to aldehydes, and thereby circumventing cell stress due to acid accumulation in the culture.

[0060] A successful syntrophic relationship between the different microorganisms require that the homoacetogens and the butyrogens are brought into close physical association with each other. In particular embodiments the C1-converting homoacetogens with the Wood-Ljundahl pathway and the tungsten-dependent aldehyde ferredoxin oxidoreductase genes AOR1 and AOR2 are brought together in an intimately mixed co-culture with the butyrogens having the BuCoAAT or the BuK genes or both sets of genes. In another embodiment of the invention the C1-converting homoacetogens will have tungsten-dependent aldehyde ferredoxin oxidoreductase genes to further increase the production of butanol in the homoacetogens. In one method of the invention, the co-culture is formed by first growing the homoacetogen species on a syngas feed. Growth of the homoacetogens continues until they produce ethanol and acetate, normally at a concentration of at least 1 g/L and more typically in a moderate concentration range of 8 to 15 g/L and preferably at a concentration of 10 g/L and a cell concentration producing an optical density (O.D.) of about 2.0. Once the homoacetogens have produced a desired concentration of ethanol and acetate and the fermenter has reached a desired O.D., the homoacetogens are inoculated with one or more selected butyrogen species that are enriched from growth on acetate, ethanol and syngas. By maintaining growth and operating conditions such as pH, dilution rate, key nutrients etc., a stable co-culture is developed that forms very close associations between the different microorganisms.

[0061] Those skilled in the art will be aware of other methods to initiate and grow the co-culture. Such methods may include the use of different substrates to first grow the butyrogen and then inoculate the fermentation medium containing the butyrogen with the homoacetogen. Another method for establishing a syntrophic association capable of converting syngas to butanol involves the growing of two or more defined cultures and establishing the pairing of these separate cultures.

[0062] Another method of pairing involves first growing the C4-producing butyrogen(s) in a fermenter using ethanol and acetate as substrates until maximum productivity targets of butyric acid has been reached. Once the maximum productivity target has been reached a seed culture of the C1-fixing homoacetogen is added directly to the fermenter containing the butyrogen culture. Syngas mass transfer to the fermentation vessel is gradually increased to balance the gas consumption of the C1-fixing homoacetogen. The ethanol or acetate used to grow the butyrogen is gradually decreased to zero as the C1-fixing homoacetogen begins to provide this substrate.

[0063] A modification of this last method of establishing a syntrophic culture involves first growing the C4-producing butyrogen culture in a fermenter with a biofilm support material that is either stationary or floating within the reactor. An example of such material is the Mutag Biochips. This method allows the butyrogen microorganism to first establish a biofilm on the carrier material thereby increasing the cell retention time versus the HRT of the fermenter. Again, target butyrogen productivity is reached before seeding the fermenter with the C1-fixing homoacetogen.

[0064] Another method to establish a syntrophic culture capable of producing butanol from syngas involves the initial mixing together of two or more cultures, one of which is a C1-fixing homoacetogen capable of growing on syngas and producing ethanol and acetate. The other culture(s) is a C4-producing butyrogen capable of converting ethanol or acetate to

butyrate. The Ethanol and acetate feed can gradually be decreased to zero as the production of these substrates by the C1-fixing homoacetogens increases to balance the substrate needs of the butyrogen production.

[0065] Suitable pairings of microorganisms for the co-culture composition of this invention are identified by the presence of key genes in the pathways for the homoacetogenic and butyrogenic microorganisms. These pathways are typically identified by using targeted gene probes. The probes are targeted toward identifying the presence of genes in the consortium that encode for two tungsten-dependent aldehyde ferredoxin oxidoreductase genes, at least one BuCoAAT gene or one Buk gene. The presence or absence of these genes can be further determined using genomic DNA and suitable probes. Further description of the gene sequences are provided in the Examples.

[0066] The methods disclosed herein can be performed in any of several types of fermentation apparatus that are known to those of skill in the art, with or without additional modifications, or in other styles of fermentation equipment that are currently under development. Examples include but are not limited to conventional stirred-tank fermenters (CSTR), bubble column bioreactors (BCBR), membrane supported bioreactors (MSBR), two-stage bioreactors, trickle-bed reactors, membrane reactors, packed-bed reactors containing immobilized cells, etc. Bioreactors may also include a column fermentor with immobilized or suspended cells, a continuous flow-type reactor, a high-pressure reactor, or a suspended cell reactor with cell recycle. Furthermore, reactors may be arranged in a series and/or parallel reactor system which contains any of the above-mentioned reactors. For example, multiple reactors can be useful for growing cells under one set of conditions and generating n-butanol (or other products) with minimal growth under another set of conditions.

[0067] Establishing the necessary close association of the co-culture may be influenced by the type of bioreactor employed for practice of the invention. For example, in the case of

planktonic type bioreactors the co-culture may continue in a growth phase and be passaged up to larger fermentation vessels. In the case of an MSBR, an established co-culture from a planktonic fermenter may be used to inoculate the membranes. However, an MSBR may also be inoculated by a series of inoculations that alternate between addition of the homoacetogen and addition of the butyrogen.

[0068] These apparatuses will be used to develop and maintain the C1-fixing homoacetogen and butyrogen cultures used to establish the metabolic association. The chief requirements of such an apparatus include:

a. Axenicity;

b. Anaerobic conditions;

c. Suitable conditions for maintenance of temperature, pressure, and pH;

d. Sufficient quantities of substrates supplied to the culture;

e. Optimum mass transfer performance to supply the gases to the fermentation medium; and

f. The end products of the fermentation can be readily recovered from the bacterial broth.

[0069] Suitable gas sources of carbon and electrons are preferably added during the inoculation. In addition to those already described these gaseous sources come from a wide range of materials and include "waste" gases such as syngas, oil refinery waste gases, steel manufacturing waste gases, gases produced by steam , autothermal or combined reforming of natural gas or naphtha, biogas and products of biomass, coal or refinery residues gasification or mixtures of the latter. Sources also include gases (containing some H_2) which are produced by yeast, clostridial fermentations, and gasified cellulosic materials. Such gaseous

substrates may be produced as byproducts of other processes or may be produced specifically for use in the methods disclosed herein. Those of skill in the art will recognize that any source of substrate gas may be used in the practice of the methods disclosed herein, so long as it is possible to provide the microorganisms of the co-culture with sufficient quantities of the substrate gases under conditions suitable for the bacterium to carry out the fermentation reactions.

[0070] In one embodiment of the invention, the source of CO, CO_2 and H_2 is syngas. Syngas for use as a substrate may be obtained, for example, as a gaseous product of coal or refinery residues gasification.

[0071] In addition to those sources as described, syngas can be produced by gasification of readily available low-cost agricultural raw materials expressly for the purpose of bacterial fermentation, thereby providing a route for indirect fermentation of biomass to alcohol. There are numerous examples of raw materials which can be converted to syngas, as most types of vegetation could be used for this purpose. Suitable raw materials include, but are not limited to, perennial grasses such as switchgrass, crop residues such as corn stover, processing wastes such as sawdust, byproducts from sugar cane harvesting (bagasse) or palm oil production, etc. Those of skill in the art are familiar with the generation of syngas from such starting materials. In general, syngas is generated in a gasifier from dried biomass primarily by pyrolysis, partial oxidation, and steam reforming, the primary products being CO, H_2 and CO_2. The terms "gasification" and "pyrolysis" refer to similar processes; both processes limit the amount of oxygen to which the biomass is exposed. The term "gasification" is sometimes used to include both gasification and pyrolysis.

[0072] Combinations of sources for substrate gases fed into the fermentation process may also be utilized to alter the concentration of components in the feed stream to the bioreactor. For example, the primary source of CO, CO_2 and H_2 may be syngas, which

typically exhibits a concentration ratio of 37% CO, 35% H_2, and 18% CO_2, but the syngas may be supplemented with gas from other sources to enrich the level of CO (i.e., steel mill waste gas is enriched in CO) or H_2.

[0073] The co-cultures disclosed herein are cultured and used under anaerobic conditions. As used herein, "anaerobic conditions" means the level of oxygen (O_2) is below 0.5 parts per million in the gas phase of the environment to which the microorganisms are exposed. One of skill in the art will be familiar with the standard anaerobic techniques for culturing these microorganisms (Balch and Wolfe (1976) *Appl. Environ. Microbiol.* 32:781-791; Balch et al., 1979, *Microbiol. Rev.* 43:260-296), which are incorporated herein by reference. Other operating conditions for the established co-culture will usually include a pH in a range of 5 to 7.

[0074] A suitable medium composition used to grow and maintain co-cultures or separately grown cultures used for sequential fermentations, includes a defined media formulation. The standard growth medium is made from stock solutions which result in the following final composition per liter of medium. The amounts given are in grams unless stated otherwise. Minerals: NaCl, 2; NH_4Cl, 25; KCl, 2.5; KH_2PO_4, 2.5; $MgSO_4 \cdot 7H_2O$, 0.5; $CaCl_2 \cdot 2H_2O$, 0.1. Trace metals: $MnSO_4 \cdot H_2O$, 0.01; $Fe(NH_4)_2(SO_4)_2 \cdot 6H2O$, 0.008; $CoCl_2 \cdot 6H2O$, 0.002; $ZnSO_4 \cdot 7H2O$, 0.01; $NiCl_2 \cdot 6H2O$, 0.002; $Na_2MoO_4 \cdot 2H_2O$, 0.0002; Na_2SeO_4, 0.001; Na_2WO_4, 0.002. Vitamins (in mg): Pyridoxine-HCl, 0.10; thiamine-HCl, 0.05; riboflavin, 0.05; calcium pantothenate, 0.05; thioctic acid, 0.05; p-aminobenzoic acid, 0.05; nicotinic acid, 0.05; vitamin B_{12}, 0.05; mercaptoethane sulfonic acid, 0.05; biotin, 0.02; folic acid, 0.02. A reducing agent mixture is added to the medium at a final concentration of 0.1 g/L of cysteine (free base); and 0.1 $Na_2S \cdot 2H_2O$. Medium compositions can also be provided by yeast extract or corn steep liquor or supplemented with such liquids.

[0075] In general, fermentation of the co-culture will be allowed to proceed until a desired level of butanol is produced in the culture media. Preferably, the level of butanol produced is in the range of 2 g/L to 7.5 g/L and most preferably in the range of 6 g/L to 15 g/L. Alternatively, production may be halted when a certain rate of production is achieved, e.g. when the rate of production of a desired product has declined due to, for example, build-up of bacterial waste products, reduction in substrate availability, feedback inhibition of by-products, reduction in the number of viable bacteria, or for any of several other reasons known to those of skill in the art. In addition, continuous culture techniques exist which allow the continual replenishment of fresh culture medium with concurrent removal of used medium, including any liquid products therein (i.e. the chemostat mode). Also, techniques of cell recycle may be employed to control the cell density and hence the volumetric productivity of the fermenter.

[0076] The products that are produced by the microorganisms of this invention can be removed from the culture and purified by any of several methods that are known to those of skill in the art. For example, butanol can be removed by distillation at atmospheric pressure or under vacuum, by adsorption or by other membrane based separations processes such as pervaporation, vapor permeation and the like.

[0077] This invention is more particularly described below in the Examples set forth herein and are intended as illustrative only, as numerous modifications and variations therein will be apparent to those skilled in the art. As used in the description herein and throughout the claims that follow, the meaning of "a", "an", and "the" includes plural reference unless the context clearly dictates otherwise. The terms used in the specification generally have their ordinary meanings in the art, within the context of the invention, and in the specific context where each term is used. Some terms have been more specifically defined to provide additional guidance to the practitioner regarding the description of the invention.

EXAMPLES

Genome sequence analysis.

[0078] Analysis of the *C. autoethanogenum, C. ragsdalei, C. ljungdahlii, C. coskatii* and *C. carboxidivorans* genome sequences was performed using the ERGO bioinformatics suite (Integrated Genomics, Arlington Heights, IL) and BLASTX. The genome analysis revealed two genes annotated as having tungsten-dependent, aldehyde ferredoxin oxidoreductase activity (Ec 1.2.7.5).

Strains

[0079] *Clostridium autoethanogenum* strain DSM 10061 (Abrini J, Naveau H, Nyns E-J, *Arch. Microbiol.* 1994, 4:345-351), *C. ragsdalei* P11 (ATCC BAA-622) (WO 2008/028055), *C. coskatii* PS-02 (ATCC-PTA10522) (US 2011/0229947), *C. carboxidivorans* strain "P7" (ATCC-BAA-624) (US 2007/0275447), *C. ljungdahlii* was obtained from ATCC (strain ATCC 55383 / DSM 13528 / PETC; Tanner RS et al. *Int J Syst Bacteriol.* (1993) 43(2):232-236).

[0080] E. coli F^- JW5272-1 Δ(araD-araB)567, ΔlacZ4787(::rrnB–3), λ^-, ΔydhV748::Kan, rph-1, Δ(rhaD-rhaB)568, hsdR514. Baba et al (2006) Mol. Syst. Biol. 2:1-11 (The Keio Collection). The E. coli AOR mutant was obtained from the Yale Culture Stock Collection (New Haven CT). E. coli DH10B (Life Technologies, Carlsbad, CA; dam^+, dcm^+, mcrABmrr$^-$) and E. coli BL21 (Life Technologies, Carlsbad, CA; dcm^-, dam^+, mcrAB mrr^+) were routinely used for propagating replicating plasmids.

Vectors

[0081] Suitable vectors are described in Table 1 with some general features. The vectors were used to express Clostridial (e.g. *C. autoethanogenum*) aldehyde ferredoxin

oxidoreductase genes in an *E. coli ydhV* mutant to demonstrate butyrate conversion to butanol function *in vivo*. High-copy expression vector pCKAR192 contained AOR2 expressed from the strong Clostridial promoter PcooS. *E. coli-Clostridium* conjugative vector pCKAR192T2 contained the same expression construct as pCKAR192 but also expressed the basis of mobilization derived from RP4 (Parke 1990, *Gene* 93:135-137). High-copy expression vector pCKAR214 contained AOR1 expressed from the strong Clostridial promoter Ppta-ack. *E. coli-Clostridium* conjugative vector pCKAR214T1 contained the same expression construct as pCKAR214 but also expressed the basis of mobilization derived from RP4 (Parke 1990, *Gene* 93:135-137). *E.coli-clostridium* shuttle vector pCK232AOR1, contained the same elements as pCKAR214 but additionally contained the rep1 origin of replication and the Ppta-AOR1 expression cassette from *C. autoethanogenum*. pCK232 lacks the Ppta-AOR1 expression cassette.

Table 1. Plasmids used and their key gene elements.

Plasmids	Key genetic elements	Antibiotic resistance
pCKAR192	pMB1, Pcoos-AOR	Ampicillin
pCKAR192T2	pMB1, RP4 oriT, Pcoos-AOR	chloramphenicol, Ampicillin
pCKAR214	pMB1, Ppta-AOR	Ampicillin
pCKAR214T1	pMB1, RP4 oriT, Ppta-AOR	chloramphenicol, Ampicillin
pCK232	pMB1, rep1, RP4 oriT	Kanamycin, Ampicillin
pCK232AOR1	pMB1, rep1, RP4 oriT, Ppta-AOR	Kanamycin, Ampicillin

[0082] *Media*

[0083] *E. coli* was grown in LB medium which was composed of 10 g/L tryptone, 5 g/L yeast extract and 10 g/L NaCl.

[0084] LB ampicillin and LB chloramphenicol medium was composed of LB medium containing filter sterilized 100 μg ampicillin sulfate and 50 μg water soluble chloramphenicol per ml, respectively.

[0085] *E. coli* was grown on 2xYT agar which was composed of 16 g per liter tryptone, 10 g per liter yeast extract and 5.0 g per liter NaCl.

[0086] *E. coli* minimal medium (M9) contained the following per liter: 200 ml 5x M9 salts; 2 ml 1 M $MgSO_4$; 20 ml of a sterile 20% glucose solution; 100 ul of 1 M $CaCL_2$. The 5x M9 salts contained the following (per liter): Na_2HPO_4:$7H_2O$, 64 g; KH_2PO_4, 15 g; NaCl, 2.5 g; NH_4CL, 5 g.

[0087] For M9 medium containing ampicillin or chloramphenicol, 50 ml M9 medium was prepared as above and was supplemented with ampicillin sulfate or water soluble chloramphenicol at 50 μg per liter medium.

[0088] 2xYT ampicillin (Ap) and chloramphenicol (Cm) were composed of 2xYT agar (1.5%) containing 100 μg (Ap) and 50 μg (Cm) respectively, per ml of agar.

[0089] *Clostridium* fermentation medium was made anaerobically from concentrated vitamin, mineral and metals stocks with the compositions shown in Tables 2 and 3a-d.

[0090] *Clostridium* plating medium was composed of fermentation medium containing 5 g/L fructose (filter-sterilized and added post autoclaving), 10 g/L yeast extract and 15-20 g/L agar.

Table 2. Fermentation Medium Compositions

Components	Amount per liter
Mineral solution, See Table 2(a)	25 ml

Trace metal solution, See Table 2(b)	10 ml
Vitamins solution, See Table 2(c)	10 ml
Yeast Extract	0.5 g
Adjust pH with NaOH	5.8
Reducing agent, See Table 2(d)	2.5 ml

Table 3(a). Mineral Solution

Components	Concentration (g/L)
NaCl	80
NH_4Cl	100
KCl	10
KH_2PO_4	10
$MgSO_4 \cdot 7H_2O$	20
$CaCl_2 \cdot 2H_2O$	4

Table 3(b). Trace Metals Solution

Components	Concentration (g/L)
Nitrilotriacetic acid	2.0
Adjust the pH to 6.0 with KOH	
$MnSO_4 \cdot H_2O$	1.0
$Fe(NH_4)_2(SO_4)_2 \cdot 6H_2O$	0.8
$CoCl_2 \cdot 6H_2O$	0.2
$ZnSO_4 \cdot 7H_2O$	1.0
$NiCl_2 \cdot 6H_2O$	0.2

$Na_2MoO_4 \cdot 2H_2O$	0.02
Na_2SeO_4	0.1
Na_2WO_4	0.2

Table 3(c). Vitamin Solution

Components	Concentration (mg/L)
Pyridoxine·HCl	10
Thiamine·HCl	5
Riboflavin	5
Calcium Pantothenate	5
Thioctic acid	5
p-Aminobenzoic acid	5
Nicotinic acid	5
Vitamin B_{12}	5
Mercaptoethanesulfonic acid	5
Biotin	2
Folic acid	2

Table 3(d). Reducing Agent

Components	Concentration (g/L)
Cysteine (free base)	40
$Na_2S \cdot 9H_2O$	40

[0091] **Example 1.** Acetate assimilation pathways in syngas-utilizing Clostridia.

Metabolic reconstruction of the Wood-Ljungdahl pathway from genome sequences of

syngas-utilizing Clostridia (including *C. autoethanogenum*) identified three possible routes to carboxylic acids and alcohols (Fig. 1). The product of the Wood-Ljungdahl pathway is acetyl-CoA which results from the three-component reaction of CoA, methyl-corrinoid iron sulfur protein and CO of the carbon monoxide dehydrogenase acetyl-CoA synthase. Acetyl-CoA can be directly converted to ethanol by the acetyl-CoA reductase (US patent 8,039,239), a bifunctional enzyme requiring NADPH and zinc in a four-electron reduction reaction. It can alternatively be converted to the aldehyde by an acetylating aldehyde dehydrogenase. The aldehyde can subsequently be converted to ethanol by variety of alcohol dehydrogenases. An alternate route for acetate assimilation involves the conversion of acetyl-CoA to acetyl-phosphate by phosphotransacetylase which in turn generates acetate by acetate kinase. The latter reaction generates an ATP which potentially acts as an important valve for energy balance by substrate-level phosphorylation when cells are growing on dissolved gases. While generation of an ATP is beneficial to the cell, by balancing the ATP used in the generation of formyl-tetrahydrofolate of the methyl branch of the pathway, acetate accumulation is detrimental to cells particularly when the free acid concentration reaches a threshold level, typically at about 5 g/L when fermentations are run at or near the pKa of the acids (~4.75). Another route to acetate assimilation can occur by conversion to the aldehyde using aldehyde ferredoxin oxidoreductase. Two genes annotated as tungsten-dependent, aldehyde ferredoxin oxidoreductase (1.2.7.5) were identified in the *C. autoethanogenum et al.* genomes and their DNA and amino acid sequences are shown in Figs. 3 and 4, respectively. These genes encode a somewhat unusual function by performing the conversion of a non-activated carboxylic acid (e.g. acetate and butyrate) directly to an aldehyde, thus potentially extending the metabolic versatility of *C. autoethanogenum et al.* to generate alcohols. While this reaction is thermodynamically uphill, the coupling of the downstream dehydrogenation reaction makes it very favorable. The aldehydes formed in turn get converted to alcohol

using nicotinamide cofactors (NADH(PH) by the wide variety of alcohol dehydrogenases identified in the *C. autoethanogenum et al.* genomes. While the conversion of an acid to an aldehyde is endergonic and likely requiring high energy electrons to be transferred to ferredoxin, this strategy would also reduce acid stress and the consequent dissipation of the chemiosmotic potential, which is known to be the major source of energy for these organisms. Moreover, the assumed *in vivo* ratio of acid to aldehyde is predicted to be very high and thus would also favor the conversion of the acid to alcohol by the aldehyde ferredoxin oxidoreductase. The energetics of the reaction requiring ferredoxin suggest that a high energy electron carrier is required for these reactions to take place but recent experimental data suggest that these anaerobes have ways of altering the redox requirements to achieve highly desirable endergonic reactions by redox coupling and bifurcation (Wang et al. *J. Bacteriol.* 195 (19): 4373-4386).

[0092] **Example 2:** Construction of AOR2 expression plasmid pCKAR192. To confirm functional activity of the open-reading frames (ORF) of the AOR2 gene a novel construct was generated whereby the AOR2 ORF was operably linked to a constitutively expressed, strong heterologous Clostridial promoter, Pcoos (derived from the carbon-monoxide-electron transfer protein region US Patent 8,039,239). The 2.307 kb Pcoos-AOR2 expression cassette was synthesized with *Sph*I, *Hin*dIII and *Kpn*I engineered at the 5' and 3' ends and ligated to a modified pMB1-based high-copy number vector for replication in *E. coli*. Ampicillin was used as the selection agent at 100 ppm (Fig. 2A). This expression vector was designated pCKAR192 (Table 1).

