

**GENE EXPRESSION SYSTEM FOR PROBIOTIC MICROORGANISMS****STATEMENT AS TO FEDERALLY SPONSORED RESEARCH**

**[0001]** None.

**REFERENCE TO RELATED APPLICATIONS**

**[0002]** This application claims the benefit of the priority dates of U.S. provisional application, 62/558,346, filed September 13, 2017, and United States application 16/048,147, filed July 27, 2018, the contents all of which are incorporated herein by reference in their entirety.

**SEQUENCE LISTING**

**[0003]** The instant application contains a Sequence Listing which has been submitted electronically in ASCII format and is hereby incorporated by reference in its entirety. Said ASCII copy, created on June 21, 2018, is named 1642-001-US\_SL.TXT and is 29,323 bytes in size.

**BACKGROUND**

**[0004]** When one consumes alcohol, the major pathway for its removal from the bloodstream is via oxidation in the liver to acetaldehyde via alcohol dehydrogenase enzymes. Acetaldehyde is subsequently oxidized in the liver to acetate via acetaldehyde dehydrogenase enzymes. When alcohol is consumed in large quantities and/or at a fast rate, the highly toxic intermediary acetaldehyde can accumulate and subsequently be released back into the bloodstream, where – abetted by its high solubility – it can act throughout the body. Not only is acetaldehyde a known carcinogen, its toxic effects are a well-studied and documented cause of many of the effects of an alcohol hangover. Indeed, removing acetaldehyde has been demonstrated to reduce hangover symptoms. PMID: 16554376 (PubMed ID). Conversely, when the body's ability to oxidize acetaldehyde to acetate is inhibited either chemically (e.g. with disulfiram) or genetically (e.g., single nucleotide polymorphisms in acetaldehyde dehydrogenase genes common in East Asian populations), extremely amplified hangover symptoms are experienced.

**[0005]** There have been several previous attempts to reduce, eliminate, or prevent the effects of a hangover. Some have attempted to enzymatically or otherwise reduce the amount of ethanol absorbed or increase the rate of ethanol removal from the body (patent publication US2009-0060894A1). However, these methods have at least two potential pitfalls: (1) their methods of action effect the consumer's blood alcohol content itself, which may be undesirable; and (2) accelerated ethanol metabolism into acetaldehyde could increase the body's exposure to acetaldehyde and exacerbate symptoms and/or merely

induce hangover symptoms earlier rather than preventing them. In short, ethanol-centered strategies do not directly address the problem of acetaldehyde toxicity.

**[0006]** More directly, other groups have attempted to address acetaldehyde itself. US patent publication 2013-0089535A1 used enzyme preparations in the oral cavity. While there is evidence that alcohol consumption does raise the acetaldehyde concentration in saliva, it is unlikely that acetaldehyde removed from the oral cavity would have any significant effect on the systemic acetaldehyde toxicity responsible for an alcohol hangover.

**[0007]** Sprince, H., et al. (*Protective action of ascorbic acid and sulfur compounds against acetaldehyde toxicity: implications in alcoholism and smoking*. Agents Actions, 1975. 5(2): p. 164-73) refers to using small molecules to bind and remove acetaldehyde [1].

**[0008]** US patent publication 2015-0087702A1 refers to using small molecules increase the rate or expression of human enzymes to remove acetaldehyde.

**[0009]** Other groups have developed expression systems based on *Bacillus* or other bacterial flagellar regulatory regions (U.S. patent 7888064B2, European patent EP2235045B1), or by manipulating specifically *flgM* and *CsrA* (Japanese patent JP5881352B2).

**[0010]** Liu, Y. et al. ("Heterologous Expression of Aldehyde Dehydrogenase in *Lactococcus lactis* for Acetaldehyde Detoxification at Low pH", *Appl Biochem Biotechnol*. 8 August 2017) refers to application of *Lactococcus lactis* with the Nlsin Controlled Expression (NICE) System to express the aldehyde dehydrogenase gene (*istALDH*) in order to catalyze oxidation of acetaldehyde at low pH.

#### BRIEF DESCRIPTION OF THE DRAWINGS

**[0011]** The accompanying drawings, which are incorporated herein and form a part of the specification, illustrate exemplary embodiments and, together with the description, further serve to enable a person skilled in the pertinent art to make and use these embodiments and others that will be apparent to those skilled in the art. The invention will be more particularly described in conjunction with the following drawings wherein:

**[0012]** **FIG. 1** shows that induction of the aldehyde dehydrogenases AldB or AcoD results in rapid removal of acetaldehyde. Samples were incubated for 30 minutes with 200mM acetaldehyde at 37°C. Samples were then filtered, diluted 1:10, and frozen. Thawed samples were then run on HPLC and compared to a standard curve to calculate remaining acetaldehyde concentration, and those values are presented here. "Acetaldehyde 10mM" refers to 10mM acetaldehyde standard that was run at the start of the HPLC run, and then again at the end of the HPLC run to ensure no stochastic loss of acetaldehyde while on the machine.

**[0013]** FIGs. 2A and 2B show expression of GFP is 5-10 times stronger under control of the modified *hag* promoter than under the control of pHyspank promoter. (A) Frozen stocks of *Bacillus subtilis* strain PY79 (negative control; left plate at "2-o'clock" position), *B. subtilis* expressing GFP under pHyspank promoter with *lacI* repression removed (right plate at "11-o'clock" position), and *B. subtilis* expressing GFP under the modified *hag* promoter (left plate at "11-o'clock" position, and right plate at "2-o'clock" position) were patched on LB and grown overnight. The green fluorescence was visualized with a blue LED flashlight and orange filter. (B) Cultures of *B. subtilis* expressing GFP under pHyspank promoter with *lacI* repression removed (dark grey bottom line in chart; right plate at "11-o'clock" position in (A)), and *B. subtilis* expressing GFP under the modified *hag* promoter (light grey top line in graph; left plate at "11-o'clock" position in (A)) were grown from individual colonies in LB broth for 7 hours and fluorescence and OD<sub>600</sub> readings were taken every 90 minutes starting at 2.5 hours.

**[0014]** FIG. 3 shows expression of GFP, AcoD, and AldB by presence of protein in pelleted cells from cultures grown for 4.5 hours in LB, lysed and assessed for total protein by Coomassie staining on 8% SDS-PAGE gel. Lane 1 is the Bio-Rad Kaleidoscope protein standard ladder. Lanes 2 and 3 are uninduced and induced cultures (respectively) of *B. subtilis* expressing AldB under pHyspank promoter. Lanes 4 and 5 are uninduced and induced cultures (respectively) of *B. subtilis* expressing AcoD under pHyspank promoter. Lane 6 is a culture of *B. subtilis* with *flgM* deleted but *hag* still intact (negative control for heterologous protein expression). Lane 7 is a culture of *B. subtilis* expressing GFP (protein band of interest indicated by arrowhead) under the modified *hag* promoter. Lanes 8, 9, and 10 are cultures of *B. subtilis* expressing n-terminal 6x-histidine-tagged AcoD ("6x-histidine" disclosed as SEQ ID NO: 22), untagged AcoD, and n-terminal 6x-histidine-tagged AldB ("6x-histidine" disclosed as SEQ ID NO: 22) (respectively; protein band of interest indicated by arrowhead) under the modified *hag* promoter.

**[0015]** FIG. 4 shows a promoter structure for *hag* gene [SEQ ID NO: 7]. Identified in the figure are: the promoter start position (indicated by arrow and "-182"), the SigD RNA polymerase binding site (underlined and labelled "-10" and "-35"), the transcriptional start site (indicated by arrow and "+1"), CsrA binding sites 1 and 2 (bracketed and labeled "CsrA BS1/2"), the residue targeted for G-to-A point mutation to prevent CsrA binding as discussed in methods (circled "g"), and the Shine-Dalgarno sequence (boxed). ATG (capitalized in sequence) is the start codon for *hag* protein or any heterologous polypeptide.