[0093] **Example 3:** Construction of *E. coli-Clostridium* shuttle vector pCK192T2. For generation of an *E. coli-Clostridium* shuttle vector, pCKAR149 was used as the cloning vehicle. pCKAR149 contains dual replicons for replication in *E. coli* and *C. autoethanogenum et al.* and dual antibiotic resistances. The resistance gene for Clostridium

selection, chloramphenicol, was expressed from a strong, constitutively-expressed native promoter and can be used as primary selection in *C. autoethanogenum*. The 2.307 kb Pcoos-AOR2 cassette was released from pCKAR192 with *SphI* and ligated into *SphI*-digested pCKAR149. Putative recombinant plasmids were screened and confirmed by restriction digestion using *Sph*I and PCR using primers that amplified the cassette (Fig. 2A). The cassette was also subjected to double-strand gene sequencing to confirm authenticity of the Pcoos-AOR2 DNA. The final construct was designated pCKAR192T2 (Table 1).

[0094] **Example 4**: Construction of Ppta-AOR1 expression plasmid pCKAR214. For AOR1 expression a novel construct was generated using the strong clostridial promoter Ppta (derived from the phosphotransacetylase/acetate kinase region) operably linked to the AOR1 ORF. The 2.298 kb Ppta-AOR1 expression cassette was synthesized and ligated to a modified pMB1-based high-copy number vector for replication in *E. coli*. Ampicillin was used as the selection agent at 100 ppm. The restriction enzyme site *Eco*RI was engineered at the 5' ends for downstream sub-cloning into suitable *E. coli-Clostridium* shuttle vectors. The final construct expressing the AOR1 gene was designated pCKAR214 (Fig. 3, left; Table 1).

[0095] **Example 5:** Construction of Ppta-AOR1 *E. coli-Clostridium* shuttle vector, pCKAR214T1. For generation of an *E. coli-Clostridium* shuttle vector containing Ppta-AOR1, pCKAR149 was used as the cloning vehicle as described in Example 3. It contains dual replicons for replication in *E. coli* and *C. autoethanogenum* and dual antibiotic resistances as well as the basis of mobilization for DNA transfer via conjugation but can also be electroporated into recipient cells. The 2.307 kb Ppta-AOR1 cassette was released from pCKAR214 with *Eco*RI and ligated to *Eco*RI-digested pCKAR149. Putative recombinant plasmids were selected on chloramphenicol, screened and confirmed by restriction digestion using *Eco*RI and PCR using primers that amplified the cassette. The cassette was also

subjected to double-strand gene sequencing to confirm authenticity of the Ppta-AOR1 DNA. The final construct was designated pCKAR214T1 (Fig. 3B; Table 1).

[0096] **Example 6:** Butyrate production using the butyryl-CoA acetate transferase pathway. The butyrogens in co-culture with *C. autoethanogenum* have two pathways for the generation of butyrate. Ethanol-oxidizing butyrogens such as *C. kluyveri* will consume six moles ethanol and 4 moles acetate to form 5 moles acetoacetyl-CoA via thiolase (2.3.1.16). The acetoacetyl-CoA is then reduced to 3-hydroxybutyryl-CoA in a reaction catalyzed by 3-hydroxybutyryl-CoA dehydrogenase (1.1.1.157). 3-hydroxybutyryl-CoA is dehydrated by crotonase (4.2.1.17) to form crotonyl-CoA. The conversion of crotonyl-CoA involves electron bifurcation and a complex of ferredoxin, NADH, and FAD to generate the chemiosmotic potential essential for cell viability (Wang et al., *J. Bacteriol.* 192:5115-5123, 2010). This reaction is carried out by butyryl-CoA dehydrogenase (1.3.8.1 or formerly 1.3.99.2) and the EtfAB proteins. The final step to butyrate production is shown in Fig. 6. and involves a transferase reaction that converts free acetate to acetyl-CoA with the release of butyrate. One of the butyrogen (*C. pharus*) genomes has been sequenced and the gene encoding a butyryl-CoA acetate transferase activity (2.8.3.8) has been identified (Seq. ID No. 7).

[0097] **Example 7:** Butyrate production using the phosphotransbutyrylase-butyrate kinase pathway. A second pathway butyrogens use for butyrate production involves phosphotransbutyrylase and butyrate kinase (Fig. 7). Some organisms contain both butyryl-CoA acetate transferase and butyrate kinase genes such as identified in the *Clostridium carboxidivorans* genome sequence (A. Reeves unpublished). Other butyrogenic organisms have one or the other system. The steps in the formation of butyryl-CoA are more or less identical in the two systems and vary only in the final steps of butyrate production (Figs. 6-7). Butyryl-CoA is converted to butyryl-phosphate catalyzed by phosphotransbutyrylase

(2.3.1.19). Butyrate kinase (2.7.2.7) converts butyryl-P to butyrate while generating ATP, similar to the acetate kinase reaction. In our butanol-producing co-cultures the genes encoding phosphotransbutyrylase and butyrate kinase activity have been detected using gene-specific primers.

[0098] Example 8: AOR1, AOR2, BCoAAT, and Buk gene amplification with co-culture DNA. PCR primers were designed from the genome sequences of AOR1 (SEQ ID NO: 1), AOR2 (SEQ ID No: 3), BCoAAT (SEQ ID NOS: 7-8), and Buk (SEQ ID NOS. 9-11) to confirm the presence or absence of these genes in mono and co-culture fermentations. The primer sequences were: AOR1-A-forward (SEQ ID NO: 19): 5'-ACTTGGATTATGTATTTTTACA-3'; AOR1-A-reverse (SEQ ID NO: 20): 5'-TGAACTATCTATGCCTGCTTTT-3'; AOR2-A-forward (SEQ ID NO: 21): 5'-AAGAAAGAACTTGCAAATCA-3'; AOR2-A reverse (SEQ ID NO: 22): 5'-CGGAGCTCCAGTTAAAGGA-3'; BCoAAT forward (SEQ ID NO: 23): 5'-AGCCATGCTAGC TCCTCTCATGTA-3';BCoAAT reverse (SEQ ID NO: 24): 5'-GGAGTATCAACCGAT TATTCACAG-3'; Buk forward (SEQ ID NO: 25): 5'-GATGGTTCTACACTT CAGCTT-3'; and Buk reverse (SEQ ID NO: 26): 5'-GAT ATC ATT TCT GAA TGT ATA CCC-3'. Primers were targeted to unique regions within each AOR gene where little or no overlap occurred so that amplicons would yield unique products. For the butyryl-CoA acetate transferase and butyrate kinase genes primers were designed from multiple sequences with degeneracies to allow amplification of the broadest variety of butyrate-producing organisms in the consortium. The cycling reactions tested a gradient of temperatures from 45°C to 61°C. For all primer sets tested, the temperature range of 53°C-56°C gave strong amplicons with no observable background products seen on agarose gels or using melting curve analysis in qPCR. The cycling conditions were: 1 cycle at 94°C for 3 minutes followed by 30 cycles at 94°C for 1 minute to denature gDNA, 30 seconds annealing

at a gradient of 45°C-61°C and an extension at 72°C for 30 seconds. The final extension step was at 72°C for 7 minutes. The results of PCR using the four best primer sets and consortia gDNA are shown in Fig. 8. For AOR1, using primer set AOR1-A the amplicon was the expected size of 198 bp (Fig. 8, top, left) and for AOR2, primer set AOR2-A gave a unique amplicon of the expected size of 201 bp (Fig. 8, bottom, left). The butyrogen primer sets BCoAAT and Buk were also tested with DNA extracted from the co-culture fermentation. The BCoAAT primer set gave the expected size amplicon of 715 bp (Fig. 8, top, right) while the Buk primer set gave the expected amplicon of 168 bp (Fig. 8, bottom, right).

Example 9: Monoculture fermentation conditions for testing AOR1 and AOR2 gene expression. A 10-liter syngas fermentation was run containing *C. autoethanogenum* DSM 10061 for more than 2000 hours (Fig. 9). The *C. autoethanogenum* culture reached an optical density of 2.2 after several hundred hours growing on minimal medium containing syngas with a composition of H_2-64.7%, CO-16.2%, CO_2-11.7%, and CH_4-6.04% (mol%), a gas flow rate of 300-500 mL/min with agitation between 500-600 rpm, and a HRT of 2.5 days during the time period for which samples for gene expression were collected. The ethanol concentrations fluctuated between 20 and 30 g/L and the acetate concentrations fluctuated between 5 and 10 g/L over the course of the fermentation. Samples for AOR1 and AOR2 gene expression were taken at nine different time points between 1700 and 2050 hours into the fermentation when the volumetric H_2 gas uptake started to fluctuate considerably (Fig. 9). The primers sets and sequences used for gene expression analysis (presented as fold change) shown in the figure are: 467/469, AOR2-forward (SEQ ID NO: 21): 5'-AAGAAAGAACTTGCAAATCA-3'; AOR2-reverse (SEQ ID NO: 22): 5'-CGGAGCTCCAGTTAAAGGA-3'; 474/478, AOR1-forward (SEQ ID NO: 19): 5'-ACTTGGATTATGTATTTTTACA-3'; AOR1-reverse (SEQ ID NO: 20): 5'-

TGAACTATCTATGCCTGCTTTT-3'; 441, iron-only hydrogenase-forward (SEQ ID NO: 27): 5'-TGT GAA CGT CCT GAA ATG AAA G-3'; iron-only hydrogenase reverse (SEQ ID NO: 28): 5'-AGT GCC TGC ACC AGA ATA AGT T-3'.

Example 10: AOR1 and AOR2 gene expression in a *C. autoethanogenum* monoculture fermentation. During most of the 10-L syngas fermentation using a monoculture of *C. autoethanogenum* the volumetric productivity rate remained high (>0.4 g/L/h). Samples were taken between 1700 and 2050 hours for RT-PCR analysis using primers AOR1-A and AOR2-A. Additionally, four *C. autoethanogenum* hydrogenase genes were tested for gene expression during the time course fermentation to see how they trended with AOR gene expression (Fig. 9). Cells were quickly pelleted and re-suspended in RNA protect solution for 5 minutes and quick frozen in liquid nitrogen. Frozen cells were kept at -80 ° C until needed. Total RNA was extracted using the RNA easy kit (Qiagen, Valencia, CA). A DNAse treatment step was included to ensure that all genomic DNA was degraded before cDNA was generated. A total cDNA reaction was performed on 50 ηg of RNA and purified using column chromatography. Fifty ηg of purified cDNA from nine sample points were used in RT-PCR (in triplicate reactions) using an Eppendorf Realplex with built-in statistical software (Eppendorf NA, Happauge, NY) and primer sets AOR1-A (primer set 474/478), AOR2-A (primer set 467/469), Hyg-1 (primer set 431), Hyg-2 (primer set 435), Hyg-3 (primer set 437) and Hyg-4 (primer set 441). The 16s rDNA and *rec*A genes were used as reference and internal standards, respectively (data not shown). Primer sequences 474/478, 467/469, and 441 see Example 9. Primer sequences for the three additional hydrogenases are: Hyg-1 (431-forward) (SEQ ID NO: 29): 5'-GCCCGATATAAATCCTCTTT-3'; (431-reverse; (SEQ ID NO: 30): 5'-CCAACAAAAATTCCATGATT-3'; Hyg-2 (435-forward; (SEQ ID NO: 31): 5'-CTACAATTTTAAACGCTGCA-3'; (435-reverse; (SEQ ID NO: 32)): 5'-GCTCTGGCACTGTTTGTTCTA-3'; Hyg-3 (437-forward; (SEQ ID NO: 33): 5'-TGA

TAC AAA CTT TGG TGC AG-3'; and (437-reverse; (SEQ ID NO: 34): 5'-ATA TAG CTC CAG CCA TCT GA-3'.

[0099] During the 350-hour sampling period, the transcript levels were normalized using 16s RNA as a reference gene allowing only fold changes to be determined. Relative expression levels were discernible and trends determined indicating the level of expression during the course of the sampling period. During the first 150 hours of sampling both AOR genes are expressed at a higher level (2-6-fold) than the preceding period. During this time the H_2 gas uptake was declining rapidly. The increase in AOR gene expression is followed by a 150-hour period of a discernible decrease in expression of AOR2 gene and a slight change in AOR1 gene expression. H_2 gas uptake stabilized during most of this period before increasing significantly. Preceding the AOR2 decline was a dramatic H_2 gas uptake decline. At 2000 hours both AORs are dramatically upregulated along with the H_2 gas uptake before the fermentation ceased. The cell density did not change significantly during this fermentation period whereas gas uptake fluctuated indicating a regulated response of the AOR genes to process parameters and metabolite production (Fig. 9).

[00100] **Example 11:** Co-culture fermentation conditions for testing AOR gene expression. A 100-liter fermentation experiment was run in order to establish a syntrophic pairing of a type strain homoacetogen, *C. autoethanogenum*, and a mixed culture of butyrogens. The fermentation was seeded with 8.5 L of a previously combined co-culture of the homoacetogen, *C. autoethanogenum*, and butyrogen that had already been running under stable conditions. The fermentation was maintained on minimal medium (Table 2) and fed syngas with a composition of H_2-64.2%, CO-16%, CO_2-12.4%, and CH_4-5.9% (mol%), a gas flow rate of 2400-2900 mL/min with an agitation of between 360 and 500 rpm. Cell recycle was used on this fermentation with a hydraulic retention time of 1.6 days and a mean cell retention time of 2.5 days. The ethanol and acetate concentrations during the 280 hr

sampling period were between 2-3.5 g/L and 2.5-4 g/L, respectively, prior to the decline in H_2 gas uptake. Butanol and butyrate concentrations of 5-7.4 g/L and 2.1-4 g/L, respectively, were achieved during this same period prior to H_2 gas uptake decline. On average more than 70% of the syngas electrons consumed were converted to butanol and butyrate over the sampling period. For RT-PCR experiments, six samples were taken during the co-culture fermentation from 144-1008 hours (Fig. 10).

[00101] Example 12: AOR2 gene expression in a butanol co-culture fermentation. The *C. autoethanogenum* AOR2 gene was highly expressed (~1 x 10^9 copies/µg RNA) throughout the 860 hours of the sampling period until the end of the fermentation but was noticeably down-regulated from 480 hours into the fermentation until the end of the fermentation (Fig.10; primer set 467/469). The volumetric gas uptake and most other measured parameters were in decline including ethanol, butanol and butyrate production during this period. The acetate concentration increased significantly (>10 g/L) which could be explained by the lowered AOR expression (i.e. poor carboxylic acid conversion). Gene expression analysis was also performed on the six identified *C. autoethanogenum* hydrogenase genes (primer sets 431, 433, 435, 437, 440, 441). All six hydrogenases showed the same general expression pattern (data only shown for 431, 435, 437, 441). During the first 150 hours of sampling (from 144-294 hours into the fermentation) the H_2 expression increased slightly as gas consumption increased. From 480-720 hours into the fermentation expression of hydrogenase expression noticeably decreased. The hydrogenase expression levels remained flat from 720 hours to the end of the fermentation while the AOR expression levels continued to decrease (Fig. 10).

[00102] Example 13: *E. coli* growth conditions for testing heterologous AOR activity. *E. coli* JW5272-1 is an aldehyde ferredoxin oxidoreductase mutant containing an inserted Kn-resistance gene in the *ydhV* gene. *ydhV* is part of a seven-gene operon (*ydhY-ydhT*)

encoding a heptameric oxidoreductase that is responsive to FNR, the anaerobic metabolism regulator. It is not known what the substrate(s) are involved in the redox reactions but there is an apparent defect in the anaerobic metabolism of five sulfur sources (Partridge J, *.et al.*, 2008, *Microbiology,* 154:608-618). While annotated as such, the phenotype does not appear to be involved in the conversion of short-chained carboxylic acids to aldehydes. Nonetheless, pCKAR192 was transformed into *E. coli* JW5272-1 to remove the possibility of any putative background aldehyde ferredoxin oxidoreductase activity. M9 minimal medium was used throughout to grow the *E. coli* strain containing pCKAR192 but the medium was supplemented with the following nutrients to take into account the likely anaerobic requirements for highly expressed AOR activity: 1x trace metals mix as described in Table 3b, 1x vitamin mix as described in Table 3c, Na_2WO_4 at a final concentration of 20 μM, $NaNO_3$ at a final concentration of 100 μM, and Na_2S at a final concentration of 1 mM. For testing the conversion of acetate and butyrate to alcohols both were added as the sodium salt at a final concentration of 22 mM and 27 mM, respectively. Cells were first grown to mid-log phase (OD ~0.8) anaerobically using glucose as the sole carbon and energy source in M9 minimal medium without supplementation before transfer into the same medium containing the supplements described above. Either acetate or butyrate was added to the *E. coli* cultures containing plasmids pCKAR162 and pCKAR192. Cultures contained ampicillin sulfate at 20 ppm for plasmid maintenance. Cells were grown under strictly anaerobic conditions without shaking with added sodium sulfide (up to 2 mM throughout the testing period) to remove any traces of oxygen.

[00103] Example 14: Heterologous expression of Ppta-AOR2 in *E. coli*. *E. coli* strains pCKAR162 and pCKAR192 grew well under the anaerobic conditions described in Example 13. Cells reached a maximum OD_{600} after about 30 hours. Results of GC analysis of culture metabolites are shown in Fig. 11 for duplicate strains grown either in the presence or absence

of acetate and butyrate. When *E. coli* strain pCKAR162 was grown only with glucose as sole carbon and energy source it made about 0.25 g/l ethanol and about 1.75 g/l acetate after 24 hours. The strain containing the Ppta-AOR gene (*E. coli* strain pCKAR192) made the same amount of ethanol and acetate. When trace metals were added along with sodium acetate the production of ethanol increased. The same results were observed with *E. coli* strain pCKAR192 when trace metals and sodium acetate were added to the medium. When 27 mM sodium butyrate was added (~3 g/l) to *E. coli* pCKAR162 none of the butyrate was converted to the alcohol. Some acetate and ethanol were observed as expected since *E. coli* produces these metabolites through known metabolic pathways. In *E. coli* pCKAR192 supplemented with 27 mM of sodium butyrate and tungsten some of the butyrate was converted to 1-butanol at a final concentration of about 0.5 g/L. The butanol production was plasmid dependent and also metal dependent, since no butanol was observed with *E. coli* strain pCKAR192 containing plasmid but no trace metals (Fig. 11). These results confirm the functional activity of the Ppta-AOR expression cassette through the conversion of a non-activated C_4 acid requiring strictly anaerobic conditions and very low redox. The fact that *E. coli* strain pCKAR162 did not make any butanol under any of the growth conditions, while having a native CoA synthetase gene, indicates that AOR gene function was solely responsible for converting the supplemented acid to the aldehyde using a non-activated substrate.

[00104] Example 15: Vectors for the generation of a recombinant Clostridial biocatalyst containing constitutively expressed aldehyde ferredoxin oxidoreductase. The AOR genes have been shown to be regulated at the transcriptional level and can vary with the growth stage (Figs. 9, 10 and Coskata, unpublished results). Efficient conversion of carboxylic acid to aldehyde in mono- and co-culture by the AOR enzymes has been observed in 100-L fermentations over more than 1000 hours and is predicted to be closely correlated to

maintenance of low free acid concentrations. In this regard, down-regulation of AOR expression could be correlated to an increase in the culture free acid concentration (Fig. 9) by an as yet unknown mechanism. Both AOR genes appear to be regulated similarly (although at different absolute transcript levels). A vector construction strategy was developed that involved autonomously replicating plasmids containing the Ppta-ack operon promoter expressing AOR genes. This vector would express one or both of the AOR genes and be fully operative at the enzymatic level when the native genes are down-regulated, thereby potentially reducing free acid accumulation, which can be deleterious to cells and fermenter productivity.

[00105] *E. coli-Clostridium* shuttle vectors were constructed by first using a mobilizable plasmid, pARO190 (Parke, 1990), as the backbone. pARO190 contains the basis of mobilization region from pSUP2020 and the origin of transfer from RP4 (Parke, 1990). A kanamycin-resistance gene expressed from a moderately strong Clostridium promoter, Psadh, was cloned into the *Pst*I-*Sph*I sites of pARO190. This generated a vector designated pCKAR28, which is expressed in high copy number in *E. coli*. The pMB1 origin contained on pCKAR28 is not functional in Clostridium and thus a suitable replicon able to autonomously replicate in Clostridium was required. Scanning the closed genome sequence of *C. autoethanogenum* (Coskata, unpublished results) identified two contigs, one of which was the chromosome of >4.3 Mb and the other a 5.456 kb autonomously replicating plasmid. ORF analysis revealed an origin of replication related to the rep1 superfamily involved in a rolling-circle replication mechanism. A 1.5 kb PstI region containing the ~800 bp rep1-homologous region and 700 bp of surrounding DNA was cloned into the unique *Pst*I site contained on pCKAR28. This *E. coli-Clostridium* shuttle vector was designated pCK231 (counterclockwise orientation with respect to the kanamycin resistance gene) and pCK232 (clockwise orientation; Fig. 12, left).

[00106] To clone the constitutively expressed AOR1 gene, plasmid pCKAR214, which contains a synthesized Ppta-AOR gene cassette (Fig. 3), was used as template for PCR. PCR primers containing engineered *Bam*HI sites at each 5' end were used to amplify a 2.385 kb fragment containing the promoter, Ppta-ack, the AOR1 gene and downstream sequences including its transcriptional terminator. After amplification, gel purification, and isolation the 2.385 kb product was digested with *Bam*HI and ligated to dephosphorylated pCK232 which was previously digested with *Bam*HI. Putative *E. coli* transformants containing recombinant plasmids were screened by *Bam*HI restriction digestion for the presence of the 2.385 kb insert. Three such recombinant plasmids containing the 2.385 kb insert in the pCK232 backbone were isolated, purified, and the cloned Ppta-AOR cassette additionally confirmed using PCR, additional restriction enzyme analysis and sequencing. This shuttle vector was designated pCK232AOR1 (Fig. 12, right) and was used to transform *C. autoethanogenum* DSM 11061.

[00107] **Example 16:** DNA transformation. Shuttle vectors pCK232 and pCK232AOR1 were transferred into *C. autoethanogenum* DSM 11061 by electroporation using essentially (with minor modifications) the method reported by Liu et al., (2006) for *C. tyrobutyricum* with scale-down adaptations implemented by Leang et al. (*Applied and Environmental Microbiology*, 79:1102-1109, 2013) for *C. ljungdahlii*. *C. autoethanogenum* DSM 11061 cells were grown in the fermentation medium described in Table 2, containing 5 g/L yeast extract without added DL-threonine and containing fructose as main carbon source (0.5%). Cells were harvested at mid-log phase (OD 600nm 0.3-0.4) and all subsequent steps including resuspensions and electrotransformation were carried out under strict anoxic conditions in an anaerobic chamber (Coy Products, MI). Washed cells were resuspended in SMP buffer pH 5.8 and aliquoted for individual electrotransformation (25 µl) into a 0.1 µm cuvette or stored frozen for near-term use at -80^{0}C in 10% DMSO and SMP buffer, pH 5.8.

The electroporation parameters used were 625 V, 100 Ω and 25 μF. Putative transformants were selected on fermentation medium agar (2%) containing 100 μg/ml kanamycin sulfate. Agar plates were incubated anaerobically in an anaerobic jar at 37°C under a N_2.CO_2 headspace.

[00108] **Example 17:** Recombinant AOR expression analysis by RT-qPCR. DNA transformation of pCK232AOR1 in *C. autoethanogenum* was confirmed by plasmid rescue into *E. coli* followed by restriction analysis and PCR amplification of a unique junction region of total DNA isolated from biocatalyst strain TF18 (data not shown). *In vivo* AOR1 expression from pCK232AOR1 in strain TF18 was shown in RT-qPCR tests using AOR1-specific primers and RNA isolated from cultures growing on syngas and fructose as sole carbon and energy sources. Wild-type and TF18 strains were grown in serum bottles (158 ml) in 20 ml of fermentation medium containing 5 g/l yeast extract and either syngas (65% H_2, 16% CO, and 15% CO_2) or 5 g/l fructose. One milliliter of cells was harvested at different optical densities corresponding to different growth phases, rapidly pelleted and flash frozen in an RNA protection solution. Frozen cells were stored at -80°C until needed. Total RNA was isolated using the RNA easy kit (Qiagen, Carlsbad, CA). Typically 3-5 μg of RNA was extracted/ml cell mass using this procedure. Complementary DNA was generated using reverse transcriptase as described by the manufacturer (ReverseScript, Ambion, Grand Island, NY). For the RT-qPCR, 50 ηg RNA was used for each reaction. A standard curve was set up using plasmid DNA that contained the cloned AOR1 gene target to quantitate transcript levels in the samples. Cycling conditions were: 95°C for 2 minutes (1 cycle, hot start activation); 40 cycles of 95°C (denaturation) for 15 sec. and 60°C (annealing and extension) for 60 sec. A final melting curve (dissociation) cycle of 55-95°C was performed to confirm specificity.

[00109] On fructose medium, both TF18 and wild type cells grew to a final OD of about 0.8 after 48-60 hours. Kanamycin at 10 ppm was used for plasmid maintenance in liquid growth for TF18. On syngas medium, the wild-type strain grew normally to a final OD of 0.7 but TF18 only reached an OD of 0.2 over two days and reached its maximum of 0.25 at 60 hours. Transcript levels and specific productivity rates were compared over the course of five days incubation (Figs. 13, 14). RT-qPCR analysis of the wild-type strain showed that the AOR1 gene is essentially not expressed when cells are grown on fructose but that recombinant strain TF18 expressed AOR1 abundantly (contained on pCK232AOR1) from the constitutive Ppta promoter when growing on fructose (Fig. 13). There was a greater than 30-fold difference between the wild-type and TF18 AOR1 transcript levels when cells were compared in mid-log phase (OD 0.4). In older, stationary phase TP18 cells (OD 0.8) AOR1 transcript declined significantly, although the gene expression levels were still about 4-fold higher than the wild-type strain. In both strains grown on syngas the Wood-Ljungdahl pathway is expected to be operative and the native AOR genes are expressed (Fig. 13), although minor differences between the wild-type and TF18 strains were likely due to differences in growth rates and cell yields under the conditions used.

[00110] **Example 18:** Productivity of *C. autoethanogenum* strain TF18 grown on syngas. Gas chromatography was performed on samples taken after 1, 2, and 5 days for the wild-type and TF18 (expressing recombinant Ppta-AOR) strains grown in syngas. The results showed that the TF18 strain had a more than 2-fold higher specific ethanol productivity rate (grams/OD) throughout the 5-day fermentation period than the wild-type strain (Fig. 14, left). The acetate productivity rate (grams/OD) was lower in the TF18 strain than the wild-type after 1 day of incubation, indicating a more rapid conversion of acid (Fig. 14, right).

[00111] **Example 19:** Scale up fermentation testing of recombinant strain TF18. The recombinant *C. autoethenogenum* strain TF18 was tested in a 5-L tank growing in

fermentation medium at pH 5.2 (Table 2) with kanamycin sulfate added at a final concentration of 25 μg/ml. Frozen stocks of TF18 were grown in 20 ml of fermentation medium in serum bottles containing 0.2% fructose to regenerate the strain. After reaching an OD of 0.8 (about 30 hours) cultures were transferred to 100 ml fermentation medium containing kanamycin sulfate (100 μg/ml) in a syngas headspace (65% H_2, 20% CO, and 8% CO_2). The 100 ml culture was transferred to 1900 ml fermentation medium in a 5-L reactor. When the culture optical density (600nm) reached 0.5 an additional 3000 ml medium was added for a final volume of 5 liters. Agitation was ramped up from 225 rpm to 525 rpm over the course of 7 days. Gas flow rate was 100 ml/min early in the fermentation and increased to 141 ml/min during steady state. Product selectivity of *C. autoethanogenum* strain TF18 is shown during a 400 hour period (Fig. 15). Only ethanol and acetate were produced, as expected for this culture. The selectivity in steady state ranged from 83.25% to 86.47% on a molar basis for ethanol and 16.75% to 13.53% for acetate. During the first 100 hours the ratio of ethanol to acetate (molar basis) was better than that seen for the parent strain which would contain only the two native AORs. The more rapid shift to solventogenesis is viewed as a strain and process improvement imparted by the additional AOR gene expression. The improvement in selectivity during this period was between 10-15% when compared to the average of many previous fermentations run solely with the parent strain. Improved selectivity would be one of the expected strain improvements of increased AOR expression, since the amount of total AOR protein would be expected to increase rapidly during growth phase and eventually settle into about $1x10^9$ copies/μg RNA in the parent strain (Fig. 10). Additional copies of AOR expressed off an autonomously replicating plasmid would likely increase the RNA levels further possibly into the $1x10^{10}$ copies/μg RNA. Increased transcript likely leads to increased levels of protein which would drive additional acetate to aldehyde conversion. This would in turn have the overall cellular effect of lessening the possibility of

acetate accumulation (increased free acid) and cell stress by diminution of the cell's membrane potential, which is primarily where *C. autoethanogenum* obtains its energy. Volumetric gas uptake was similar to that observed for the parental strain and depending on the gas flow ranged from 45 mM/l/h to over 52 mM/l/h. Once the culture was fully acclimated to the fermenter conditions, the growth rate of strain TF18 in the presence of 25 mg/L of kanamycin was easily maintained with a 2-day hydraulic retention time.

Seq ID No. 1 *C. autoethanogenum* AOR1 DNA sequence including native promoter.

TCTGAAATGGAACAAAATTGCACCAAGATGGAACAGTTTGCTTGGCGCAATACC
CTCAAAACACAATATTTTATTATTGGCATGGTTTATGCTATATATATAATTGTTTG
AATAATCAATTTTTTTTAGGAGGGTTTTTATGTACGGATATAAGGGTAAGGTATTA
AGAATTAATCTAAGTAGTAAAACTTATATAGTGGAAGAATTGAAAATTGACAAA
GCTAAAAAAATTTATAGGTGCAAGAGGGTTAGGCGTAAAAACCTTATTTGACGAA
GTAGATCCAAAGGTAGATCCATTATCACCTGATAACAAATTTATTATAGCAGCGG
GACCACTTACAGGTGCACCTGTTCCAACAAGCGGAAGATTCATGGTAGTTACTAA
ATCACCTTTAACAGGAACTATTGCTATTGCAAATTCAGGTGGAAAATGGGGAGC
AGAATTCAAAGCAGCTGGATACGATATGATAATCGTTGAAGGTAAATCTGATAA
AGAAGTTTATGTAAATATAGTAGATGATAAAGTAGAATTTAGGGATGCTTCTCAT
GTTTGGGGAAAACTAACAGAAGAAACTACAAAAATGCTTCAACAGGAAACAGAT
TCGAGAGCTAAGGTTTTATGCATAGGACCAGCTGGGGAAAAGTTATCACTTATGG
CAGCAGTTATGAATGATGTTGATAGAACAGCAGGACGTGGTGGTGTTGGAGCTG
TTATGGGTTCAAAGAACTTAAAAGCTATTGTAGTTAAAGGAAGCGGAAAAGTAA
AATTATTTGATGAACAAAAAGTGAAGGAAGTAGCACTTGAGAAAACAAATATTT
TAAGAAAAGATCCAGTAGCTGGTGGAGGACTTCCAACATACGGAACAGCTGTAC
TTGTTAATATTATAAATGAAAATGGTGTACATCCAGTAAAGAATTTTCAAAAATC
TTATACAGATCAAGCAGATAAGATCAGTGGAGAAACTTTAACTAAAGATTGCTT
AGTTAGAAAAAATCCTTGCTATAGGTGTCCAATTGCCTGTGGAAGATGGGTAAA
ACTTGATGATGGAACTGAATGTGGAGGACCAGAATATGAAACATTATGGTCATTT
GGATCTGATTGTGATGTATACGATATAAATGCTGTAAATACAGCAAATATGTTGT
GTAATGAATATGGACTAGATACCATTACAGCAGGATGTACTATTGCAGCAGCTAT
GGAACTTTATCAAAGAGGTTATATTAAGGATGAAGAAATAGCAGCAGATGGATT
GTCACTTAATTGGGGAGATGCTAAGTCCATGGTTGAATGGGTAAAGAAAATGGG
ACTTAGAGAAGGATTTGGAGACAAGATGGCAGATGGTTCATACAGACTTTGTGA
CTCATACGGTGTACCTGAGTATTCAATGACTGTAAAAAAACAGGAACTTCCAGCA
TATGACCCAAGAGGAATACAGGGACATGGCATTACTTATGCTGTTAACAATAGG
GGAGGATGTCACATTAAGGGATATATGGTAAGTCCTGAAATACTTGGCTATCCAG
AAAAACTTGATAGACTTGCAGTGGAAGGAAAAGCAGGATATGCTAGAGTATTCC
ATGATTTAACAGCTGTTATAGATTCACTTGGATTATGTATTTTTACAACATTTGGT
CTTGGTGCACAGGATTATGTTGATATGTATAATGCAGTAGTTGGTGGAGAATTAC
ATGATGTAAATTCTTTAATGTTAGCTGGAGATAGAATATGGACTTTAGAAAAAAT
ATTTAACTTAAAAGCAGGCATAGATAGTTCACAGGATACTCTTCCAAAGAGATTG
CTTGAAGAACAAATTCCAGAAGGACCATCAAAAGGAGAAGTTCATAAGTTAGAT

GTACTACTACCTGAATATTATTCAGTACGTGGATGGGATAAAAATGGTATTCCTA
CAGAGGAAACGTTAAAGAAATTAGGATTAGATGAATACGTAGGTAAGCTTTAG

Seq. ID No. 2 *C. autoethanogenum* amino acid sequence of AOR1.

M Y G Y K G K V L R I N L S S K T Y I V E E L K I D K A K K F I G
A R G L G V K T L F D E V D P K V D P L S P D N K F I I A A G P L
T G A P V P T S G R F M V V T K S P L T G T I A I A N S G G K W G
A E F K A A G Y D M I I V E G K S D K E V Y V N I V D D K V E F
R D A S H V W G K L T E E T T K M L Q Q E T D S R A K V L C I G
P A G E K L S L M A A V M N D V D R T A G R G G V G A V M G S
K N L K A I V V K G S G K V K L F D E Q K V K E V A L E K T N I
L R K D P V A G G G L P T Y G T A V L V N I I N E N G V H P V K
N F Q K S Y T D Q A D K I S G E T L T K D C L V R K N P C Y R C
P I A C G R W V K L D D G T E C G G P E Y E T L W S F G S D C D
V Y D I N A V N T A N M L C N E Y G L D T I T A G C T I A A A M
E L Y Q R G Y I K D E E I A A D G L S L N W G D A K S M V E W V
K K M G L R E G F G D K M A D G S Y R L C D S Y G V P E Y S M T
V K K Q E L P A Y D P R G I Q G H G I T Y A V N N R G G C H I K
G Y M V S P E I L G Y P E K L D R L A V E G K A G Y A R V F H D
L T A V I D S L G L C I F T T F G L G A Q D Y V D M Y N A V V G
G E L H D V N S L M L A G D R I W T L E K I F N L K A G I D S S Q
D T L P K R L L E E Q I P E G P S K G E V H K L D V L L P E Y Y S
V R G W D K N G I P T E E T L K K L G L D E Y V G K L

Seq. ID No. 3 DNA sequence of *C. autoethanogenum* AOR2 including native promoter.

AAATAGTATTGTACGTTATTTGATCACTTTTTTAAAAATAAAATGATAATATTGTA
TTTTTTTTGAGCATTTTGGCAATACAAAGAACTGTATTGTTTATTATTTAAGCATT
TTATTATAAAACAAAAAAACGTTATTAAATTATTTGCTATGAATTCACTTGATAA
TCAATGCATTGCATGTGATGTTGATTATTGAGTGTTTTTTTTGTAACCATATTTGG
CACAATTTATGCTCTATAATATTTCTGAAATAAATACATTTATATGAGGAGGAAT
TTCAATGTATGGTTATGATGGTAAAGTATTAAGAATTAATTTAAAAGAAAGAACT
TGCAAATCAGAAAATTTAGATTTAGATAAAGCTAAAAAGTTTATAGGTTGTAGG
GGACTAGGTGTTAAAACTTTATTTGATGAAATAGATCCTAAAATAGATGCATTAT
CACCAGAAAATAAATTTATAATTGTAACAGGTCCTTTAACTGGAGCTCCGGTTCC
AACTAGTGGAAGGTTTATGGTAGTTACTAAAGCACCGCTTACAGGAACTATAGG
AATTTCAAATTCGGGTGGAAAATGGGGAGTAGACTTAAAAAAAGCTGGTTGGGA
TATGATAATAGTAGAGGATAAGGCTGATTCACCAGTTTACATTGAAATAGTAGAT
GATAAGGTAGAAATTAAAGACGCGTCACAGCTTTGGGGAAAAGTTACATCAGAA
ACTACAAAAGAGTTAGAAAAGATAACTGAGAATAAATCAAAGGTATTATGTATA
GGACCTGCTGGTGAACGATTGTCTCTTATGGCAGCAGTTATGAATGATGTAGATA
GAACTGCAGCAAGAGGCGGCGTTGGTGCAGTTATGGGATCTAAAAACTTAAAAG
CTATTACAGTTAAAGGAACTGGAAAAATAGCTTTAGCTGATAAAGAAAAAGTAA
AAAAAGTGTCCGTAGAAAAAATTACAACATTAAAAAATGATCCAGTAGCTGGTC

AGGGAATGCCAACTTATGGTACAGCTATACTGGTTAATATAATAAATGAAAATG
GAGTTCATCCTGTAAAGAATTTTCAAGAGTCTTATACGAATCAAGCAGATAAAAT
AAGTGGAGAGACTCTTACTGCTAACCAACTAGTAAGGAAAAATCCTTGTTACAG
CTGTCCTATAGGTTGTGGAAGATGGGTTAGACTAAAAGATGGCACAGAGTGCGG
AGGACCAGAATATGAAACACTGTGGTGTTTTGGATCTGACTGTGGTTCATATGAT
TTAGATGCTATAAATGAAGCTAATATGTTATGTAATGAATATGGTATTGATACTA
TTACTTGTGGTGCAACAATTGCTGCAGCTATGGAACTTTATCAAAGAGGATATAT
AAAAGACGAAGAAATAGCTGGAGATAACCTATCTCTCAAGTGGGGTGATACGGA
ATCTATGATTGGCTGGATAAAGAGAATGGTATATAGTGAAGGCTTTGGAGCAAA
GATGACAAATGGTTCATATAGGCTTTGTGAAGGTTATGGAGCACCGGAGTATTCT
ATGACAGTTAAAAAGCAGGAAATTCCAGCATATGATCCAAGGGGAATACAGGGA
CACGGTATTACCTATGCAGTTAATAATAGAGGAGGCTGTCATATTAAGGGATACA
TGATTAACCCTGAAATATTAGGTTATCCTGAAAAACTTGATAGATTTGCATTAGA
TGGTAAAGCAGCTTATGCCAAATTATTTCATGATTTAACTGCTGTAATTGATTCTT
TAGGATTGTGCATATTCACTACATTTGGGCTTGGAATACAGGATTATGTAGATAT
GTATAATGCAGTAGTAGGAGAATCTACTTATGATGCAGATTCACTATTAGAGGCA
GGAGATAGAATCTGGACTCTTGAGAAATTATTTAATCTTGCAGCTGGAATAGACA
GCAGCCAGGATACTCTACCAAAGAGATTGTTAGAAGAACCTATTCCAGATGGCC
CATCAAAGGGAGAAGTTCATAGGCTAGATGTTCTTCTGCCAGAATATTACTCAGT
ACGAGGATGGAGTAAAGAGGGTATACCTACAGAAGAAACATTAAAGAAATTAG
GATTAGATGAATATATAGGTAAGTTCTAG

Seq ID No. 4 Amino acid sequence of *C. autoethanogenum* AOR2.

M Y G Y D G K V L R I N L K E R T C K S E N L D L D K A K K F I
G C R G L G V K T L F D E I D P K I D A L S P E N K F I I V T G P L
T G A P V P T S G R F M V V T K A P L T G T I G I S N S G G K W G
V D L K K A G W D M I I V E D K A D S P V Y I E I V D D K V E I K
D A S Q L W G K V T S E T T K E L E K I T E N K S K V L C I G P A
G E R L S L M A A V M N D V D R T A A R G G V G A V M G S K N
L K A I T V K G T G K I A L A D K E K V K K V S V E K I T T L K N
D P V A G Q G M P T Y G T A I L V N I I N E N G V H P V K N F Q E
S Y T N Q A D K I S G E T L T A N Q L V R K N P C Y S C P I G C G
R W V R L K D G T E C G G P E Y E T L W C F G S D C G S Y D L D
A I N E A N M L C N E Y G I D T I T C G A T I A A A M E L Y Q R G
Y I K D E E I A G D N L S L K W G D T E S M I G W I K R M V Y S
E G F G A K M T N G S Y R L C E G Y G A P E Y S M T V K K Q E I
P A Y D P R G I Q G H G I T Y A V N N R G G C H I K G Y M I N P E
I L G Y P E K L D R F A L D G K A A Y A K L F H D L T A V I D S L
G L C I F T T F G L G I Q D Y V D M Y N A V V G E S T Y D A D S
L L E A G D R I W T L E K L F N L A A G I D S S Q D T L P K R L L
E E P I P D G P S K G E V H R L D V L L P E Y Y S V R G W S K E G
I P T E E T L K K L G L D E Y I G K F

Seq ID No. 5 DNA sequence of AOR2 including Pptaack and terminator region

GAATTCTGATTGATTATTTATTTTAAAATGCCTAAGTGAAATATATACATATTATA
ACAATAAAATAAGTATTAGTGTAGGATTTTTAAATAGAGTATCTATTTTCAGATT

AAATTTTTGATTATTTGATTTACATTATATAATATTGAGTAAAGTATTGACTAGCA
AAATTTTTTGATACTTTAATTTGTGAAATTTCTTATCAAAAGTTATATTTTTGAAT
AATTTTTATTGAAAAATACAACTAAAAAGGATTATAGTATAAGTGTGTGTAATTT
TGTGTTAAATTTAAAGGGAGGAAAATGTATGGTTATGATGGTAAAGTATTAAGA
ATTAATTTAAAAGAAAGAACTTGCAAATCAGAAAATTTAGATTTAGATAAAGCT
AAAAAGTTTATAGGTTGTAGGGGACTAGGTGTTAAAACTTTATTTGATGAAATAG
ATCCTAAAATAGATGCATTATCACCAGAAAATAAATTTATAATTGTAACAGGTCC
TTTAACTGGAGCTCCGGTTCCAACTAGTGGAAGGTTTATGGTAGTTACTAAAGCA
CCGCTTACAGGAACTATAGGAATTTCAAATTCGGGTGGAAAATGGGGAGTAGAC
TTAAAAAAAGCTGGTTGGGATATGATAATAGTAGAGGATAAGGCTGATTCACCA
GTTTACATTGAAATAGTAGATGATAAGGTAGAAATTAAAGACGCGTCACAGCTTT
GGGGAAAAGTTACATCAGAAACTACAAAAGAGTTAGAAAAGATAACTGAGAAT
AAATCAAAGGTATTATGTATAGGACCTGCTGGTGAACGATTGTCTCTTATGGCAG
CAGTTATGAATGATGTAGATAGAACTGCAGCAAGAGGCGGCGTTGGTGCAGTTA
TGGGATCTAAAAACTTAAAAGCTATTACAGTTAAAGGAACTGGAAAAATAGCTT
TAGCTGATAAAGAAAAAGTAAAAAAAGTGTCCGTAGAAAAAATTACAACATTAA
AAAATGATCCAGTAGCTGGTCAGGGAATGCCAACTTATGGTACAGCTATACTGGT
TAATATAATAAATGAAAATGGAGTTCATCCTGTAAAGAATTTTCAAGAGTCTTAT
ACGAATCAAGCAGATAAAATAAGTGGAGAGACTCTTACTGCTAACCAACTAGTA
AGGAAAAATCCTTGTTACAGCTGTCCTATAGGTTGTGGAAGATGGGTTAGACTAA
AAGATGGCACAGAGTGCGGAGGACCAGAATATGAAACACTGTGGTGTTTTGGAT
CTGACTGTGGTTCATATGATTTAGATGCTATAAATGAAGCTAATATGTTATGTAA
TGAATATGGTATTGATACTATTACTTGTGGTGCAACAATTGCTGCAGCTATGGAA
CTTTATCAAAGAGGATATATAAAAGACGAAGAAATAGCTGGAGATAACCTATCT
CTCAAGTGGGGTGATACGGAATCTATGATTGGCTGGATAAAGAGAATGGTATAT
AGTGAAGGCTTTGGAGCAAAGATGACAAATGGTTCATATAGGCTTTGTGAAGGT
TATGGAGCACCGGAGTATTCTATGACAGTTAAAAAGCAGGAAATTCCAGCATAT
GATCCAAGGGGAATACAGGGACACGGTATTACCTATGCAGTTAATAATAGAGGA
GGCTGTCATATTAAGGGATACATGATTAACCCTGAAATATTAGGTTATCCTGAAA
AACTTGATAGATTTGCATTAGATGGTAAAGCAGCTTATGCCAAATTATTTCATGA
TTTAACTGCTGTAATTGATTCTTTAGGATTGTGCATATTCACTACATTTGGGCTTG
GAATACAGGATTATGTAGATATGTATAATGCAGTAGTAGGAGAATCTACTTATGA
TGCAGATTCACTATTAGAGGCAGGAGATAGAATCTGGACTCTTGAGAAATTATTT
AATCTTGCAGCTGGAATAGACAGCAGCCAGGATACTCTACCAAAGAGATTGTTA
GAAGAACCTATTCCAGATGGCCCATCAAAGGGAGAAGTTCATAGGCTAGATGTT
CTTCTGCCAGAATATTACTCAGTACGAGGATGGAGTAAAGAGGGTATACCTACA
GAAGAAACATTAAAGAAATTAGGATTAGATGAATATATAGGTAAGTTCTAGTTT
GATTCGGTAAACTAGAAAGCAGACTTTATGTGTTAAAGAAGATAGCTTCTCTCTA
TATATGAAGTCTGTTTTTAATAGAAAGATATGAATTTGAGAATAGAAGTTAGATT
AGTTGCTTATATATTTGCAAAAGTGTTAAGTTCTGCTTGGATAAGTTCGGGAGAT
GAAATTTAGTTGTTATGATAAACTTCAATGAATTC

Seq ID No. 6. DNA sequence of AOR2 including Pcoos with terminator region.