**[0016]** FIG. 5 shows predicted secondary structure of *hag* mRNA [SEQ ID NO: 21], including 5'-UTR (fold prediction generated by University of Vienna's RNA fold software; citation: Gruber AR, Lorenz R, Bernhart SH, Neuböck R, Hofacker IL. The Vienna RNA Websuite. Nucleic Acids Res. 2008). (A) Folded structure. (B) Magnified view of the region of

the mRNA containing the two CsrA binding sites. The "AGGA" motif in the loop of the stem-loop secondary structure for both binding sites recognized by CsrA is outlined with thick black lines. The secondary structure is crucial, as it can clearly be seen that the two binding sites are adjacent. Thus, modifications that affect the secondary structure in a way that would prevent the proximity of these two sites could have deleterious effects on CsrA binding potentially equal to modifications to the binding sites or "AGGA" recognition sequences themselves.

**[0017]** FIG. 6 shows a strategy for de-repression of expression from the hag promoter. Knock-out ("X") of the FlgM gene prevents FlgM inhibition of SigD activity. SigD activates (bold arrow) transcription of the hag promoter and operatively linked ALDH gene. Mutation ("X") of the csrA binding site on the hag promoter reduces inhibition of ribosome binding and translation of the ALDH gene. This results in robust expression (bold arrow) of ALDH protein.

#### SUMMARY

**[0018]** In one aspect disclosed herein is a recombinant microorganism comprising: a) a polynucleotide comprising an expression construct comprising a flagellin gene promoter operatively linked with a heterologous nucleotide sequence encoding a subject polypeptide, wherein the flagellin gene promoter comprises one or more genetic modifications that reduce CsrA repression of translation of an mRNA transcribed from the flagellin gene promoter; and b) a genetic modification of a *flgM* gene that reduces inhibition of SigD initiation of transcription. In one embodiment, the recombinant microorganism constitutively expresses the polypeptide. In another embodiment, the microorganism is probiotic. In another embodiment, the microorganism belongs to genus selected from *Bacillus*, *Bifidobacterium*, *Enterococcus*, *Escherichia*, *Lactobacillus*, *Leuconostoc*, *Pediococcus*, *Pediococcus* and *Streptococcus*. In another embodiment, the microorganism belongs to genus *Bacillus*. In another embodiment, the microorganism is *Bacillus subtilis* (*B. subtilis*). In another embodiment, the flagellin gene promoter is a *hag* promoter. In another embodiment, the *hag* promoter comprises one or more genetic modifications that reduce CsrA repression of translation of an mRNA transcribed from the flagellin gene promoter, wherein the genetic modification comprises modification of the CsrA BS1 binding site and/or CsrA BS2 binding site (e.g., nucleotide substitution, insertion or deletion). In another embodiment, the one or more genetic modifications comprise one or a plurality (e.g., two, three or four) of genetic modifications to the CsrA BS1 recognition sequence, AGGA, e.g., to the sequence AGAA. In another embodiment, the one or more genetic modifications comprise one or more genetic modifications in the 12-base-pair BS1 binding site or in the surrounding bases on either side of the binding site that form the stem of the stem-loop secondary structure of BS1. In another

embodiment, the one or more genetic modifications comprise one or a plurality (e.g., at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12) of genetic modifications in the BS1 binding site, gcacaaggacgt (SEQ ID NO: 8). In another embodiment, the one or more genetic modifications disrupt the stem and loop structure of BS1 by eliminating complementarity that allows hydrogen bonding. In another embodiment, the one or more genetic modifications comprise one or a plurality of genetic modifications in the sequence taagggcacaaggacgtgcctta (SEQ ID NO: 1) that are involved in hydrogen bonding, for example, to eliminate one, two, three, four or more hydrogen bond pairs. In another embodiment, the modified BS1 has the nucleotide sequence GCACAAGAACGT (SEQ ID NO: 2). In another embodiment, the one or more genetic modifications comprise one or more point mutations to the CsrA BS2 binding site. In another embodiment, the one or more genetic modifications comprise one or a plurality (e.g., at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13) of genetic modification in the 13-base-pair BS2 binding site or in the surrounding bases on either side of the binding site that form the stem of the stem-loop secondary structure of BS2 ATTCAGGGAGGAA (SEQ ID NO: 9). In another embodiment, the one or more genetic modifications disrupt the stem and loop structure of BS2 by eliminating complementarity that allows hydrogen bonding. In another embodiment, the modified BS2 has the nucleotide sequence ATTTAGGGAGGAA (SEQ ID NO: 3). In another embodiment, the one or more genetic modifications to the BS2 binding site does not include an alteration of nucleotides in the Shine-Dalgarno sequence agggagga. In another embodiment, wherein the flagellin gene promoter is located in a bacterial chromosome or in a plasmid. In another embodiment, the subject polypeptide is an aldehyde dehydrogenase. In another embodiment, the aldehyde dehydrogenase is AcoD from *Cupriavidus necator* and comprises an amino acid sequence identical to or substantially identical to:

MNMAEIAQLGVSNPYKQQYENYIGGAWVPPAGGEYFESTTPITGKPFTRVPRSGQQQDVDA  
ALDAAHAAKAAWARTSTTERANILNRIADRIEANLKLVAESIDNGKPVRETTAADLPLAVD  
HFRYFAGCIRAQEGGISEIDADTIAYHFHEPLGVVGQIIPWNFLLMATWKLAPALAAGNCVV  
LKPAEQTPASILVLMEVIGDLLPPGVNVINGFGLEAGKPLASSPRISKVAFTGETTTGRLIM  
QYASQNLIPVTLELGKSPNIFEDVLAADDAFFDKALEGFAMFALNQGEVCTCPSRALIQE  
SIYDRFMERALKRVAAIRQGHPLDTGTMIGAQASAEQLEKILSYIDLGRKEGAQCLTGGERN  
VLDGDLAGGYVVKPTVFAGHNKMRIFQEEIFGPVVSVTTFKDEEEALAIANDTLYGLGAGV  
WTRDGARAFRMGRGIQAGRVWTNCYHAYPAHAAFGGYKQSGIGRENHRMMLDHYQQTK  
NLLVSYSPNALGFF [SEQ ID NO: 4]. In another embodiment, the aldehyde dehydrogenase is a human aldehyde dehydrogenase, e.g., having an amino acid sequence identical to or substantially identical to:

MLRAARFGPRLGRRLLSAAATQAVPAPNQQPEVFCNQIFINNEWHDAVSRKTFPTVNPS  
TGEVICQVAEGDKEDVDKAVKAARAAAFQLGSPWRRMDASHRGRLNRLADLIERDRTYLA  
ALETLDNGKPYVISYLVLDMLVKCLRYAGWADKYHGKTIPIDGFFSYTRHEPVGVC  
QIIPWNFPLLMQAWKLGPAATGNVVMKVAEQTPLTALYVANLIKEAGFPPGVVNIVPG  
FGPTAGAAIASHEDVDKVAFTGSTEIGRVIQVAAGSSNLKRVTELGKSPNIIMSDADM  
DWAVEQAHFALFFNQGCCAGSRTFVQEDIYDEFVERSVARAKSRVVGPNPFDKTEQG  
P

QVDETQFKKILGYINTGKQEGAKLLCGGGIAADRGYFIQPTVFGDVQDGMTIAKEEIFGP  
VMQILKFKTIEEVGRANNSTYGLAAVFTKDLKANYLSQALQAGTVVWNCYDVFGAQS  
PFGGYKMSGSGRELGEYGLQAYTEVKT VTKVPQKNS [SEQ ID NO: 5]. In another  
embodiment, the genetic modification in a *flgM* gene comprises deletion of all or part of the  
*flgM* gene. In another embodiment, the genetic modification in a *flgM* gene comprises a  
single mutation or series of mutations in the sequence encoding active sites of *flgM*. In  
another embodiment, the genetic modification in a *flgM* gene disrupts secondary or tertiary  
structure, such as in one of the helices that defines FlgM secondary structure. In another  
embodiment, the genetic modification in a *flgM* gene comprises altering an amino acid in the  
3<sup>rd</sup> or 4<sup>th</sup> helix at the C-terminal end of the FlgM protein, e.g., selected from I-58, K-62, I-65,  
G-68, D-73, A-78 of the *B. subtilis* FlgM. In another embodiment, the genetic modification in  
a *flgM* gene comprises altering one or more amino acids predicted to participate in FlgM  
binding to SigD, e.g., selected from I-3, G-7, S-10, V-11, A-40, K-41, M43, I-58, L-61, K-62,  
I-65, Y-70, K-71, V-72, D-73, A-74, H-76, I-77, A-78, N-80, M-81, I-82, N-83, F-84, Y-85, and  
K-86 of the *B. subtilis* FlgM.

**[0019]** In another aspect disclosed herein is a recombinant probiotic microorganism that  
constitutively expresses an aldehyde dehydrogenase. In one embodiment, the  
microorganism comprises: a) a polynucleotide comprising an expression construct  
comprising a flagellin gene promoter operatively linked with a heterologous nucleotide  
sequence encoding a subject polypeptide; and b) a genetic modification in a FlgM gene that  
reduces inhibition of SigD expression. In another embodiment, the flagellin gene promoter  
comprises one or more genetic modifications that reduce CsrA repression of translation of  
an mRNA transcribed from the promoter.

**[0020]** In another aspect disclosed herein is a polynucleotide comprising an expression  
construct comprising a flagellin gene promoter operatively linked with a heterologous  
nucleotide sequence encoding a subject polypeptide, wherein the flagellin gene promoter  
comprises one or more genetic modifications that reduce CsrA repression of translation of  
an mRNA transcribed from the flagellin gene promoter.

**[0021]** In another aspect disclosed herein is a method of making a polypeptide comprising culturing a recombinant microorganism as disclosed herein. In one embodiment, the method further comprises isolating the polypeptide.

**[0022]** In another aspect disclosed herein is a composition comprising a physiologically acceptable carrier and a recombinant probiotic microorganism, wherein the recombinant probiotic microorganism comprises: a) a polynucleotide comprising an expression construct comprising a flagellin gene promoter operatively linked with a heterologous nucleotide sequence encoding a subject polypeptide; and b) a genetic modification of a FlgM gene that reduces inhibition of SigD initiation of transcription. In one embodiment, the flagellin gene promoter comprises one or more genetic modifications that reduce CsrA repression of translation of an mRNA transcribed from the flagellin gene promoter. In another embodiment, the subject polypeptide is an aldehyde dehydrogenase. In another embodiment, the physiologically acceptable carrier is selected from lactic acid fermented foods, fermented dairy products, resistant starch, dietary fibers, carbohydrates, proteins and glycosylated proteins, water, capsule filler, and a gummy material.

**[0023]** In another aspect disclosed herein is a unit dose of a composition as disclosed herein, comprising about  $10^4$  to about  $10^{12}$  colony forming units of the recombinant probiotic microorganisms.

**[0024]** In another aspect disclosed herein is a method for preventing or treating an alcohol hangover, comprising administering to a subject in need thereof an effective amount of a composition comprising microorganisms that constitutively expresses aldehyde dehydrogenase. In another embodiment, the composition is administered before during or after the subject has consumed alcohol.

**[0025]** In another aspect disclosed herein is a method of metabolizing an analyte in a gut or in circulation in a subject comprising administering to the subject an effective amount of a composition comprising microorganisms that constitutively express an enzyme that metabolizes the analyte.

**[0026]** In another aspect disclosed herein is a method of producing a target compound comprising: a) contacting a culture comprising a recombinant microorganism expressing a subject polypeptide as disclosed herein with an analyte, wherein the subject polypeptide is an enzyme for which the analyte is a substrate, and b) culturing the microorganism, wherein the enzyme catalyzes the conversion of the analyte into the target compound. In one embodiment, the enzyme is selected from an amylase, a lipase and a protease.

## DETAILED DESCRIPTION

### I. Introduction

**[0027]** Disclosed herein is a microorganism, e.g., a bacterium, that has been genetically engineered to constitutively express an acetaldehyde dehydrogenase enzyme. The acetaldehyde dehydrogenase (ALDH) is expressed internally in the bacterial cell and gains access to its substrate, acetaldehyde, not by secretion of the enzyme, but by diffusion of the acetaldehyde into the bacteria. Acetaldehyde is a highly soluble molecule and can passively diffuse across cellular membranes. This internal location of the enzyme provides a functional advantage over secretion, as the interior of the bacterial cell is protected from the harsh and variable environment in the lumen of the gut, characterized by: low pH, hostile bacteria and eukaryotic cells that are looking to degrade free floating proteins for defense or nutritional purposes, high competition for enzymatic co-factors such as NAD, and extracellular proteases.