GATTCTTAGTATAAGTATTCTTAGTATCTTTAGCACTTAGAATACGTTATCCTTTA
GGAGAATAATCCTAATCAGTAGTTCTAATAATTTAATAGTATACTTAAATAGTAT
ATTTTGGAGGTTTTATTATGTATGGTTATGATGGTAAAGTATTAAGAATTAATTTA
AAAGAAAGAACTTGCAAATCAGAAAATTTAGATTTAGATAAAGCTAAAAAGTTT
ATAGGTTGTAGGGGACTAGGTGTTAAAACTTTATTTGATGAAATAGATCCTAAAA

TAGATGCATTATCACCAGAAAATAAATTTATAATTGTAACAGGTCCTTTAACTGG
AGCTCCGGTTCCAACTAGTGGAAGGTTTATGGTAGTTACTAAAGCACCGCTTACA
GGAACTATAGGAATTTCAAATTCGGGTGGAAAATGGGGAGTAGACTTAAAAAAA
GCTGGTTGGGATATGATAATAGTAGAGGATAAGGCTGATTCACCAGTTTACATTG
AAATAGTAGATGATAAGGTAGAAATTAAAGACGCGTCACAGCTTTGGGGAAAAG
TTACATCAGAAACTACAAAAGAGTTAGAAAAGATAACTGAGAATAAATCAAAGG
TATTATGTATAGGACCTGCTGGTGAACGATTGTCTCTTATGGCAGCAGTTATGAA
TGATGTAGATAGAACTGCAGCAAGAGGCGGCGTTGGTGCAGTTATGGGATCTAA
AAACTTAAAAGCTATTACAGTTAAAGGAACTGGAAAAATAGCTTTAGCTGATAA
AGAAAAAGTAAAAAAAAGTGTCCGTAGAAAAAATTACAACATTAAAAAATGATCC
AGTAGCTGGTCAGGGAATGCCAACTTATGGTACAGCTATACTGGTTAATATAATA
AATGAAAATGGAGTTCATCCTGTAAAGAATTTTCAAGAGTCTTATACGAATCAAG
CAGATAAAATAAGTGGAGAGACTCTTACTGCTAACCAACTAGTAAGGAAAAATC
CTTGTTACAGCTGTCCTATAGGTTGTGGAAGATGGGTTAGACTAAAAGATGGCAC
AGAGTGCGGAGGACCAGAATATGAAACACTGTGGTGTTTTGGATCTGACTGTGG
TTCATATGATTTAGATGCTATAAATGAAGCTAATATGTTATGTAATGAATATGGT
ATTGATACTATTACTTGTGGTGCAACAATTGCTGCAGCTATGGAACTTTATCAAA
GAGGATATATAAAAGACGAAGAAATAGCTGGAGATAACCTATCTCTCAAGTGGG
GTGATACGGAATCTATGATTGGCTGGATAAAGAGAATGGTATATAGTGAAGGCT
TTGGAGCAAAGATGACAAATGGTTCATATAGGCTTTGTGAAGGTTATGGAGCAC
CGGAGTATTCTATGACAGTTAAAAAGCAGGAAATTCCAGCATATGATCCAAGGG
GAATACAGGGACACGGTATTACCTATGCAGTTAATAATAGAGGAGGCTGTCATA
TTAAGGGATACATGATTAACCCTGAAATATTAGGTTATCCTGAAAAACTTGATAG
ATTTGCATTAGATGGTAAAGCAGCTTATGCCAAATTATTTCATGATTTAACTGCT
GTAATTGATTCTTTAGGATTGTGCATATTCACTACATTTGGGCTTGGAATACAGG
ATTATGTAGATATGTATAATGCAGTAGTAGGAGAATCTACTTATGATGCAGATTC
ACTATTAGAGGCAGGAGATAGAATCTGGACTCTTGAGAAATTATTTAATCTTGCA
GCTGGAATAGACAGCAGCCAGGATACTCTACCAAAGAGATTGTTAGAAGAACCT
ATTCCAGATGGCCCATCAAAGGGAGAAGTTCATAGGCTAGATGTTCTTCTGCCAG
AATATTACTCAGTACGAGGATGGAGTAAAGAGGGTATACCTACAGAAGAAACAT
TAAAGAAATTAGGATTAGATGAATATATAGGTAAGTTCTAGTTTGATTCGGTAAA
CTAGAAAGCAGACTTTATGTGTTAAAGAAGATAGCTTCTCTCTATATATGAAGTC
TGTTTTTAATAGAAAGATATGAATTTGAGAATAGAAGTTAGATTAGTTGCTTATA
TATTTGCAAAAGTGTTAAGTTCTGCTTGGATAAGTTCGGGAGATGAAATTTAGTT
GTTATGATAAACTTCAATGAATTC

Seq. ID No. 7 *C. pharus* Butyryl-CoA acetate transferase

ATGAGGAAGGTGTTTTTATTTAAAATATTAATAAATTTTTTGGAAGGGGTTTTAA
AAATGGTTTTTAAAAATTGGCAGGATCTTTATAAAAGTAAATTTGTTAGTGCAGA
TGAAGCTGTATCTAAGGTAAACTGTGGGGATACTATAGTTTTAGGTAATGCTTGT
GGAGCACCTCTTACACTTTTAGATGCTTTGGCTGCAAATAAGGAAAAGTATAAGA
GTGTACAGATATATAATCTTATACTGAACTATAAAAGTGATATATATGCTGAACC
AGGTGCAGAAAAGTATATTCATGGAAATACTTTTTTTGTAAGTGGAGGTACTAAG
GAAGCTGTTAATTGTAACAGAACAGATTATACCCCATGCTTTTTTTATGAAATAC
CAAAATTAATAAAACAAAAATATATACATATAGATGTAGCATTTATTCATGTAAG
TAAACCTGATAAGCATGGTTATTGTAGTTTTGGAGTATCAACTGATTATTCACAG
GCAATGGTACAGGGTGCTAAACTTGTAATTGCAGAAGTAAACGATCAAATGCCA
AGAGTTTTTGGAGACAATTTTATACACATTTCTGATATTGATTACATAGTAAAGA

CTTCACGCCCAATTCTTGAGTTGGCACCTCCTAAAATAGGAGAAGTAGAAAAAA
CAATAGGAAAATATTGTGCATCTCTTATAGAAGATGGTTCTACACTTCAACTTGG
AATAGGCGCTATTCCAGATGCAGTACTTTTGTTTTTGAAAGACAAAAAAGATTTG
GGAATACACTCAGAAATGATATCTGATGGTGTTGTAGAATTAGTTGAATCAGGA
GTAATTACAAATAAGAAAAAAGCTCTTCATCCAGGAAAAATAATTGTTACATTCT
TAATGGGAACTAAAAAATTATATGATTTTATAGATGATAATCCTATGGTAGAAGG
TTATCCAGTAGATTATGTAAATGATCCTAAAGTTATTATGCAAAATTCCAAGATG
GTATGTATAAATTCTTGTGTAGAGGTAGATTTTACAGGACAGGTGTGTGCTGAAA
GTATAGGATTTAAGCAGATAAGCGGAGTAGGCGGACAAGTTGATTATATGAGAG
GTGCTAGCATGTCTGATGGAGGAAAATCAATTCTTGCTATACCATCTACTGCAGC
TGGCGGCAAAATTTCAAGAATAGTTCCGATGTTGACTGAAGGAGCAGGAGTTAC
TACTTCAAGATATGATGTTCAATATGTTGTTACAGAGTATGGTATTGCACTTCTCA
AGGGTAAATCCATAAGAGAAAGAGCTAAAGCTCTTATAAACATTGCACATCCTA
AATTTAGAGAACAATTAGAAAAATCGTTTGAAGAAAGATTTAGTTGTAAACTTTA
A

Seq. ID No. 8 *C. kluyveri* Butyryl-CoA acetate transferase-1

ATGGTTTTTAAAAATTGGCAGGATCTTTATAAAAGTAAAATTGTTAGTGCAGACG
AAGCTGTATCTAAAGTAAGCTGTGGAGATAGCATAATTTTAGGCAATGCTTGTGG
AGCATCTCTTACACTTTTAGATGCCTTGGCTGCAAATAAGGAAAAGTATAAGAGT
GTAAAGATACACAATCTTATACTTAATTATAAAAATGATATATATACTGATCCGG
AATCAGAAAAGTATATTCATGGAAATACTTTCTTTGTAAGTGGAGGTACAAAGG
AAGCAGTTAATTGTAATAGAACAGATTATACTCCATGCTTTTTTTATGAAATACC
AAAATTATTAAAACAAAAGTATATAAATGCAGATGTAGCTTTTATTCAAGTAAGT
AAGCCTGATAGCCATGGATACTGTAGCTTTGGAGTATCAACCGATTATTCACAGG
CAATGGTACAGTCTGCAAAGCTTATAATTGCAGAAGTAAACGATCAGATGCCAA
GAGTTTTAGGAGACAATTTTATACACATTTCTGATATGGATTACATAGTAGAAAG
TTCACGTCCAATTCTAGAATTGACTCCTCCTAAAATAGGAGAAGTAGAGAAGAC
AATAGGAAAATACTGTGCATCTCTTGTAGAAGATGGTTCTACACTTCAGCTTGGA
ATAGGAGCTATTCCAGATGCAGTACTTTTATTCTTGAAGGATAAAAAGGATTTGG
GTATACATTCAGAAATGATATCCGATGGTGTTGTTGAATTAGTTGAAGCAGGGGT
AATTACAAATAAGAAAAAGTCCCTTCATCCAGGAAAAATAATTATTACATTCTTA
ATGGGAACTAAGAAATTATATGATTTCATAAATGATAATCCTATGGTAGAAGGAT
ACCCTGTAGATTATGTAAATGATCCTAAGGTTATTATGCAAAATTCTAAGATGGT
ATGTATAAACTCCTGTGTAGAAGTGGATTTCACAGGACAAGTGTGTGCTGAAAGT
GTAGGATTTAAACAAATAAGCGGTGTAGGTGGACAAGTTGATTACATGAGAGGA
GCTAGCATGGCTGATGGAGGAAAATCAATTCTTGCTATACCATCTACTGCAGCTG
GCGGCAAAATTTCAAGAATAGTTCCTATTTTAACTGAAGGAGCGGGGGTTACTAC
TTCAAGATATGATGTTCAATATGTTGTTACAGAATATGGTATTGCACTTCTCAAG
GGCAAATCCATAAGAGAAAGAGCTAAGGAGCTTATAAAAATTGCACATCCTAAA
TTTAGGGAAGAATTAACAGCTCAATTTGAAAAAAGATTCAGTTGTAAGCTTTAA

Seq. ID No. 9 *C. carboxidivorans* Butyrate kinase-1

ATGAGTTATAAGATATTAGCAATTAACCCAGGATCTACTTCTACAAAAATAGCTT
TATACGAAGATGAAAAAGAAATATTTTGCAAAACGTTAGAGCATCCAGTTGAAC
AAATTGAAAAATATGAGAATGTGGCAGATCAATTTGATATGAGAAAAGAAGTTG
TTCTTTCATTTTTAAAGCAAAATGGATATGAAGTTAAAGAATTAGCTGCAGTTGT

TGGAAGAGGTGGAATGGTTCCAAAAGTAAAATCTGGAGCTTATAAAGTTAATGA
AACAATGGTAGATAGATTAAAAAATAATCCAGTAGTAGAACATGCTTCAAATTT
AGGAGCTTTAATTGCTTATGAAATAGCAAATTCTATTGGAGTATCAGCCTATATA
TATGACTCTGTTAGAGTAGATGAATTAGAGGATATAGCTCGTATATCAGGTATGC
CGGATATACCAAGAACAAGTACTAGTCATGCATTAAATACAAGGGCAATGGCAA
TGAAGGTTGCAAAAAATTATGGTAAAAAGTATTCAGATATGAACTTTATTGTAGC
TCATCTAGGTGGAGGAATATCAGTAAATGTTCATAGAAAAGGACAAATGGTAGA
TATAATGGCAGATGACGAAGGACCATTTTCACCTGAAAGAGCTGGAAAAGTTCC
TTGCAATGCACTTATAGATCTTTGCTATTCAGGAAAATTTGATAAAAAAACTACG
AAGAAAAAATTAAGGGGAAATGGTGGATTAAAAGCTTATCTTAACACTGTTGAT
GCTAGAGAAGTTGAAAGAATGATTGAAAGTGGAGATGAAAAAGCAAAGCTTGTT
TATGAAGCTATGGCTTATCAGGTTGCTAAGGGAATAGGAGAACTTGCAACAGTA
GTAGAAGGTAAGGTTGATGCTATCGTTATTACAGGAGGTATAGCATATTCTGATA
TGATAACTAACTGGATTAAAAAGCGTGTAGAGTTTATTGCGCCTGTTGAGATTAT
GCCTGGTGAAAATGAAATGGAATCTTTGGCTTTGGGAACTCTTAGAGTGTTAAAG
GGTGAAGAAGAAGCAAGAGAATATGTTGAATAA

Sequence ID No. 10 *C. carboxidivorans* butyrate kinase-2

TTGCTAATTAAAATATTTATTAAGTATTGCTATAATCAGGAGGGTAAAATAATGT
ACAAAATACTAGCAATAAATCCAGGTTCAACTTCAACTAAAATAGCTATTTATGA
TGACACAGAGGAATTATTTAAAACCACTATAGAACATTCTAGTGAAGAAGTGAA
AAAATATGAAAACATAGCTGATCAATATAGTATGAGATATGAAGCTATAATGAA
ATTTTTAAAAGAAGTAGATTTTGATGTCAAAGCTTTATCTGCAGTAGTTGGAAGA
GGAGGAATTCTGCCTCCAGTTAAATCAGGAGCTTACAGAGTAAATGATTCTATGG
TAGAAAGACTGGCTAAAAGACCTGTAGTAGAGCATGCTTCAAATTTAGGAGCTA
TAATTTCATATGCAATAGCAAAACCTTTAAATATACCAGCTTTCATATATGATTCT
GTAGCTGTAGATGAATTTGAGGATATTGCAAGAATATCAGGACTTGCAGATATA
AAAAGAGAGAGTTTTATTCATGCTTTAAATATGAGAGCTGCAGCAATAAAAACA
GCAAAAAAACTAGGTAAACCTTATGAACAATGTAATTTAGTTGTTGCTCATTTAG
GAGGCGGAATATCTCTTACTGTACATAAAGGTGGAAAAATGATAGACGCTGTTA
CTGATGAAGAAGGACCGTTTTCACCAGAAAGGTCAGGTAGAGTACCTTGTAAGC
GCTTAATAGAAATGTGTTATAAAAATGATGAACGCACAATGAAAAAGAAAATAA
GAGGAGATGGTGGATTAATCTCTTATTTAGGAACTAATAGTGCATTAGATGTAGA
AAAAAGAATTGAAAATGGAGATGCTGAAGCCAAATTAGTTTATGAAGCTATGGC
ATATCAAATTGCAAAAGCAATAGGAGAACTTGCAACTGTAGTAAAGGGAAAGGT
TGATGCAGTAGTAATTACAGGGGGAATTGCCTATTCAAAAATGATGACAGGATG
GATAAAAGAAAGAGTAGAATTTATAGCACCTGTAGAGATATTGCCAGGAGAAAA
TGAATTAGAATCTCTTGCTTTAGGTACGCTTAGAGTTATAAAGGGAGAAGAAAA
AGCACACGAATATGATTTAGATTAG

Sequence ID No. 11 *C. carboxidivorans* Butyrate Kinase-3.

ATGTCATATAAATTATTAATATTAAATCCAGGATCTACATCTACCAAAATAGGAG
TATATGATGGAGAAAATGAAATTTTAGAAGAAACTTTAAGACATTCTTCAGAAG
AAATTGAGAAATATGCTACTATTTATGATCAATTTGAATTTAGAAAAGAAGTTAT
ATTGAAGGTTTTAAAAGAAAAGAATTTTGATATTAATACATTAGACGGAGTAGTA
GGCAGAGGTGGATTATTAAAACCAATTGAAAGTGGAACTTATAAAGTCAATGAT
GCTATGTTAGAAGACCTAAAAGTTGGAGTGCAAGGACAGCATGCTTCAAATTTA

GGTGGAATAATAGCTAATGAAATAGGAAAATCTATAAATAAACCAGCATTTATA
GTAGACCCAGTTGTTGTTGATGAATTAGATGAAGCAGCTAGAATATCCGGAATGC
CTGAAATAGAAAGAATAAGTATATTCCATGCTTTAAATCAAAAAGCAGTAGCAA
AGAGATATGCAAAAGAAAACAATAAGAAGTATGATGAATTAAATTTAGTAGTGA
CACACATGGGTGGCGGAGTAACTGTTGGAGCTCACAAAAAAGGAAGAGTTGTAG
ATGTAGCCAATGGTTTAGATGGAGATGGACCATTTTCACCAGAAAGAACAGGAG
GACTTCCTGTAGGAGGTTTAATAAAGCTTTGCTATAGTGGAAAATATACTTTAGA
AGAAATGAAGAAAAAGATAAGTGGAAAAGGTGGAATTGTAGCTTATCTAAATAC
AAATGATTTTAGGGAAGTAGAACAAAAAGCAGAAAGTGGAGATAAAAAGGCAA
AGTTAGTATTTGATGCTTTCATATTACAAGTAGGTAAAGAAATTGGTAAATGTGC
TGCAGTTTTACATGGAAAAGTAGATGCTTTAATTTTAACTGGAGGAATAGCTTAT
AGTAAAACTGTTACAGCTGCAATAAAAGACATGGTAGAATTTATTGCACCAGTTG
TAGTTTATCCAGGAGAAGATGAATTATTAGCATTAGCACAAGGCGGACTTAGAG
TACTAGGTGGAGAAGAACAAGCAAAAGAATATAAGTAA

Sequence ID No. 12 *C. difficile* butyrate kinase

ATGACTTACAGAATATTAGCCATAAATCCAGGTTCTACTTCTACAAAAATAGCAG
TATATGATGGAGAAGAACAAATTCTTGTGAAGACGATAGACCATCCGGCTGAAG
AGATTGCAAAATATAATACTATACAAGACCAGTTTGAAATGCGTAAGGAAGCAG
TTTTGAATATTCTTAAAGAAAATAGTATAGACTTAAAATCTCTTAGTGCAATAGT
AGGAAGAGGTGGAGTTTTACCACCAGTAAAATCAGGAGCATATTTAGTAAATGA
AGAAATGATTGATGTACTAAGACATAGACCAGTACTTGAACACGCTTCCAATTTA
GGTGCTGTTGTGGCACATGCAATATCAGAACCTCTTGGAATCAACTCATATATTT
ATGATTCTGTTGCAGTAGATGAGCTTATAGATGTAGCGAGAATATCTGGACTTTG
TGGAATGGATAGATCAAGTGCAGGGCATGCATTAAATACTAGAGCAATGGCTTT
AAAATATGCTAAGGATAAAGGAAAAGATTATAAGAGCTTAAACTTAATAGTAGC
TCACATTGGTGGAGGAGTAAGTATTTATCTTCATGAAAAAGGAAGAATGGTTGAT
ATGCTATCTGATGATGAAGGACCATTTTCTCCAGAAAGGTCAGGAAGAGTACCTG
CTACAAAATTAGTGGCTGCCTGTTATTCAGGTCAATATTCAGAAAGAGAAATGAC
TAAAAAGATAAGAGGTAAAGGTGGTATAGTTTCATACCTAAATACTGTAGATGC
TAGAGAAGTTGAAAAAATGATAGCAGAAGGAAATGAAGAAGCAAAAATTATTT
ATGAAGCAATGGCTTATCAGTTAGCAAAAGGTATTGGAGAGTTAGCAACTGTAG
TAGATGGAAAGGTAGATGCTATAATTATAACAGGTGGAATTGCATATTCTGAAAT
GTTTACTTCAATGGTTAAAAAGAAAGTTGAGTTTATAGCACCAGTAGAAATTATG
GCAGGAGAAAATGAGTTGGATAATCACTTGCTTTTGGAACTTTAAGAGTACTAAA
TGGAGAAGAAGAAGCTAGAATTTATAGTGAAA

Seq ID No. 13 171 bp region in tungsten-dependent aldehyde ferredoxin oxidoreductase -1 (AOR1) covered by the probes found in Seq. ID No.1

5'-
ACTTGGATTATGTATTTTTACAACATTTGGTCTTGGTGCACAGGATTATGTTGATA
TGTATAATGCAGTAGTTGGTGGAGAATTA
CATGATGTAAATTCTTTAATGTTAGCTGGAGATAGAATATGGACTTTAGAAAAAA
TATTTAACTTAAAAGCAGGCATAGATAGTTCA-3'

Seq ID No. 14 169 bp region in tungsten-dependent aldehyde ferredoxin oxidoreductase #2 (AOR2) covered by the probes found in Seq. ID NO: 3.

5'-
AAGAAAGAACTTGCAAATCAGAAAATTTAGATTTAGATAAAGCTAAAAAGTTTA
TAGGTTGTAGGGGACTAGGT
GTTAAAACTTTATTTGATGAAATAGATCCTAAAATAGATGCATTATCACCAGAAA
ATAAATTTATAATTGTAACAGGTCCTTTAACTGGAGCTCCG-3'

Seq ID No. 15 *C. ljungdahlii* AOR1

MYGYKGKVLRINLSSKTYIVEELKIDKAKKFIGARGLGVKTLFDEVDPKVDPLSPDN
KFIIAAGPLTGAPVPTSGRFMVVTKSPLTGTIAIANSGGKWGAEFKAAGYDMIIVEGK
SDKEVYVNIVDDKVEFRDASHVWGKLTEETTKMLQQETDSRAKVLCIGPAGEKLSL
MAAVMNDVDRTAGRGGVGAVMGSKNLKAIVVKGSGKVKLFDEQKVKEVALEKTN
ILRKDPVAGGGLPTYGTAVLVNIINENGVHPVKNFQKSYTDQADKISGETLTKDCLV
RKNPCYRCPIACGRWVKLDDGTECGGPEYETLWSFGSDCDVYDINAVNTANMLCN
EYGLDTITAGCTIAAAMELYQRGYIKDEEIAADGLSLNWGDAKSMVEWVKKMGLR
EGFGDKMADGSYRLCDSYGVPEYSMTVKKQELPAYDPRGIQGHGITYAVNNRGGC
HIKGYMVSPEILGYPEKLDRLAVEGKAGYARVFHDLTAVIDSLGLCIFTTFGLGAQD
YVDMYNAVVGGELHDVNSLMLAGDRIWTLEKIFNLKAGIDSSQDTLPKRLLEEQIPE
GPSKGEVHKLDVLLPEYYSVRGWDKNGIPTEETLKKLGLDEYVGKL

Seq ID No. 16 *C. ragsdalei* AOR1

MYGYSGKVLRINLSNKTYKAEELKIDEAKKFIGARGLGVKTLLDEIDPKIDPLSPDNK
FIIATGPLTGAPVPTSGRFMVITKAPLTGTIGIANSGGKWGAELKTAGYDMVIVEGKS
DKPVYVNIVDDKVEFKDASHVWGKLTEETTKMLQNEIDAKAKVLCIGPAGENLSLM
AAVMNDIDRTAGRGGVGAVMGSKNLKAIVVKGSGKVKLFDEEKVKAVSLQKSDIL
RKDPVAGGGLPTYGTAVLVNIINENGINPVRNFQESYTDEADKVSGETMTQECLVRK
NPCYRCPIACGRWVRLDDGTECGGPEYETLWSFGSDCDVYDLNAVNKANMLCNEY
GLDTISAGATIASAMELYQRGYIKDEEIAADGLSLKWGDAKSMVEWVKKMGRREGF
GGKMADGSYRLCESYGVPQYSMSVKKQELPAYDPRGAQGHGLTYAVNNRGGCHIK
GYMISPEILGYPEKLDRFSIEGKPAYAKVFHDLTAVIDSLGLCIFTTFGLGAQDYVDM
YNAVVGGELHDVDSLMLAGDRVWTLEKIFNLKAGVGSSQDTLPKRLLEEEVVEGPS
KGHVHRLDELVPEYYSVRGWDKNGVPTEETLKKLGLEEYIGKI

Seq ID No. 17 *C. ljungdahlii* AOR2

MYGYDGKVLRINLKERTCKSENLDLDKAKKFIGCRGLGVKTLFDEIDPKIDALSPEN
KFIIVTGPLTGAPVPTSGRFMVVTKAPLTGTIGISNSGGKWGVDLKKAGWDMIIVED
KADSPVYIEIVDDKVEIKDASQLWGKVTSETTKELEKITENKSKVLCIGPAGERLSLM
AAVMNDVDRTAARGGVGAVMGSKNLKAITVKGTGKIALADKEKVKKVSVEKITTL
KNDPVAGQGMPTYGTAILVNIINENGVHPVKNFQESYTNQADKISGETLTANQLVRK
NPCYSCPIGCGRWVRLKDGTECGGPEYETLWCFGSDCGSYDLDAINEANMLCNEYG
IDTITCGATIAAAMELYQRGYIKDEEIAGDNLSLKWGDTESMIGWIKRMVYSEGFGA
KMTNGSYRLCEGYGAPEYSMTVKKQEIPAYDPRGIQGHGITYAVNNRGGCHIKGYM
INPEILGYPEKLDRFALDGKAAYAKLFHDLTAVIDSLGLCIFTTFGLGIQDYVDMYNA
VVGESTYDADSLLEAGDRIWTLEKLFNLAAGIDSSQDTLPKRLLEEPIPDGPSKGEVH
RLDVLLPEYYSVRGWSKEGIPTEETLKKLGLDEYIGKF

Seq ID No. 18 *C. ragsdalei* AOR2

MYGYNGKVLRINLKERTCKSENLDLDKAKKFIGCRGLGVKTLFDEIDPKIDALSPEN
KFIIVTGPLTGAPVPTSGRFMVVTKAPLTGTIGISNSGGKWGVDLKKAGWDMIIVED

KADSPVYIEIVDDKVEIKDASQLWGKVTSETTKELEKITENRSKVLCIGPAGERLSLM
AAVMNDVDRTAARGGVGAVMGSKNLKAITVKGTGKIALADKEKVKKVSVEKITTL
KNDPVAGQGMPTYGTAILVNIINENGVHPVNNFQESYTDQADKISGETLTANQLVRK
NPCYSCPIGCGRWVRLKDGTECGGPEYETLWCFGSDCGSYDLDAINEANMLCNEYG
IDTITCGATIAAAMELYQRGYVKDEEIAGDNLSLKWGDTESMIGWIKKMVYSEGFG
AKMTNGSYRLCEGYGVPEYSMTVKKQEIPAYDPRGIQGHGITYAVNNRGGCHIKGY
MINPEILGYPEKLDRFALDGKAAYAKMMHDLTAVIDSLGLCIFTTFGLGIQDYVDMY
NAVVGESTCDSDSLLEAGDRVWTLEKLFNLAAGIDSSQDTLPKRLLEEPIPDGPSKGH
VHRLDVLLPEYYSVRGWSKEGIPTEETLKKLGLDEYIGKF

Seq ID No. 19 AOR1-A-forward primer

5'-ACTTGGATTATGTATTTTTACA-3'

Seq ID No. 20 AOR1-A-reverse primer

5'- TGAACTATCTATGCCTGCTTTT-3'

Seq ID No. 21 AOR2-A-forward primer

5'-AAGAAAGAACTTGCAAATCA-3'

Seq ID No. 22 AOR2-A reverse primer

5'-CGGAGCTCCAGTTAAAGGA-3'

Seq ID No. 23 BCoAAT forward primer

5'-AGCCATGCTAGC TCCTCTCATGTA-3'

Seq ID No. 24 BCoAAT reverse primer

5'-GGAGTATCAACCGAT TATTCACAG-3'

Seq ID No. 25 Buk forward primer

5'-GATGGTTCTACACTT CAGCTT-3'

Seq ID No. 26 Buk reverse primer

5'-GAT ATC ATT TCT GAA TGT ATA CCC-3'

Seq ID No. 27 iron-only hydrogenase-forward primer

5'-TGT GAA CGT CCT GAA ATG AAA G-3'

Seq ID No. 28 iron-only hydrogenase reverse primer

5'-AGT GCC TGC ACC AGA ATA AGT T-3'

Seq ID No. 29 Hyg-1 forward primer

5'-GCCCGATATAAATCCTCTTT-3'

Seq ID No. 30 Hyg-1 reverse primer

5'-CCAACAAAAATTCCATGATT-3'

Seq ID No. 31 Hyg-2 forward primer

5'-CTACAATTTTAAACGCTGCA-3'

Seq ID No. 32 Hyg-2 reverse primer

5'-GCTCTGGCACTGTTTGTTCTA-3'

Seq ID No. 33 Hyg-3 forward primer

5'-TGA TAC AAA CTT TGG TGC AG-3'

Seq ID No. 34 Hyg-3 reverse primer

5'-ATA TAG CTC CAG CCA TCT GA-3'

CLAIMS

1. A method for the production of an alcohol, comprising: culturing a recombinant C1-fixing homoacetogen microorganism in a culture medium in the presence of syngas, wherein the C1-fixing homoacetogen microorganism comprises a recombinant gene encoding an aldehyde ferredoxin oxidoreductase polypeptide modulated by a promoter, wherein the recombinant C1-fixing microorganism produces alcohol.

2. The method of claim 1 wherein the recombinant aldehyde ferredoxin oxidoreductase gene encodes a *Clostridium* aldehyde ferredoxin oxidoreductase polypeptide.

3. The method of claim 2 wherein the aldehyde ferredoxin oxidoreductase gene is at least 95 % identical to SEQ ID NO: 1 or SEQ ID NO: 3.

4. The method of claim 1 wherein the promoter is an inducible promoter or a constitutive promoter.

5. The method of claim 1 wherein the alcohol is ethanol.

6. The method of claim 1 wherein the C1-fixing homoacetogen microorganism *is C. ljungdahlii, C. ragsdalei, C. autoethanogenum* or *C. coskatii.*

7. A method for the production of butanol, comprising: culturing a co-culture comprising a recombinant C1-fixing homoacetogen microorganism and a C4-producing butyrate microorganism in a culture medium in the presence of syngas, wherein the C1-fixing homoacetogen microorganism comprises a recombinant gene encoding an aldehyde ferredoxin oxidoreductase polypeptide modulated by a promoter.

8. The method of claim 7 wherein the C4-producing butyrate microorganism comprises a gene encoding a Butyryl-CoA acetate transferase polypeptide or a Butyrate kinase polypeptide.

9. The method of claim 7 wherein the recombinant aldehyde ferredoxin oxidoreductase gene encodes a *Clostridium* aldehyde ferredoxin oxidoreductase polypeptide.

10. The method of claim 9 wherein the aldehyde ferredoxin oxidoreductase gene is at least 95% identical to SEQ ID NO: 1 or SEQ ID NO: 3.

11. The method of claim 7 wherein the promoter is an inducible promoter or a constitutive promoter.

12. The method of claim 7 wherein the C1-fixing homoacetogen microorganism *is C. ljungdahlii, C. ragsdalei, C. autoethanogenum* or *C. coskatii.*

13. The method of claim 7 wherein the C4-producing butyrate microorganism is *Clostridium kluyveri*, *Clostridium carboxidivorans*, *Butyribacterium methylotrophicum*, or *Clostridium pharus*.

14. A microorganism co-culture for the conversion of syngas to butanol comprising: a recombinant C1-fixing homoacetogen microorganism comprising a recombinant gene encoding an aldehyde ferredoxin oxidoreductase polypeptide modulated by a promoter and a C4-producing butyrate microorganism.

15. The co-culture of claim 14 wherein the C4-producing butyrate microorganism comprises a gene encoding a Butyryl-CoA acetate transferase polypeptide or a Butyrate kinase polypeptide.

16. The co-culture of claim 14 wherein the recombinant aldehyde ferredoxin oxidoreductase gene encodes a *Clostridium* aldehyde ferredoxin oxidoreductase polypeptide.

17. The co-culture of claim 16 wherein the aldehyde ferredoxin oxidoreductase gene is at least 95 % identical to SEQ ID NO: 1 or SEQ ID NO: 3.

18. The co-culture of claim 14 wherein the promoter is an inducible promoter or a constitutive promoter.

19. The co-culture of claim 14 wherein the C1-fixing homoacetogen microorganism *is C. ljungdahlii, C. ragsdalei, C. autoethanogenum* or *C. coskatii.*

20. The co-culture of claim 14 wherein the C4-producing butyrate microorganism is *Clostridium kluyveri*, *Clostridium carboxidivorans*, *Butyribacterium methylotrophicum*, or *Clostridium pharus*.

Fig. 1

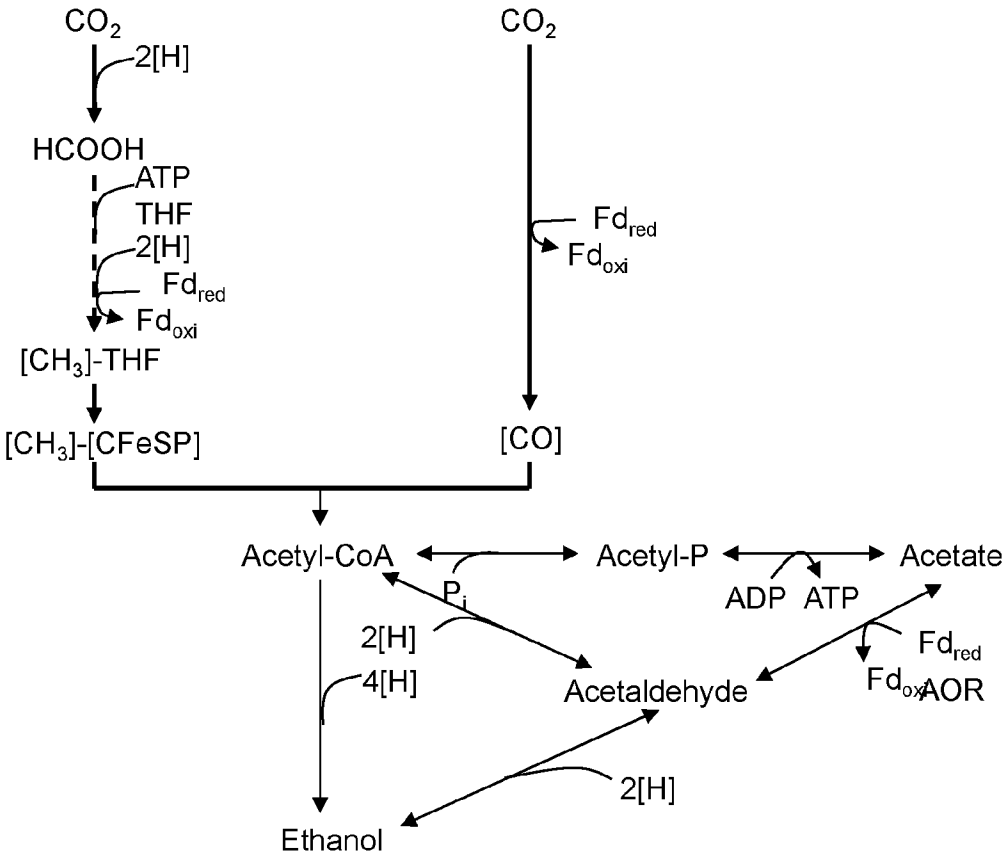

CO_2
2[H]
HCOOH
ATP
THF
2[H]
Fd_{red}
Fd_{oxi}
$[CH_3]$-THF
$[CH_3]$-[CFeSP]
CO_2
Fd_{red}
Fd_{oxi}
[CO]
Acetyl-CoA
Acetyl-P
Acetate
P_i
ADP ATP
2[H]
4[H]
Acetaldehyde
Fd_{red}
Fd_{oxi}
AOR
2[H]
Ethanol

Fig. 2

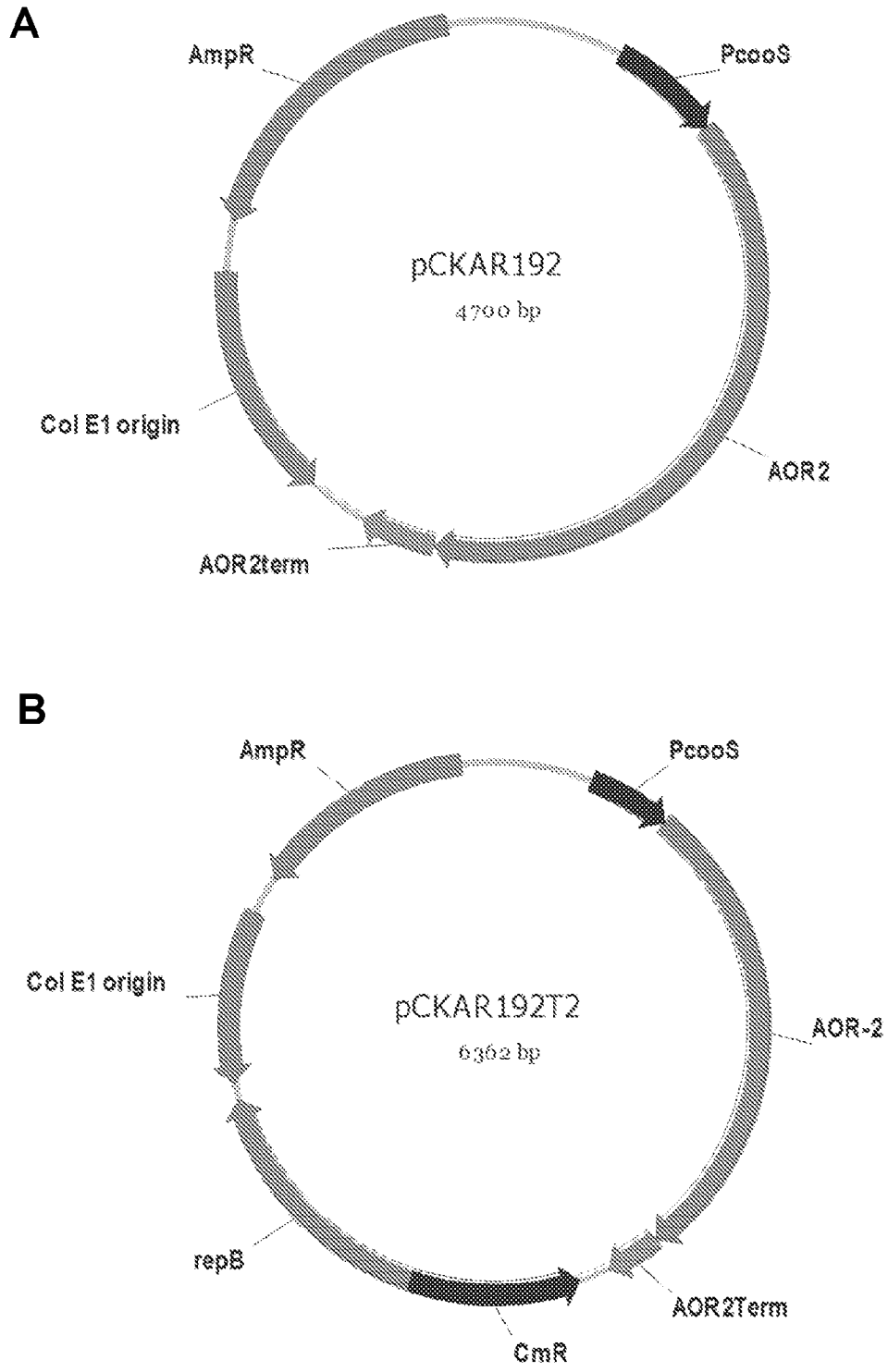

A
AmpR
PcooS
pCKAR192
4790 bp
Col E1 origin
AOR2
AOR2term
B
AmpR
PcooS
Col E1 origin
pCKAR192T2
6362 bp
AOR-2
repB
AOR2Term
CmR

Fig. 3

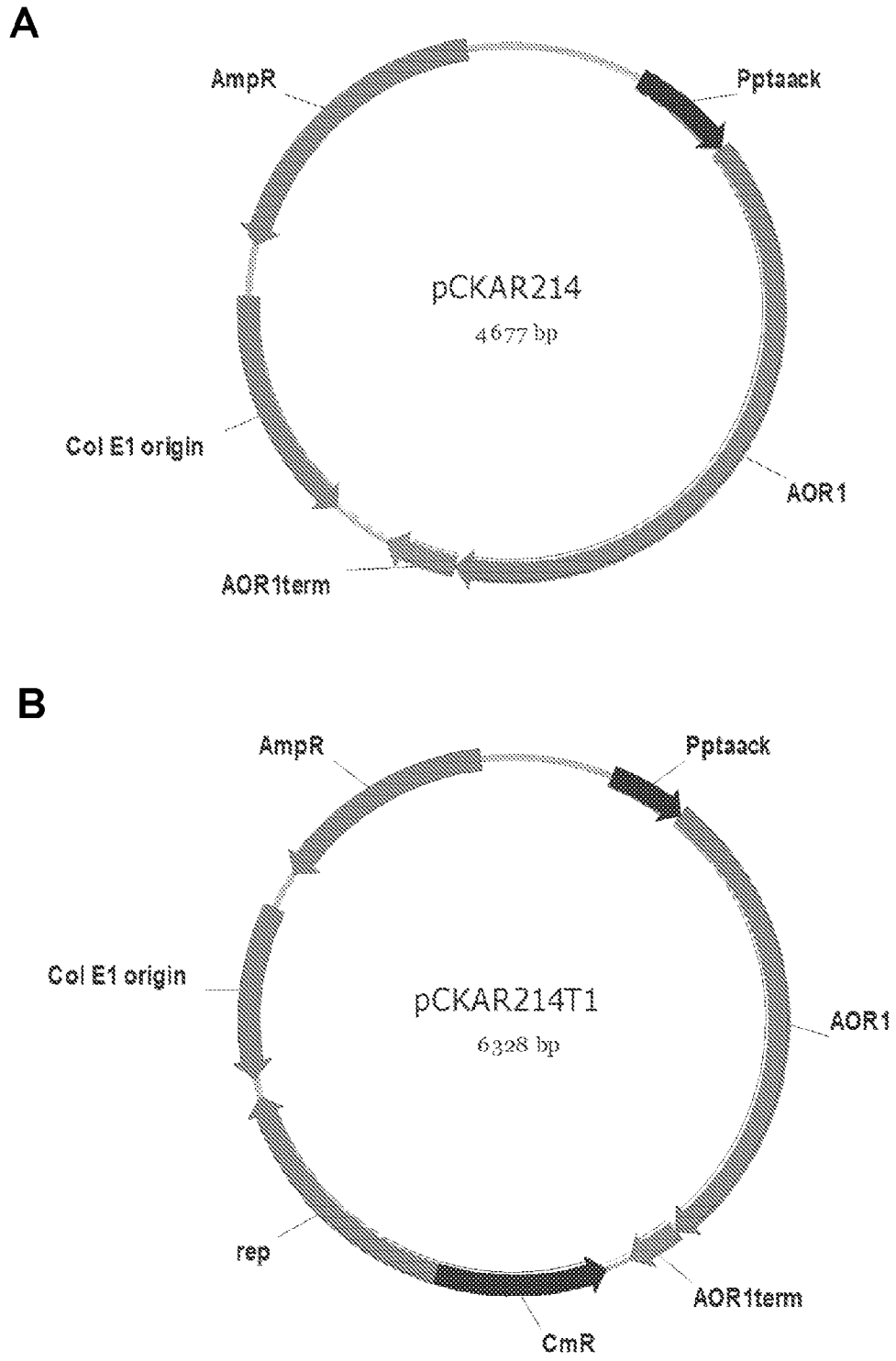

A
AmpR
Pptaack
pCKAR214
4677 bp
Col E1 origin
AOR1
AOR1term
B
AmpR
Pptaack
Col E1 origin
pCKAR214T1
6328 bp
AOR1
rep
AOR1term
CmR