**[0028]** The flagellar regulatory machinery of *B. subtilis* was adapted to accomplish constitutive and robust expression of the ALDH. *B. subtilis* regulates motility by a sophisticated system involving several positive and negative regulators [2]. The basic strategy of this system is to remove negative regulators of the highly expressed flagellin gene, called *hag*.

**[0029]** The gene encoding the flagellar subunit of the *B. subtilis* flagellin is *hag*, and thus it is produced in hundreds of thousands of copies in a single bacterium in the right conditions, using a transcriptional promoter and a ribosome binding site that are both robust [3, 4].

**[0030]** Transcription is mediated by a sigma factor, SigD, which is repressed by the FlgM protein [5]. Deletion of *flgM* greatly enhances constitutive expression and activity of SigD, and consequently results in higher and more constitutive transcription of the flagellar operon and specifically the *hag* gene.

**[0031]** Translation of the *hag* gene is enabled by a highly robust ribosome binding site that is bound and repressed post-transcriptionally by a protein called CsrA [6]. However, a single point mutation in the CsrA binding site abrogates its binding and results in constitutive translation of the Hag protein [7].

**[0032]** Using the combination of deleting the *flgM* gene and making the single point mutation in the CsrA-binding site, one can achieve extremely high levels of Hag protein constitutively during the *B. subtilis* life cycle. Similarly, if the *hag* gene is replaced with a heterologous gene encoding a protein of interest, that gene can be transcribed and translated constitutively at high levels. Thus, by replacing the *hag* gene with a gene encoding an ALDH, and then deleting the *flgM* gene and making a single point mutation at

the CsrA binding site of the *hag* promoter, robust and constitutive expression of our ALDH in *B. subtilis* was achieved.

**[0033]** Accordingly, in certain embodiments provided herein are microorganisms comprising a point mutation in the binding site of CsrA combined with a *flgM* deletion. By making these two mutations, the utility of any SigD-based or flagellin promoter systems is vastly increased by removing repression and making expression constitutive. Without wishing to be limited by theory, it is believed that this is distinct from a simple deletion of CsrA because CsrA is a pluripotent regulator in many bacterial species [8], and its deletion could have many other potentially undesirable phenotypic effects on the cell. By making a point mutation only in the binding site, the mutation precludes CsrA repression of the *hag* promoter specifically and uniquely, rather than removing CsrA repression from any other targets it may have.

**[0034]** This strategy is useful for constitutive expression of any desired polypeptide. This includes other enzymes that metabolize target analytes in the gut of an individual, or in the production of target molecules in industrial processes.

## II. Definitions

**[0035]** As used herein, the term “probiotic microorganism” or “probiotic bacterium” refers to a microorganism or bacterium, which, when administered in an effective amount, confers a health or wellness benefit on a host.

**[0036]** As used herein, the term “recombinant microorganism” or “recombinant bacterium” refers to a microorganism or bacterium that comprises a polynucleotide comprising attached nucleotide sequences not normally attached to each other in nature.

**[0037]** Terms used to describe sequence relationships between two or more nucleotide sequences or amino acid sequences include “reference sequence,” “selected from,” “comparison window,” “identical,” “percentage of sequence identity,” “substantially identical,” “complementary,” and “substantially complementary.”

**[0038]** A “reference sequence” is a defined sequence used as a basis for a sequence comparison and may be a subset of a larger sequence, e.g., a complete cDNA, protein, or gene sequence.

**[0039]** Because two polynucleotides or polypeptides each may comprise (1) a sequence (i.e., only a portion of the complete polynucleotide or polypeptide sequence) that is similar between the two polynucleotides, or (2) a sequence that is divergent between the two polynucleotides, sequence comparisons between two (or more) polynucleotides or polypeptides are typically performed by comparing sequences of the two polynucleotides over a “comparison window” to identify and compare local regions of sequence similarity.