Fig. 4A. DNA sequence alignment of *C. autoethanogenum* AOR1 and AOR2.

```
AOR1    ATGTACGGATATAAGGGTAAGGTATTAAGAATTAATCTAAGTAGTAAAACTTATATAGTGGAAGAATTGAAAATTGACAA
AOR2    ATGTATGGTTATGATGGTAAAGTATTAAGAATTAATTTAAAAGAAAGAACTTGCAAATCAGAAAATTTAGATTTAGATA
        ***** ** *** * ****** ********************** ***          * ******   * *    *** * **  *  * ** **

AOR1    AGCTAAAAAAATTTATAGGTGCAAGAGGGTTAGGCGTAAAAACCTTATTTGACGAAGTAGATCCAAAGGTAGATCCATTAT
AOR2    AGCTAAAAAGTTTATAGGTTGTAGGGGACTAGGTGTTAAAACTTTATTTGATGAAATAGATCCTAAAATAGATGCATTAT
        ********** **********    ** **   ***** ** ****** ********** *** ********* **  ***** ******

AOR1    CACCTGATAACAAATTTATTATAGCAGCGGGACCACTTACAGGTGCACCTGTTCCAACAAGCGGAAGATTCATGGTAGTT
AOR2    CACCAGAAAATAAATTTATAATTGTAACAGGGTCCTTTAACTGGAGCTCCGGTTCCAACTAGTGGAAGGTTTATGGTAGTT
        **** ** ** ********** ** * * * ** **  * ** ** ** ********** ** ***** ** **********

AOR1    ACTAAATCACCTTTAACAGGAACTATTGCTATTGCAAATTCAGGTGGAAAATGGGGAGCAGAATTCAAAGCAGCTGGATA
AOR2    ACTAAAGCACCGCTTACAGGAACTATAGGAATTTCAAATTCGGGTGGAAAATGGGGAGTAGACTTAAAAAAAGCTGGTTG
        ******  *****   * ************** *  ***  ********* ********************** *** ** ***  ******* *

AOR1    CGATATGATAATCGTTGAAGGTAAATCTGATAAAGAAGTTTATGTAAATATAGTAGATGATAAAGTAGAATTTAGGGATG
AOR2    GGATATGATAATAGTAGAGGATAAGGCTGATTCACCAGTTTACATTGAAATAGTAGATGATAAGGTAGAAATTAAAGACG
        ********** ** ** * ***   ******   *   *******  * ********************* ******* ***  ** *

AOR1    CTTCTCATGTTTGGGGAAACTAACAGAAGAAACTACAAAAATGCTTCAACAGGAAACAGATTCGAGAGCTAAGGTTTTA
AOR2    CGTCACAGCTTTGGGGAAAAGTTACATCAGAAACTACAAAAGAGTTAGAAAAGATAACTGAGAATAAATCAAAGGTATTA
        * ** **  ********** * ***  ********** * *  ** **  *** **     * * * ****** ***

AOR1    TGCATAGGACCAGCTGGGGAAAAGTTATCACTTATGGCAGCAGTTATGAATGATGTTGATAGAACAGCAGGACGTGGTGG
AOR2    TGTATAGGACCTGCTGGTGAACGATTGTCTCTTATGGCAGCAGTTATGAATGATGTAGATAGAACTGCAGCAAGAGGCGG
        ** ********** ***** ***    ** ** ********** ********** ********* **** * * ** **

AOR1    TGTTGGAGCTGTTATGGGTTCAAAGAACTTAAAAGCTATTGTAGTTAAAGGAAGCGGAAAGTAAAATTATTTGATGAAC
AOR2    CGTTGGTGCAGTTATGGGATCTATAAAAACTTTAAAAGCTATTACAGTTAAAGGAACTGGAAAAATAGCTTTAGCTGATAAAG
          ****** ** ********* ** ** ********************  ********** *******  **  ***  **** **
```

```
AOR1  AAAAAGTGAAGGAAGTAGCACTTGAGAAAACAAATATTTTAAGAAAAGATCCAGTAGCTGGTGAGGACTTCCAACATAC
AOR2  AAAAAGTAAAAAAAAAGTGTCCGTAGAAAAAAATTACAACATTAAAAAATGATCCAGTAGCTGGTCAGGGAATGCCAACTTAT
      ********  **  ****  *   *  **  ****  *   *  ****  ***  ********************    ***  *  ******  **

AOR1  GGAACAGCTGTACTTGTTAATATTATAAATGAAAATGGTGTACATCCAGTAAAGAATTTTCAAAAATCTTATACAGATCA
AOR2  GGTACAGCTATACTGGTTAATATAATAAATGAAAATGGAGTTCATCCTGTAAAGAATTTTCAAGAGTCTTATACGAATCA
      ** *******  ****  ********  ********** ** *****  ********************  *  ********   ****

AOR1  AGCAGATAAGATCAGTGGAGAAACTTTAACTAAAGATTGCTTAGTTAGAAAAAATCCTTGCTATAGGTGTCCAATTGCCT
AOR2  AGCAGATAAAATAAGTGGAGAGACTCTTACTGCTAACCAACTAGTAAGGAAAAATCCTTGTTACAGCTGTCCTATAGGTT
      **********  **  **********  ***  * ***     *       ****  **  ********************  **  **  *****  **  *  *

AOR1  GTGGAAGATGGGTAAAACTTGATGATGGAACTGAATGTGGAGGACCAGAATATGAAACATTATGGTCATTTGGATCTGAT
AOR2  GTGGAAGATGGGTTAGACTAAAAGATGGCACAGAGTGCGGAGGACCAGAATATGAAACACTGTGGTGTTTTGGATCTGAC
      **********  *  ***   *  ******  **  **  ********************  *  ****   **********

AOR1  TGTGATGTATACGATATAAATGCTGTAAATACAGCAAATATGTTGTGTAATGAATATGGACTAGATACCATTACAGCAGG
AOR2  TGTGGTTCATATGATTTAGATGCTATAAATGAAGCTAATATGTTATGTAATGAATATGGTATTGATACTATTACTTGTGG
      ****  *   ***  ***  **  ******  ******  ***  *********  **********  *  *****  *****     **

AOR1  ATGTACTATTGCAGCAGCTATGGAACTTTATCAAAGAGGTTATATTAAGGATGAAGAATAGCAGCAGATGGATTGTCAC
AOR2  TGCAACAATTGCTGCAGCTATGGAACTTTATCAAAGAGGATATATAAAAGACGAAGAAATAGCTGGAGATAACCTATCTC
        **  ******  **********************  ******  **  **  **********  *  ****      *  **  *

AOR1  TTAATTGGGGAGATGCTAAGTCCATGGTTGAATGGGTAAAGAAAATGGGACTTAGAGAAGGATTTGGAGACAAGATGGCA
AOR2  TCAAGTGGGGTGATACGGAATCTATGATTGGCTGGATAAAGAGAATGGTATATAGTGAAGGCTTTGGAGCAAAGATGACA
      * **  *****  ***  *  * **  ***  ***  ******  *****  *  ***  *****  *******  *******  **

AOR1  GATGGTTCATACAGACTTTGTGACTCATACGGTGTACCTGAGTATTCAATGACTGTAAAAAAACAGGAACTTCCAGCATA
AOR2  AATGGTTCATATAGGCTTTGTGAAGGTTATGGAGCACCGGAGTATTCTATGACAGTTAAAAAGCAGGAAATTCCAGCATA
        **********  **  **********      **  **  *  ***  *********  *****  **  *****  *******  **********

AOR1  TGACCCAAGAGGAATACAGGGACATGGCATTACTTATGCTGTTAACAATAGGGGAGGATGTCACATTAAGGGATATATGG
AOR2  TGATCCAAGGGGAATACAGGGACACGGTATTACCTATGCAGTTAATAATAGAGGAGGCTGTCATATTAAGGGATACATGA
      ***  ******  **********************  **  ******  ******  *****  *****  **********  ***
```

```
AOR1    TAAGTCCTGAAATACTTGGCTATCCAGAAAAACTTGATAGACTTGCAGTGGAAGGAAAAGCAGGATATGCTAGAGTATTC
AOR2    TTAACCCTGAAATATTAGGTTATCCTGAAAAAACTTGATAGATTTGCATTAGATGGTAAAGCAGCTTATGCCAAATTATTT
        * *  ********** * ** ****** ********************** ***** * ** ** ******* * * ****

AOR1    CATGATTTAACAGCTGTTATAGATTCACTTGGATTATGTATTTTTACAACATTTGGTCTTGGTGCACAGGATTATGTTGA
AOR2    CATGATTTAACTGCTGTAATTGATTCTTTAGGATTGTGCATATTCACTACATTTGGGCTTGGAATACAGGATTATGTAGA
        ********** ****** ** ****** * ****** ** ** ** ********* *****    ***************** **

AOR1    TATGTATAATGCAGTAGTTGGTGGAGAATTACATGATGTAAATTCTTTAATGTTAGCTGGAGATAGAAATATGGACTTTAG
AOR2    TATGTATAATGCAGTAGTAGGAGAATCTACTTATGATGCAGATTCACTATTAGAGGCAGGAGATAGAATCTGGACTCTTG
        ********************** ** * *          ******* * ****  **  * *     ** ********** ******* * *

AOR1    AAAAAATATTTAACTTAAAAGCAGGCATAGATAGTTCACAGGATACTCTTCCAAAGAGATTGCTTGAAGAACAAATTCCA
AOR2    AGAAATTATTTAATCTTGCAGCTGGAATAGACAGCCAGGATACTCTACCAAAGAGATTGTTAGAAGAACCTATTCCA
        * *** *********   *    *** ** ****** **          ********************** * ********  *******

AOR1    GAAGGACCATCAAAAGGAGAAGTTCATAAGTTAGATGTACTACTACCTGAATATTATTCAGTACGTGGATGGGATAAAAA
AOR2    GATGGCCCATCAAAGGGAGAAGTTCATAGGCTAGATGTTCTTCTGCCAGAATATTACTCAGTACGAGGATGGAGTAAAGA
        ** ** ********** ********************** * ******** ** ** ** ********** ********** ******  **** *

AOR1    TGGTATTCCTACAGAGGAAACGTTAAAGAAATTAGGATTAGATGAATACGTAGGTAAGCTTTAG
AOR2    GGGTATACCTACAGAAGAAACATTAAAGAAATTAGGATTAGATGAATATATAGGTAAGTTCTAG
          ****** ********** ****** ********************************  ********** * ***
```

Fig. 4B. Amino acid sequence alignment of AOR1 and AOR2.

```
CAUTOAOR1   MYGYDGKVLRINLKERTCKSENLDLDKAKKFIGCRGLGVKTLFDEIDPKIDALSPENKFIIVTGPLTGAPVPTSGRFMVV
CAUTOAOR2   MYGYKGKVLRINLSSKTYIVEELKIDKAKKFIGARGLGVKTLFDEVDPKVDPLSPDNKFIIAAGPLTGAPVPTSGRFMVV
            *****.**********..:*   *:*..:*********.**************:****:*.****:******.:*********************
```

```
CAUTOAOR1  TKAPLITGTIGISNSGGKWGVDLKKAGWDMIIVEDKADSPVYIEIVDDKVEIKDASQLWGKVTSETTKELEKITENKSKVL
CAUTOAOR2  TKSPLTGTIAIANSGGKWGAEFKAAGYDMIIVEGKSDKEVYVNIVDDKVEFRDASHVWGKLTEETTKMLQQETDSRAKVL
           **:********.*:*********.::*  **:*******.*:*.  **::*********::*****:*.**** *::  *:.::***

CAUTOAOR1  CIGPAGERLSLMAAVMNDVDRTAARGGVGAVMGSKNLKAITVKGTGKIALADKEKVKKVSVEKITTLKNDPVAGQGMPTY
CAUTOAOR2  CIGPAGEKLSLMAAVMNDVDRTAGRGGVGAVMGSKNLKAIVVKGSGKVKLFDEQKVKEVALEKTNILRKDPVAGGGLPTY
           ********:**********************.***:***: * *::****:*::** . *::****** *:****

CAUTOAOR1  GTAILVNIINENGVHPVKNFQESYTNQADKISGETLTANQLVRKNPCYSCPIGCGRWVRLKDGTECGGPEYETLWCFGSD
CAUTOAOR2  GTAVLVNIINENGVHPVKNFQKSYTDQADKISGETLTKDCLVRKNPCYRCPIACGRWVKLDDGTECGGPEYETLWSFGSD
           ***:**********************:***:********** : ********** ***.*****:*.*.**************.****

CAUTOAOR1  CGSYDLDAINEANMLCNEYGIDTITCGATIAAAMELYQRGYIKDEEIAGDNLSLKWGDTESMIGWIKRMVYSEGFGAKMT
CAUTOAOR2  CDVYDINAVNTANMLCNEYGLDTITAGCTIAAAMELYQRGYIKDEEIAADGLSLNWGDAKSMVEWVKKMGLREGFGDKMA
           *. **::*:* **********:*****.*.**********************.*.***:***::***: *:*:*   **** **:

CAUTOAOR1  NGSYRLCEGYGAPEYSMTVKKQEIPAYDPRGIQGHGITYAVNNRGGCHIKGYMINPEILGYPEKLDRFALDGKAAYAKLF
CAUTOAOR2  DGSYRLCDSYGVPEYSMTVKKQELPAYDPRGIQGHGITYAVNNRGGCHIKGYMVSPEILGYPEKLDRLAVEGKAGYARVF
           :*******:.***.**********:**********************************:.**************:*::***.**::*

CAUTOAOR1  HDLTAVIDSLGLCIFTTFGLGIQDYVDMYNAVVGESTYDADSLLEAGDRIWTLEKLFNLAAGIDSSQDTLPKRLLEEPIP
CAUTOAOR2  HDLTAVIDSLGLCIFTTFGLGAQDYVDMYNAVVGGELHDVNSLMLAGDRIWTLEKIFNLKAGIDSSQDTLPKRLLEEQIP
           ********************** ********** . :*.:**: ********** ********:*** ********** **

CAUTOAOR1  DGPSKGEVHRLDVLLPEYYSVRGWSKEGIPTEETLKKLGLDEYIGKF
CAUTOAOR2  EGPSKGEVHKLDVLLPEYYSVRGWDKNGIPTEETLKKLGLDEYVGKL
           :**********:**********************.*:*********************:**:
```

Fig. 5A. Multiple amino acid sequence alignment of *C. autoethanogenum* AOR1 with two other AOR from homoacetogens.

```
CAUTOAOR1  MYGYKGKVLRINLSSKTYIVEELKIDKAKKFIGARGLGVKTLFDEVDPKVDPLSPDNKFIIAAGPLTGAPVPTSGRFMVV
CLUNGAOR1  MYGYKGKVLRINLSSKTYIVEELKIDKAKKFIGARGLGVKTLFDEVDPKVDPLSPDNKFIIAAGPLTGAPVPTSGRFMVV
CRAGSAOR1  MYGYSGKVLRINLSNKTYKAEELKIDEAKKFIGARGLGVKTLLDEIDPKIDPLSPDNKFIIATGPLTGAPVPTSGRFMVI
           ****.**********.**** .*******:**********:**:***:**********:**********:**********:

CAUTOAOR1  TKSPLTGTIAIANSGGKWGAEFKAAGYDMIIVEGKSDKEVYVNIVDDKVEFRDASHVWGKLTEETTKMLQQETDSRAKVL
CLUNGAOR1  TKSPLTGTIAIANSGGKWGAEFKAAGYDMIIVEGKSDKEVYVNIVDDKVEFRDASHVWGKLTEETTKMLQQETDSRAKVL
CRAGSAOR1  TKAPLTGTIGIANSGGKWGAELKTAGYDMVIVEGKSDKPVYVNIVDDKVEFKDASHVWGKLTEETTKMLQNEIDAKAKVL
           **:*******.**********:*:******:********** **********:**********:* *::*****

CAUTOAOR1  CIGPAGEKLSLMAAVMNDVDRTAGRGGVGAVMGSKNLKAIVVKGSGKVKLFDEQKVKEVALEKTNILRKDPVAGGGLPTY
CLUNGAOR1  CIGPAGEKLSLMAAVMNDVDRTAGRGGVGAVMGSKNLKAIVVKGSGKVKLFDEQKVKEVALEKTNILRKDPVAGGGLPTY
CRAGSAOR1  CIGPAGENLSLMAAVMNDIDRTAGRGGVGAVMGSKNLKAIVVKGSGKVKLFDEEKVKAVSLQKSDILRKDPVAGGGLPTY
           ********:**********:**********:**********:*** *:*:*::**********:**********

CAUTOAOR1  GTAVLVNIINENGVHPVKNFQKSYTDQADKISGETLTKDCLVRKNPCYRCPIACGRWVKLDDGTECGGPEYETLWSFGSD
CLUNGAOR1  GTAVLVNIINENGVHPVKNFQKSYTDQADKISGETLTKDCLVRKNPCYRCPIACGRWVKLDDGTECGGPEYETLWSFGSD
CRAGSAOR1  GTAVLVNIINENGINPVRNFQESYTDEADKVSGETMTQECLVRKNPCYRCPIACGRWVRLDDGTECGGPEYETLWSFGSD
           **********:::**:***:****:****:*::**********:**********:**********:**********

CAUTOAOR1  CDVYDINAVNTANMLCNEYGLDTITAGCTIAAAMELYQRGYIKDEEIAADGLSLNWGDAKSMVEWVKKMGLREGFGDKMA
CLUNGAOR1  CDVYDINAVNTANMLCNEYGLDTITAGCTIAAAMELYQRGYIKDEEIAADGLSLNWGDAKSMVEWVKKMGLREGFGDKMA
CRAGSAOR1  CDVYDLNAVNKANMLCNEYGLDTISAGATIASAMELYQRGYIKDEEIAADGLSLKWGDAKSMVEWVKKMGRREGFGGKMA
           *****:*****.**********:**.***:**********:**********:********** ******.***

CAUTOAOR1  DGSYRLCDSYGVPEYSMTVKKQELPAYDPRGIQGHGITYAVNNRGGCHIKGYMVSPEILGYPEKLDRLAVEGKAGYARVF
CLUNGAOR1  DGSYRLCDSYGVPEYSMTVKKQELPAYDPRGIQGHGITYAVNNRGGCHIKGYMVSPEILGYPEKLDRLAVEGKAGYARVF
CRAGSAOR1  DGSYRLCESYGVPQYSMSVKKQELPAYDPRGAQGHGLTYAVNNRGGCHIKGYMISPEILGYPEKLDRFSIEGKPAYAKVF
           *******:*****:***:**********  ****:**********:**********:::***..**:**
```

```
CAUTOAOR1   HDLTAVIDSLGLCIFTTFGLGAQDYVDMYNAVVGGELHDVNSLMLAGDRIWTLEKIFNLKAGIDSSQDTLPKRLLEEQIP
CLUNGAOR1   HDLTAVIDSLGLCIFTTFGLGAQDYVDMYNAVVGGELHDVNSLMLAGDRIWTLEKIFNLKAGIDSSQDTLPKRLLEEQIP
CRAGSAOR1   HDLTAVIDSLGLCIFTTFGLGAQDYVDMYNAVVGGELHDVDSLMLAGDRVWTLEKIFNLKAGVGSSQDTLPKRLLEEEVV
            ****************************************:**********:*****************:.**************::

CAUTOAOR1   EGPSKGEVHKLDVLLPEYYSVRGWDKNGIPTEETLKKLGLDEYVGKL
CLUNGAOR1   EGPSKGEVHKLDVLLPEYYSVRGWDKNGIPTEETLKKLGLDEYVGKL
CRAGSAOR1   EGPSKGHVHRLDELVPEYYSVRGWDKNGVPTEETLKKLGLEEYIGKI
            *******.**:** *:**********:**********:**:**:**:
```

Fig. 5B. Multiple amino acid sequence alignment of *C. autoethanogenum* AOR 2 with two other AORs from homoacetogens.

```
CAUTOAOR2   MYGYDGKVLRINLKERTCKSENLDLDKAKKFTGCRGLGVKTLFDEIDPKIDALSPENKFIIVTGPLTGAPVPTSGRFMVV
CLJUNAOR2   MYGYDGKVLRINLKERTCKSENLDLDKAKKFTGCRGLGVKTLFDEIDPKIDALSPENKFIIVTGPLTGAPVPTSGRFMVV
CRAGSAOR2   MYGYNGKVLRINLKERTCKSENLDLDKAKKFTGCRGLGVKTLFDEIDPKIDALSPENKFIIVTGPLTGAPVPTSGRFMVV
            ****:**********************************************************************:*****

CAUTOAOR2   TKAPLTGTIGISNSGGKWGVDLKKAGWDMIIVEDKADSPVYIEIVDDKVEIKDASQLWGKVTSETTKELEKITENKSKVL
CLJUNAOR2   TKAPLTGTIGISNSGGKWGVDLKKAGWDMIIVEDKADSPVYIEIVDDKVEIKDASQLWGKVTSETTKELEKITENKSKVL
CRAGSAOR2   TKAPLTGTIGISNSGGKWGVDLKKAGWDMIIVEDKADSPVYIEIVDDKVEIKDASQLWGKVTSETTKELEKITENRSKVL
            **********************************************************************:*****

CAUTOAOR2   CIGPAGERLSLMAAVMNDVDRTAARGGVGAVMGSKNLKAITVKGTGKIALADKEKVKKVSVEKITTLKNDPVAGQGMPTY
CLJUNAOR2   CIGPAGERLSLMAAVMNDVDRTAARGGVGAVMGSKNLKAITVKGTGKIALADKEKVKKVSVEKITTLKNDPVAGQGMPTY
CRAGSAOR2   CIGPAGERLSLMAAVMNDVDRTAARGGVGAVMGSKNLKAITVKGTGKIALADKEKVKKVSVEKITTLKNDPVAGQGMPTY
            **********************************************************************************

CAUTOAOR2   GTAILVNIINENGVHPVKNFQESYTNQADKISGETLTANQLVRKNPCYSCPIGCGRWVRLKDGTECGGPEYETLWCFGSD
CLJUNAOR2   GTAILVNIINENGVHPVKNFQESYTNQADKISGETLTANQLVRKNPCYSCPIGCGRWVRLKDGTECGGPEYETLWCFGSD
CRAGSAOR2   GTAILVNIINENGVHPVNNFQESYTDQADKISGETLTANQLVRKNPCYSCPIGCGRWVRLKDGTECGGPEYETLWCFGSD
            **********************:***********************************************************
```

```
CAUTOAOR2  CGSYDLDAINEANMLCNEYGIDTITCGATIAAAMELYQRGYIKDEEIAGDNLSLKWGDTESMIGWIKRMVYSEGFGAKMT
CLJUNAOR2  CGSYDLDAINEANMLCNEYGIDTITCGATIAAAMELYQRGYIKDEEIAGDNLSLKWGDTESMIGWIKRMVYSEGFGAKMT
CRAGSAOR2  CGSYDLDAINEANMLCNEYGIDTITCGATIAAAMELYQRGYVKDEEIAGDNLSLKWGDTESMIGWIKKMVYSEGFGAKMT
           ************************************************************:***********************

CAUTOAOR2  NGSYRLCEGYGAPEYSMTVKKQEIPAYDPRGIQGHGITYAVNNRGGCHIKGYMINPEILGYPEKLDRFALDGKAAYAKLF
CLJUNAOR2  NGSYRLCEGYGAPEYSMTVKKQEIPAYDPRGIQGHGITYAVNNRGGCHIKGYMINPEILGYPEKLDRFALDGKAAYAKLF
CRAGSAOR2  NGSYRLCEGYGVPEYSMTVKKQEIPAYDPRGIQGHGITYAVNNRGGCHIKGYMINPEILGYPEKLDRFALDGKAAYAKMM
           **********.************************************************************************::

CAUTOAOR2  HDLTAVIDSLGLCIFTTFGLGIQDYVDMYNAVVGESTYDADSLLEAGDRIWTLEKLFNLAAGIDSSQDTLPKRLLEEPIP
CLJUNAOR2  HDLTAVIDSLGLCIFTTFGLGIQDYVDMYNAVVGESTYDADSLLEAGDRIWTLEKLFNLAAGIDSSQDTLPKRLLEEPIP
CRAGSAOR2  HDLTAVIDSLGLCIFTTFGLGIQDYVDMYNAVVGESTCDSDSLLEAGDRVWTLEKLFNLAAGIDSSQDTLPKRLLEEPIP
           ************************************ *:**********:**************************************

CAUTOAOR2  DGPSKGEVHRLDVLLPEYYSVRGWSKEGIPTEETLKKLGLDEYIGKF
CLJUNAOR2  DGPSKGEVHRLDVLLPEYYSVRGWSKEGIPTEETLKKLGLDEYIGKF
CRAGSAOR2  DGPSKGHVHRLDVLLPEYYSVRGWSKEGIPTEETLKKLGLDEYIGKF
           *******.*************************************
```

Fig. 6

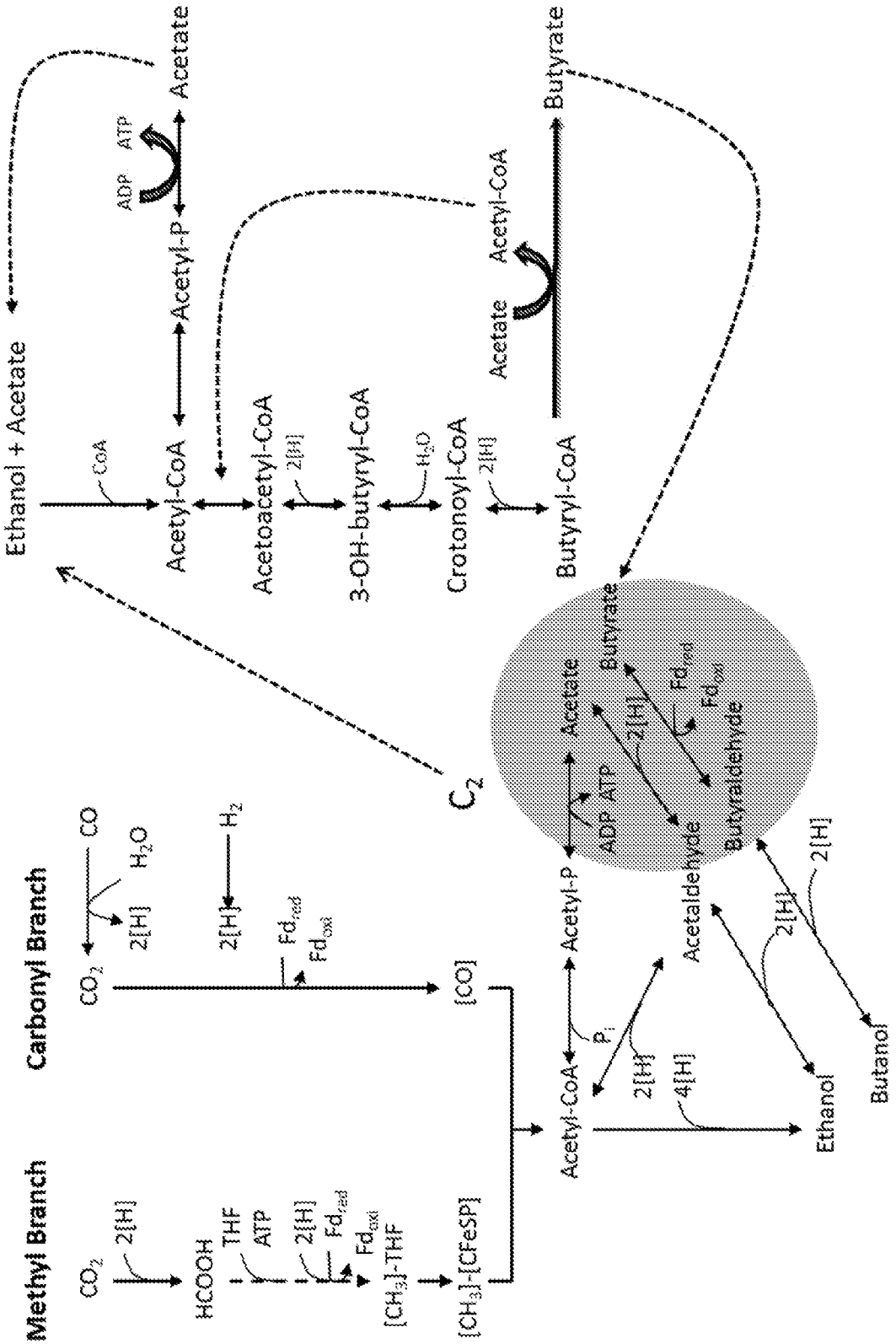

Ethanol + Acetate
CoA
Acetyl-CoA
Acetoacetyl-CoA
2[H]
3-OH-butyryl-CoA
H_2O
Crotonoyl-CoA
2[H]
Butyryl-CoA
Acetyl-P
ADP ATP
Acetate
Acetate
Acetyl-CoA
Butyrate
C_2
Acetate
Butyrate
ADP ATP
2[H]
Fd_{red}
Fd_{oxi}
Butyraldehyde
Acetaldehyde
Acetyl-P
Acetyl-CoA
P_i
2[H]
4[H]
Ethanol
2[H]
2[H]
Butanol
Carbonyl Branch
CO_2
CO
H_2O
2[H]
H_2
2[H]
Fd_{red}
Fd_{oxi}
[CO]
Methyl Branch
CO_2
2[H]
HCOOH
THF
ATP
2[H]
Fd_{red}
Fd_{oxi}
$[CH_3]$-THF
$[CH_3]$-[CFeSP]

Fig. 7

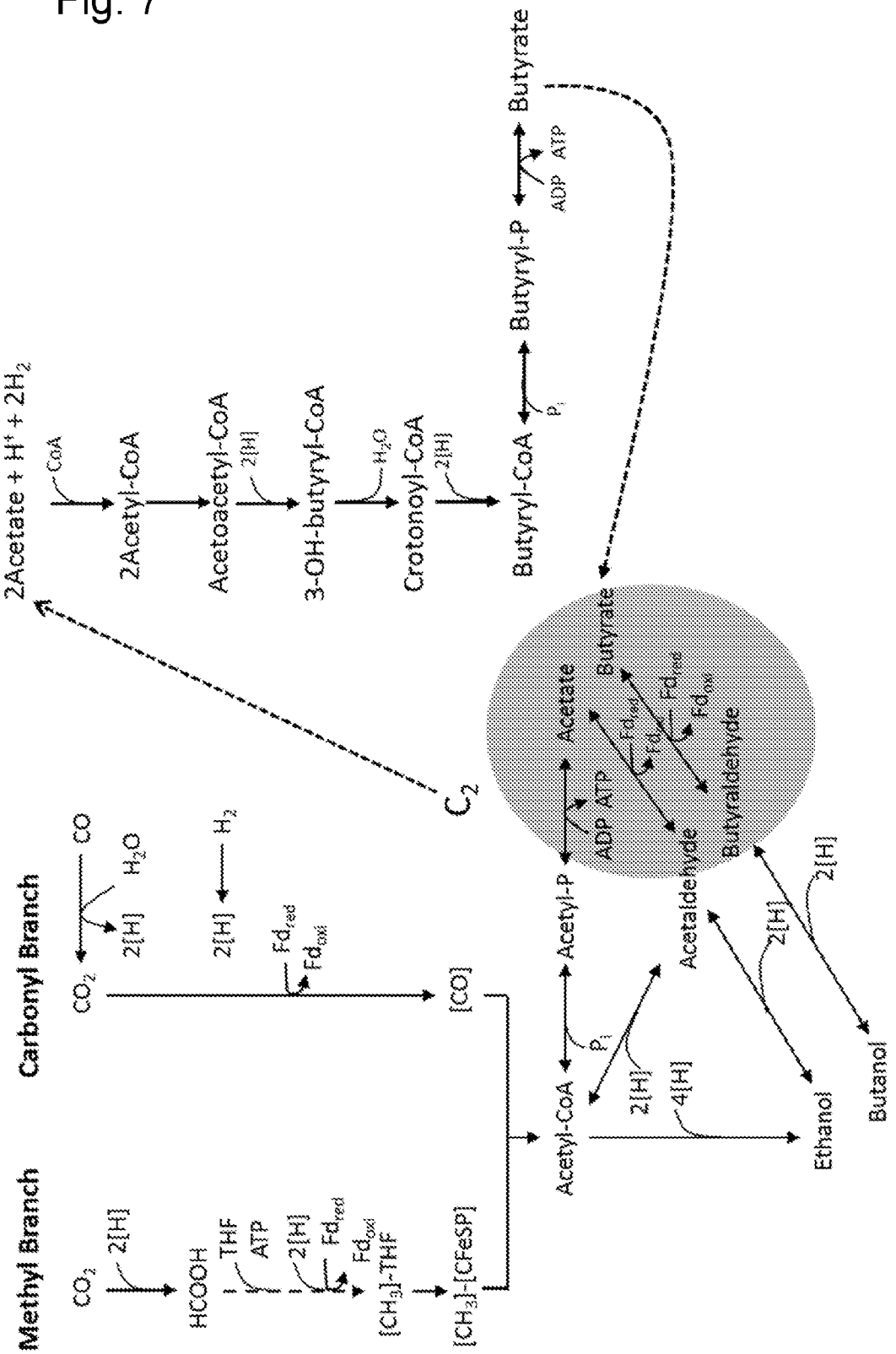

Methyl Branch
Carbonyl Branch
2Acetate + H+ + 2H2
2Acetyl-CoA
Acetoacetyl-CoA
3-OH-butyryl-CoA
Crotonoyl-CoA
Butyryl-CoA
Butyryl-P
Butyrate
CoA
2[H]
H2O
ADP
ATP
Pi
C2
CO2
CO
H2
HCOOH
THF
[CH3]-THF
[CFeSP]
[CO]
Fdred
Fdoxi
Acetyl-CoA
Acetyl-P
Acetate
Acetaldehyde
Butyraldehyde
Ethanol
Butanol
4[H]

Fig. 8

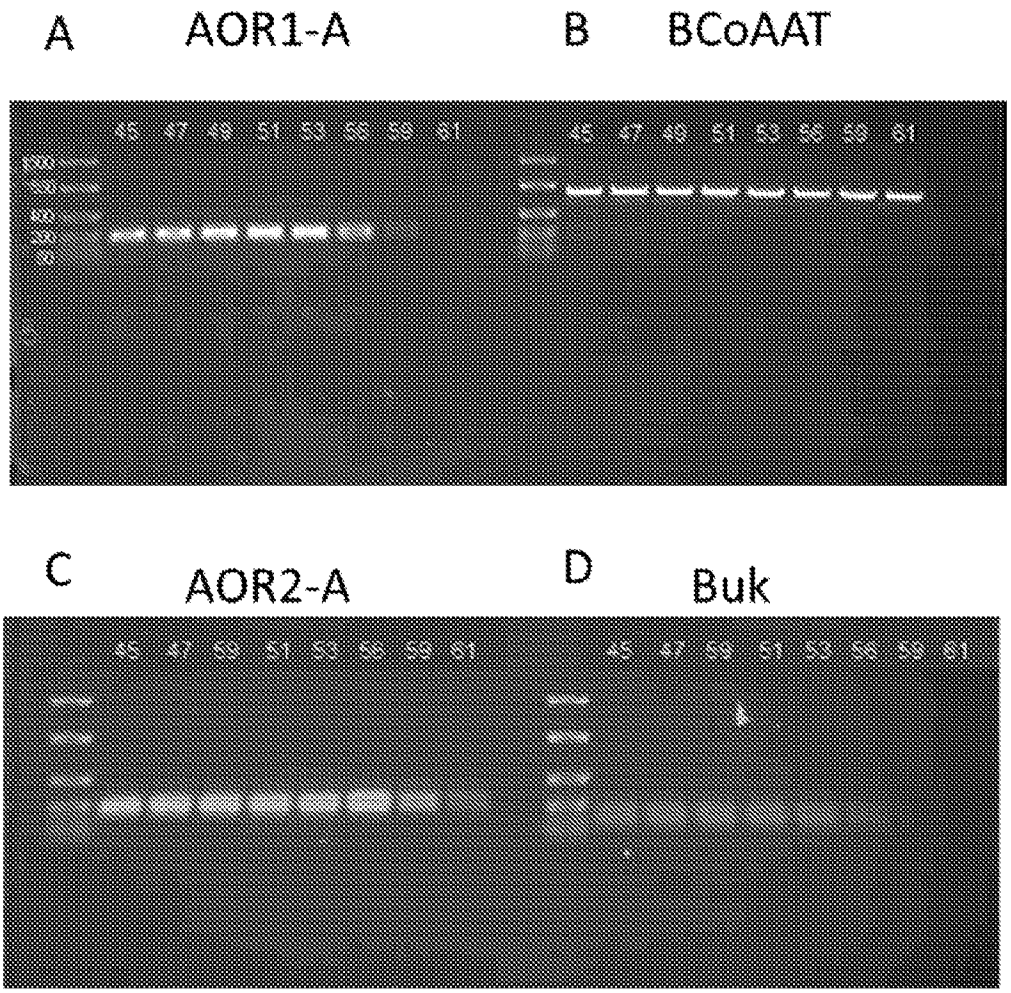


Fig. 9

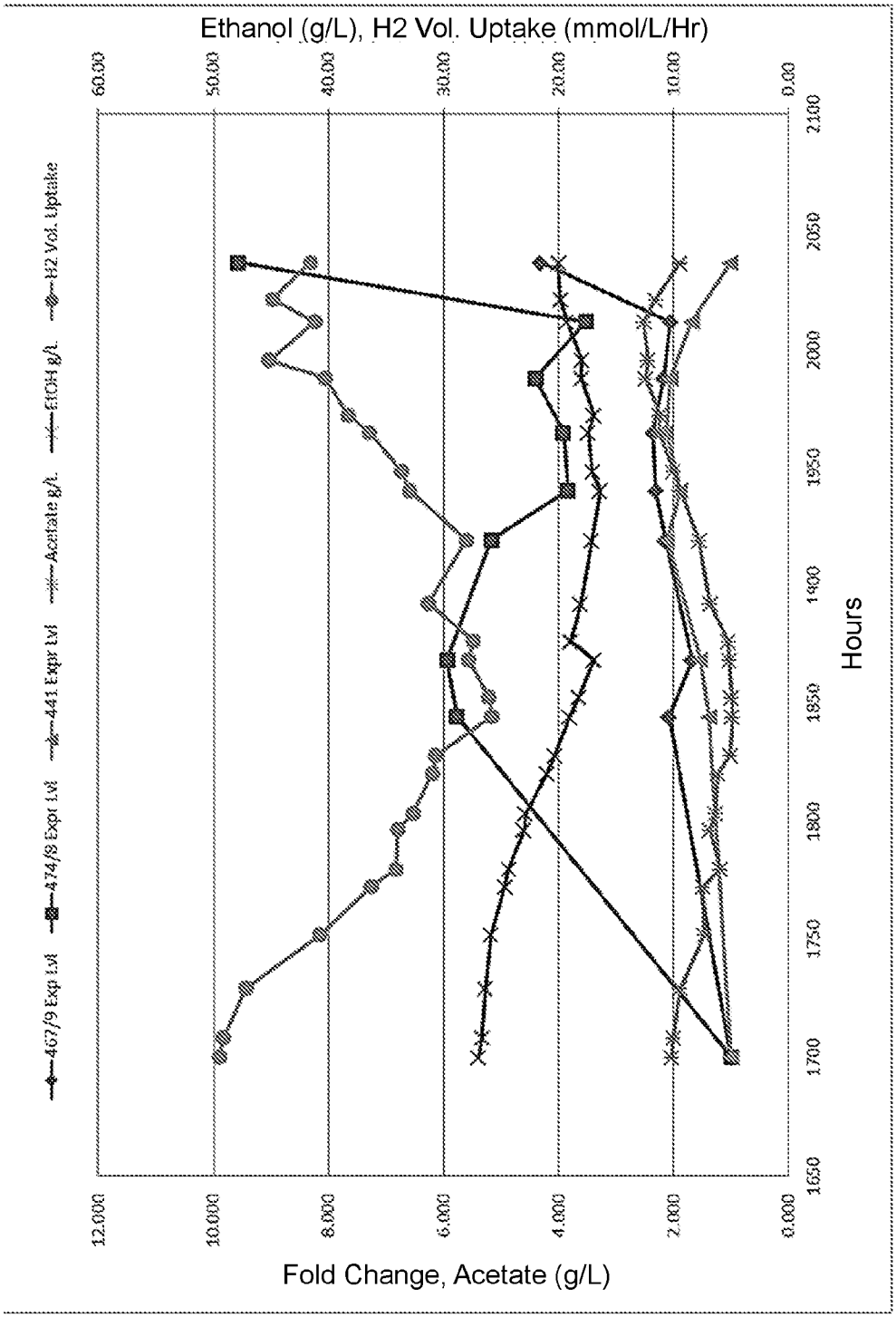

Ethanol (g/L), H2 Vol. Uptake (mmol/L/Hr)
60.00
50.00
40.00
30.00
20.00
10.00
0.00
2100
2050
2000
1950
1900
1850
1800
1750
1700
1650
Hours
12.000
10.000
8.000
6.000
4.000
2.000
0.000
Fold Change, Acetate (g/L)
467/9 Exp Lvl
474/8 Expr Lvl
441 Expr Lvl
Acetate g/L
EtOH g/L
H2 Vol. Uptake

Fig. 10

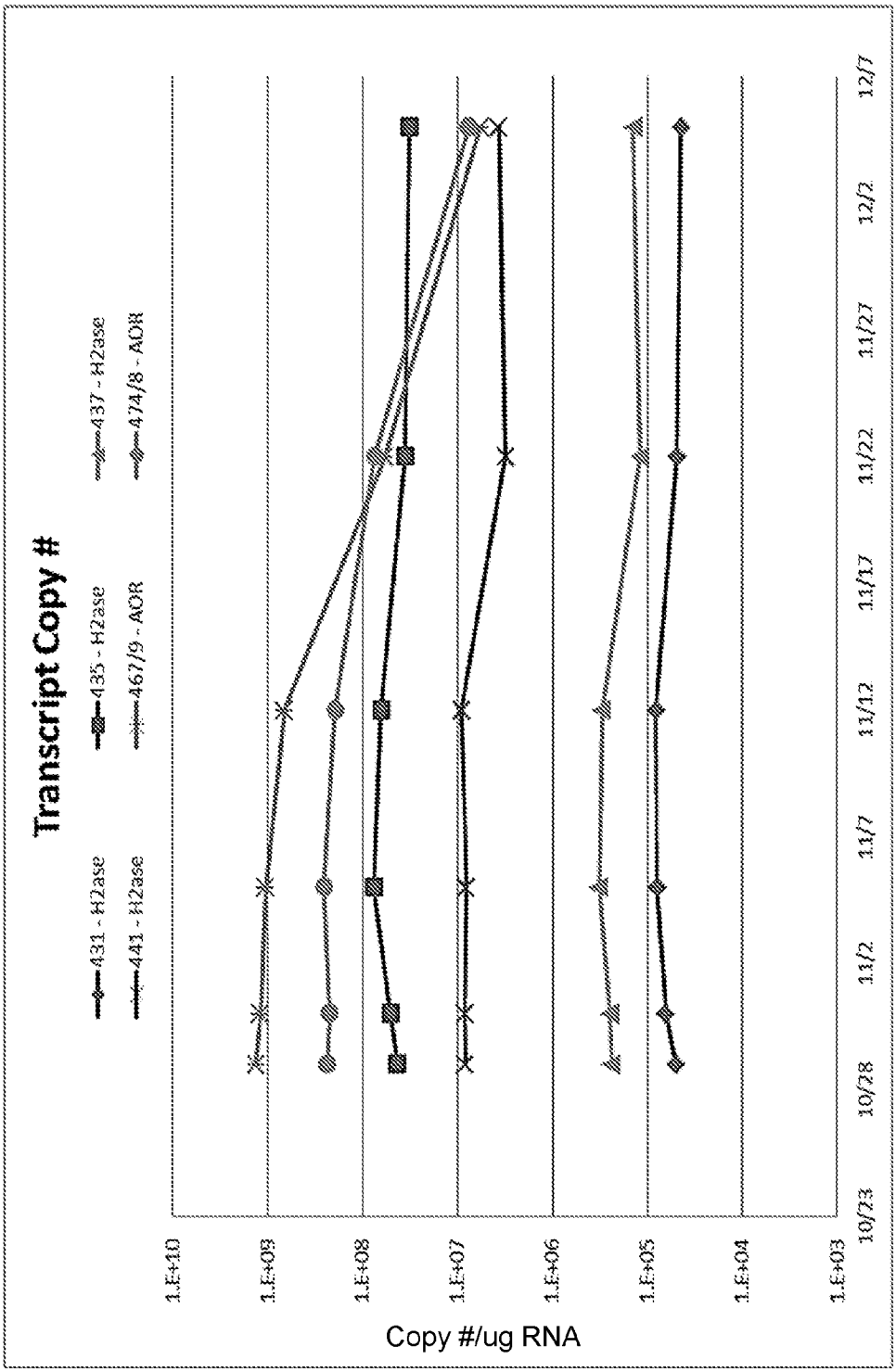

Transcript Copy #
431 - H2ase
435 - H2ase
437 - H2ase
441 - H2ase
467/9 - AOR
474/8 - AOR
1.E+10
1.E+09
1.E+08
1.E+07
1.E+06
1.E+05
1.E+04
1.E+03
10/23
10/28
11/2
11/7
11/12
11/17
11/22
11/27
12/2
12/7
Copy #/ug RNA