**[0040]** A “comparison window” refers to a conceptual segment of typically at least 12 consecutive nucleotides or 4 consecutive amino acid residues that is compared to a reference sequence. The comparison window frequently has a length of at least 15 or at least 25 nucleotides or at least 5 or at least 8 amino acids. The comparison window may comprise additions or deletions (i.e., gaps) of about 20 percent or less as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. Optimal alignment of sequences for aligning a comparison window may be conducted by computerized implementations of algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package Release 7.0, Genetics Computer Group, 575 Science Dr., Madison, WI) or by inspection, and the best alignment (i.e., resulting in the highest percentage of homology over the comparison window) generated by any of the various methods is selected.

**[0041]** A subject nucleotide sequence or amino acid sequence is “identical” to a reference sequence if the two sequences are the same when aligned for maximum correspondence over the length of the nucleotide or amino acid sequence.

**[0042]** The “percentage of sequence identity” between two sequences is calculated by comparing two optimally aligned sequences over a comparison window, determining the number of positions at which the identical nucleotide or amino acid occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison (i.e., the window size), and multiplying the result by 100 to yield the percentage of sequence identity.

**[0043]** Unless otherwise specified, the comparison window used to compare two sequences is the length of the shorter sequence.

**[0044]** Methods are described further in Natl. Acad. Sci. USA 85:2444; Higgins & Sharp (1988) Gene 73:237-244; Higgins & Sharp, CABIOS 5:151-153 (1989); Corpet et al. (1988) Nucleic Acids Research 16:10881-90; Huang et al. (1992) Computer Applications in the Biosciences 8:155-65; and Pearson et al. (1994) Methods in Molecular Biology 24:307-31. Alignment is also often performed by inspection and manual alignment.

**[0045]** A subject nucleotide sequence or amino acid sequence is “substantially identical” to a reference sequence if the subject amino acid sequence or nucleotide sequence has at least 70% sequence identity over a comparison window. Thus, sequences that have at least 80% sequence identity, at least 85% sequence identity, at least 90% sequence identity, at least 95% sequence identity, at least 98% sequence identity or at least 99% sequence identity with the reference sequence are also “substantially identical”. Two sequences that are identical to each other are, of course, also “substantially identical”.

**[0046]** In certain embodiments, the subject polynucleotide specifically hybridizes to a reference polynucleotide. “Hybridizing specifically to” or “specific hybridization” or

“selectively hybridize to”, refers to the binding, duplexing, or hybridizing of a nucleic acid molecule preferentially to a particular nucleotide sequence under stringent conditions when that sequence is present in a complex mixture (e.g., total cellular) DNA or RNA.

**[0047]** The term “stringent conditions” refers to conditions under which a probe will hybridize preferentially to its target subsequence, and to a lesser extent to, or not at all to, other sequences. “Stringent hybridization” and “stringent hybridization wash conditions” in the context of nucleic acid hybridization experiments such as Southern and northern hybridizations are sequence dependent, and are different under different environmental parameters. An extensive guide to the hybridization of nucleic acids is found in Tijssen (1993) *Laboratory Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes* part I chapter 2 “Overview of principles of hybridization and the strategy of nucleic acid probe assays”, Elsevier, New York.

**[0048]** Generally, highly stringent hybridization and wash conditions are selected to be about 5° C lower than the thermal melting point (T<sub>m</sub>) for the specific sequence at a defined ionic strength and pH. The T<sub>m</sub> is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Very stringent conditions are selected to be equal to the T<sub>m</sub> for a particular probe.

**[0049]** An example of stringent hybridization conditions for hybridization of complementary nucleic acids which have more than 100 complementary residues on a filter in a Southern or northern blot is 50% formalin with 1 mg of heparin at 42° C, with the hybridization being carried out overnight. An example of highly stringent wash conditions is 0.15 M NaCl at 72° C for about 15 minutes. An example of stringent wash conditions is a 0.2X SSC wash at 65° C for 15 minutes (see, Sambrook et al. for a description of SSC buffer). Often, a high stringency wash is preceded by a low stringency wash to remove background probe signal. An example medium stringency wash for a duplex of, e.g., more than 100 nucleotides, is 1x SSC at 45° C for 15 minutes. An example low stringency wash for a duplex of, e.g., more than 100 nucleotides, is 4-6x SSC at 40° C for 15 minutes. In general, a signal to noise ratio of 2x (or higher) than that observed for an unrelated probe in the particular hybridization assay indicates detection of a specific hybridization.