Fig. 11

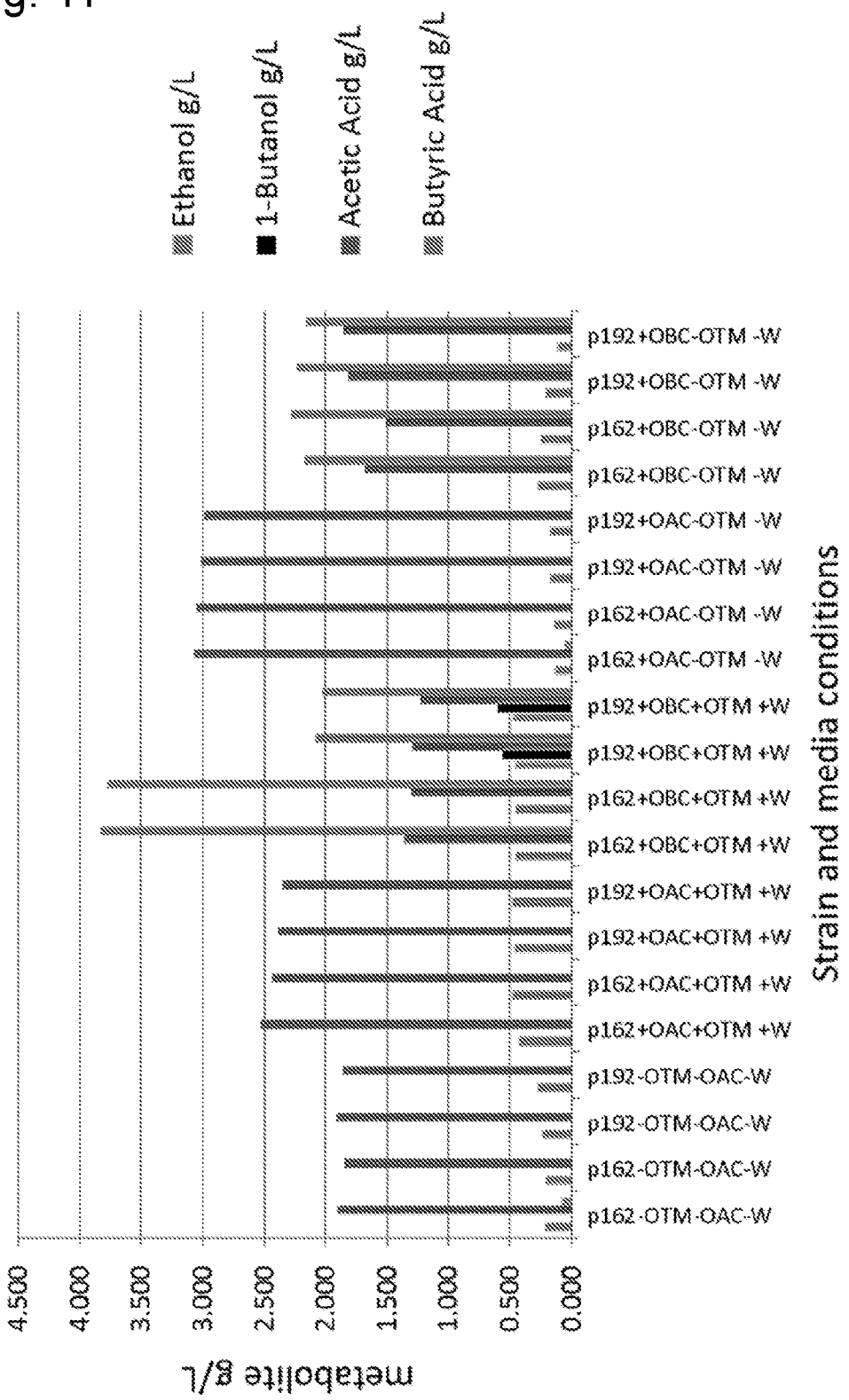

Ethanol g/L
1-Butanol g/L
Acetic Acid g/L
Butyric Acid g/L
p192+OBC-OTM -W
p192+OBC-OTM -W
p162+OBC-OTM -W
p162+OBC-OTM -W
p192+OAC-OTM -W
p192+OAC-OTM -W
p162+OAC-OTM -W
p162+OAC-OTM -W
p192+OBC+OTM +W
p192+OBC+OTM +W
p162+OBC+OTM +W
p162+OBC+OTM +W
p192+OAC+OTM +W
p192+OAC+OTM +W
p162+OAC+OTM +W
p162+OAC+OTM +W
p192-OTM-OAC-W
p192-OTM-OAC-W
p162-OTM-OAC-W
p162-OTM-OAC-W
Strain and media conditions
4,500
4,000
3,500
3,000
2,500
2,000
1,500
1,000
0,500
0,000
metabolite g/L

Fig. 12

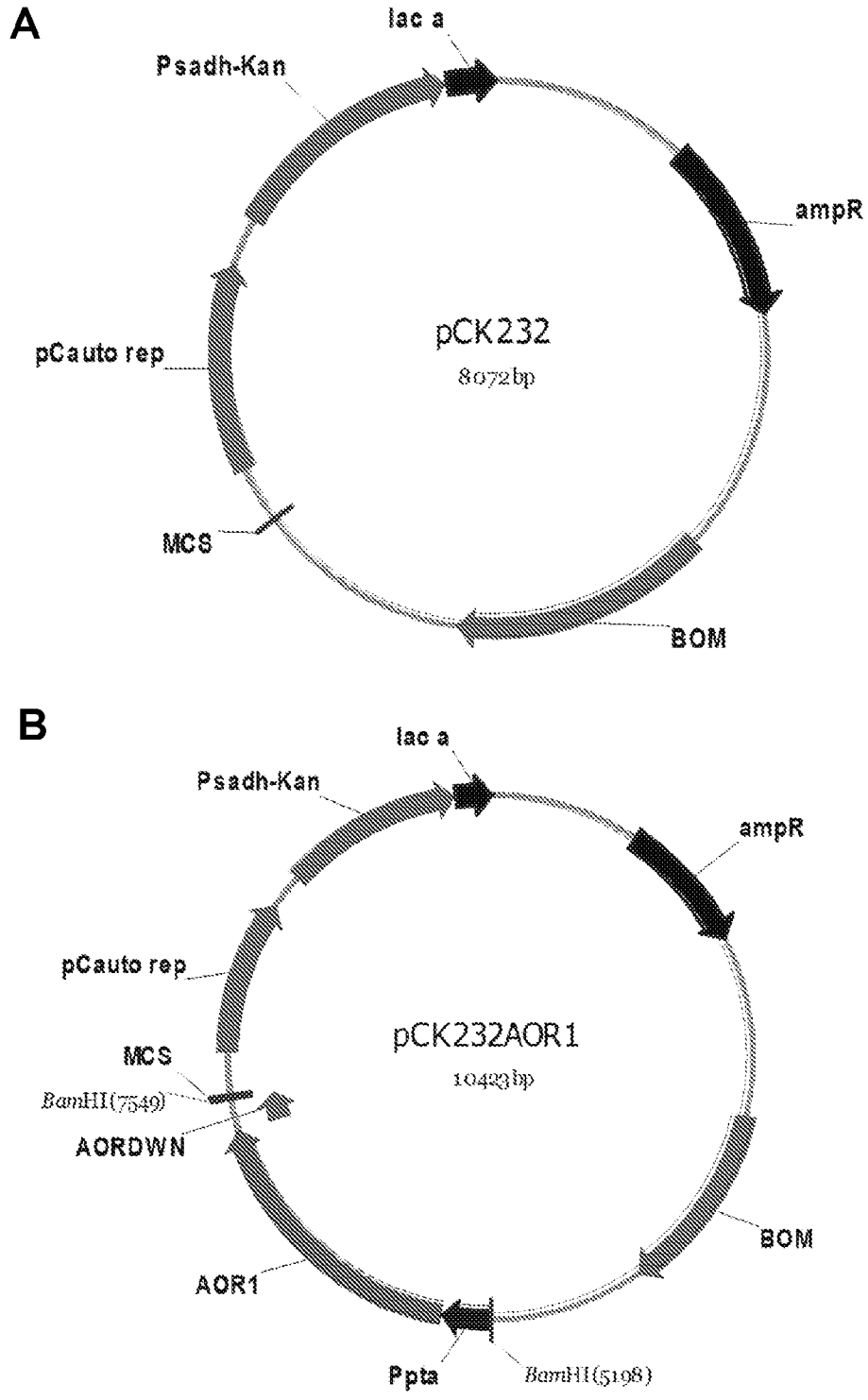

A
lac a
Psadh-Kan
ampR
pCK232
8072bp
pCauto rep
MCS
BOM
B
lac a
Psadh-Kan
ampR
pCauto rep
pCK232AOR1
10423bp
MCS
BamHI(7549)
AORDWN
BOM
AOR1
Ppta
BamHI(5198)

Fig. 13

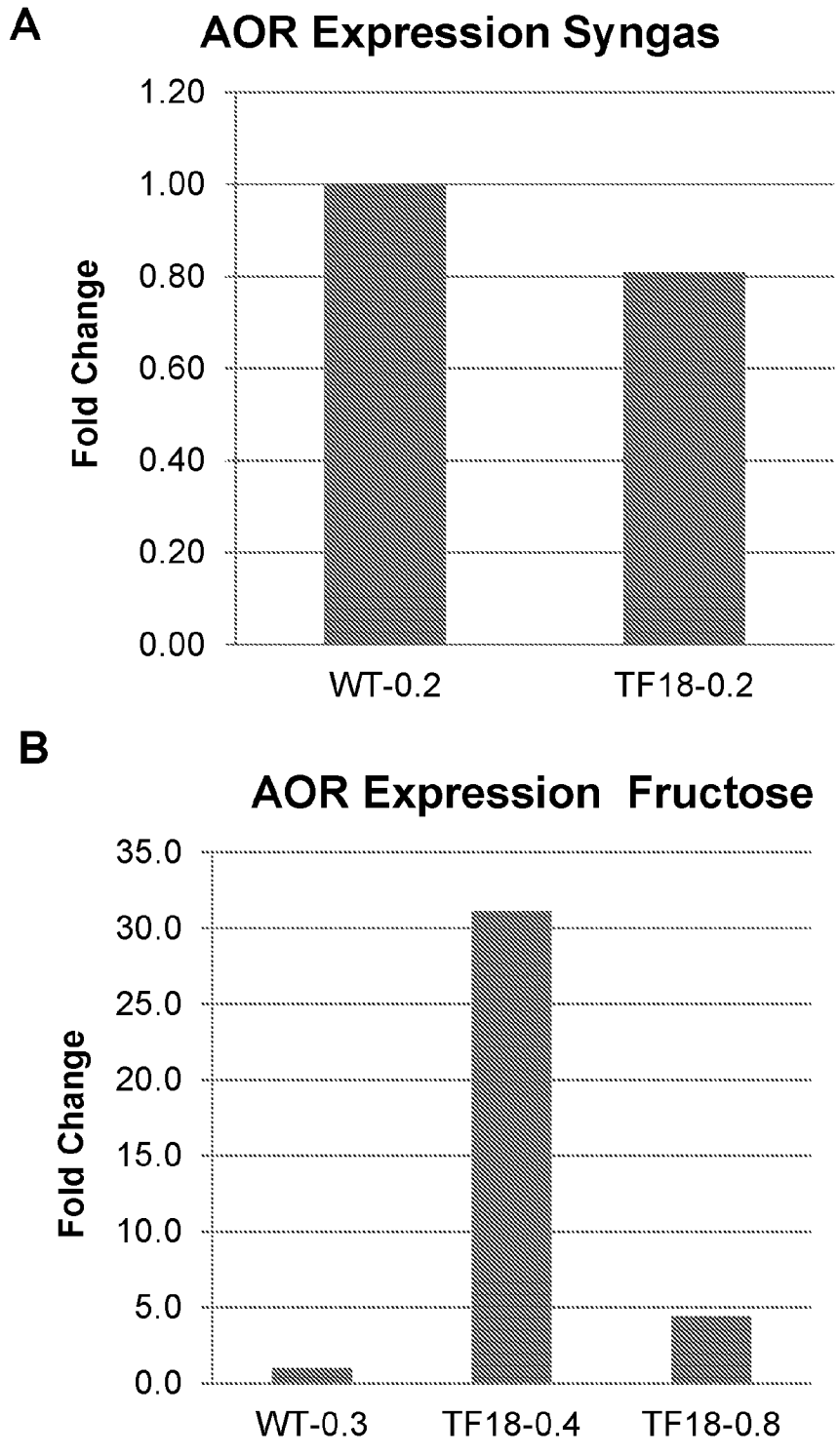
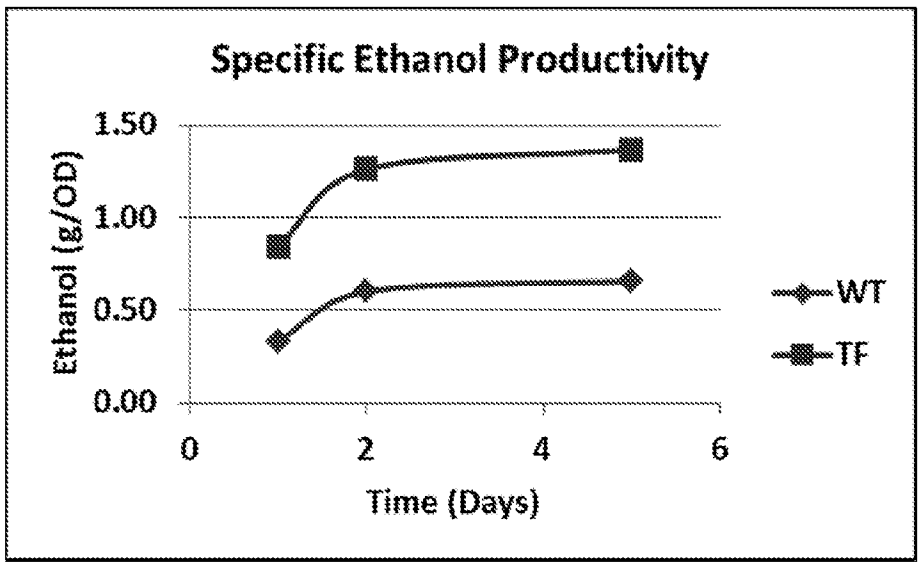

A
AOR Expression Syngas
Fold Change
1.20
1.00
0.80
0.60
0.40
0.20
0.00
WT-0.2
TF18-0.2
B
AOR Expression Fructose
Fold Change
35.0
30.0
25.0
20.0
15.0
10.0
5.0
0.0
WT-0.3
TF18-0.4
TF18-0.8

Fig. 14

A


Specific Ethanol Productivity
Ethanol (g/OD)
1.50
1.00
0.50
0.00
0
2
4
6
Time (Days)
WT
TF

B

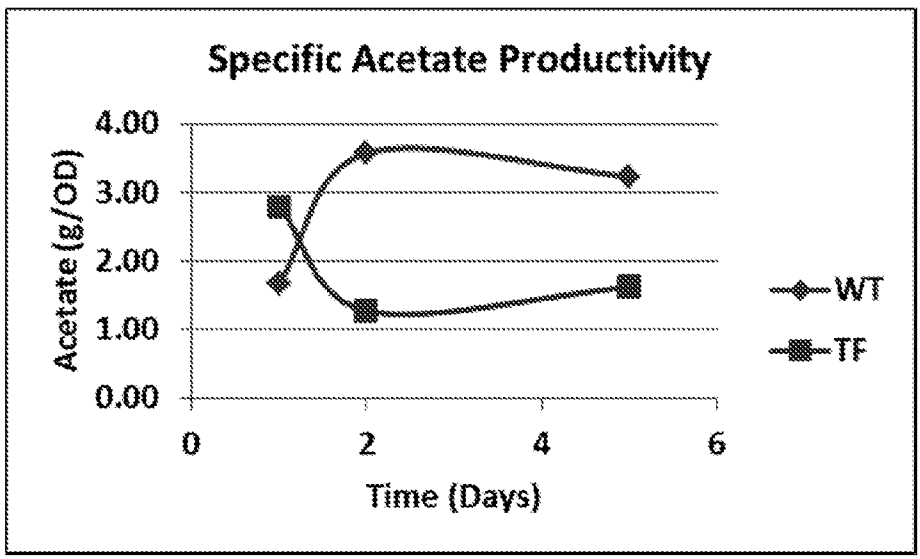
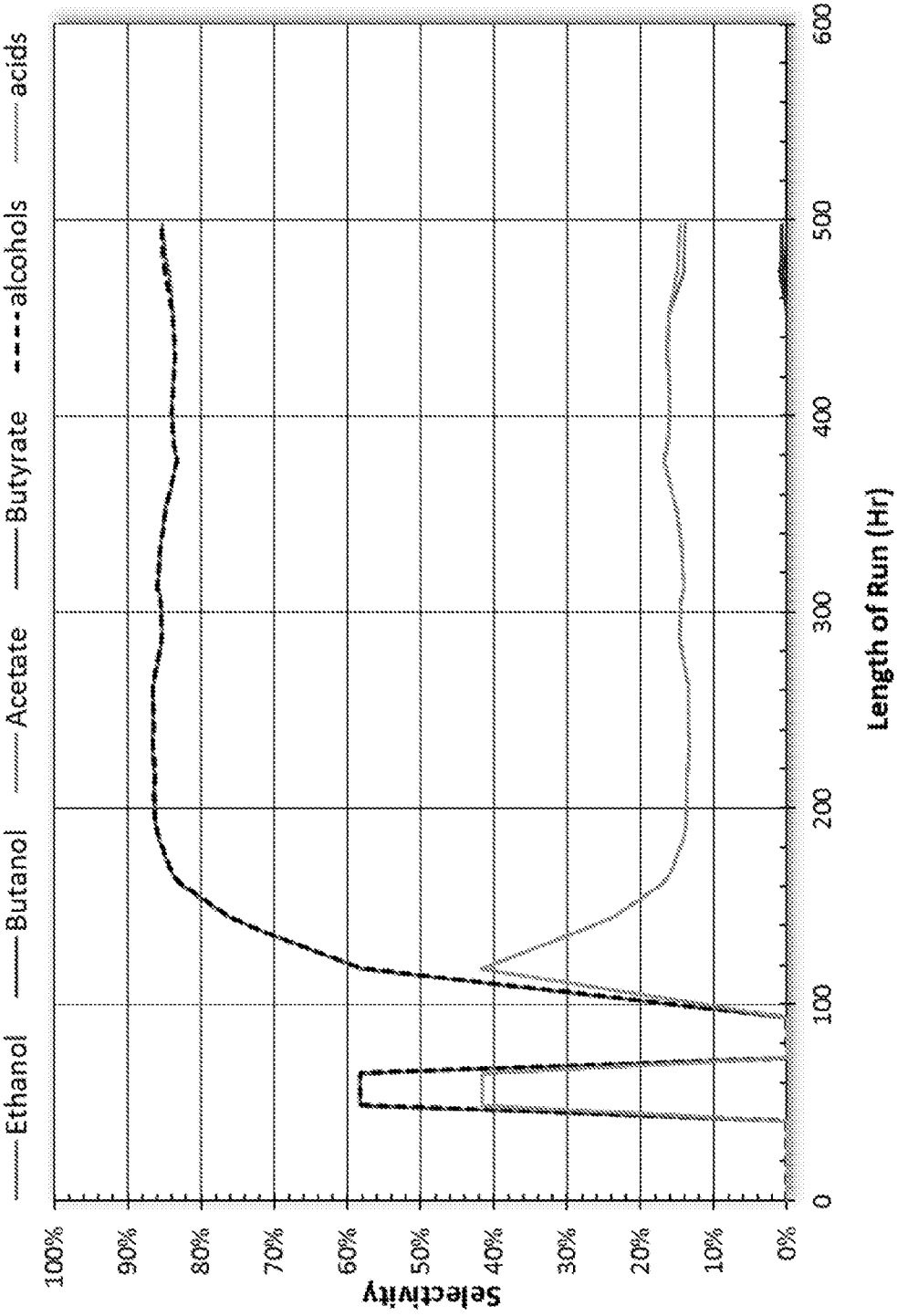

Specific Acetate Productivity
Acetate (g/OD)
4.00
3.00
2.00
1.00
0.00
0
2
4
6
Time (Days)
WT
TF

Fig. 15


Ethanol
Butanol
Acetate
Butyrate
alcohols
acids
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Selectivity
0
100
200
300
400
500
600
Length of Run (Hr)