**[0050]** As used herein, the term “transcription regulatory sequence” refers to a first nucleotide sequence that regulates transcription of a second nucleotide sequence to which it is operatively linked.

**[0051]** A “promoter” is a transcription regulatory sequence at least sufficient to promote the transcription of a nucleotide sequence in DNA into an RNA transcript. A transcript transcribed from a promoter typically includes sequences from the promoter downstream of the transcription start site, as well as downstream sequences that, in the case of mRNA, encode an amino acid sequence. Promoters are the best-characterized transcriptional

regulatory sequences because of their predictable location immediately upstream of transcription start sites. Promoters include sequences that modulate the recognition, binding and transcription initiation activity of the RNA polymerase. These sequences can be cis acting or can be responsive to trans acting factors. Promoters, depending upon the nature of the regulation, can be constitutive or regulated. They are often described as having two separate segments: core and extended promoter regions.

**[0052]** The core promoter includes sequences that are sufficient for RNA polymerase recognition, binding and transcription initiation. The core promoter includes the transcriptional start site, an RNA polymerase binding site, and other general transcription binding sites and is where the pre-initiation complex forms and the general transcription machinery assembles. The pre-initiation complex is generally within 50 nucleotides (nt) of the transcription start site (TSS).

**[0053]** The core promoter also includes a sequence for a ribosome binding site, necessary for translation of an mRNA into a polypeptide.

**[0054]** The extended promoter region includes the so-called proximal promoter, which extends to about 250 nucleotides upstream of the transcriptional start site (i.e., -250 nt). It includes primary regulatory elements such as specific transcription factor binding sites. It has been found that many genes have transcription regulatory elements located further upstream. In particular, a fragment that includes most of the transcription regulatory elements of a gene can extend up to 700 nt or more up-stream of the transcription start site. (See, e.g., U.S. 2007-0161031.) In certain genes, transcription regulatory sequences have been found thousands of nucleotides upstream of the transcriptional start site.

**[0055]** As used herein, a nucleotide sequence is "operatively linked" with a transcription regulatory sequence when the transcription regulatory sequence functions in a cell to regulate transcription of the nucleotide sequence. This includes promoting transcription of the nucleotide sequence through an interaction between a polymerase and a promoter.

**[0056]** As used herein, a first nucleotide sequence is "heterologous" to a second nucleotide sequence if the first nucleotide sequence is not operatively linked with the second nucleotide sequence in nature. By extension, a polypeptide is "heterologous" to an expression control sequence if it is encoded by nucleotide sequence heterologous the promoter.

**[0057]** As used herein, the term "homolog" refers to any naturally occurring gene from another genus or species than the one defined, or a distinct gene in the same strain or species that encodes for a protein having nearly identical folding and function. Studies have shown that between microorganisms, proteins with at least 30% amino acid identity to the gene or protein discussed have such properties (citation: PMID 23352839). Furthermore, the term "homolog" extends to genes that encode proteins with less than 30% identity to the

gene or protein discussed but have been identified in a peer-reviewed scientific journal as a homolog of said gene or protein.

**[0058]** As used herein, the term “ortholog” refers to any homolog that occurs in another genus or species from the one discussed.

**[0059]** As used herein, the term “paralog” refers to any homolog that occurs in the same strain or species as the one discussed, often the result of gene duplication.

**[0060]** As used herein, the term “allelic variant” refers to a naturally occurring variation of a gene.

**[0061]** As used herein, the term “artificial variant” refers to a gene or protein comprising one or more genetic modifications to a naturally occurring gene or protein while retaining natural function.

**[0062]** As used herein, the term “mutation” refers to an alteration in a nucleotide sequence or amino acid sequence. A mutation can include a substitution of one or more nucleotides (a single nucleotide substitution is referred to as an “SNV” or “point mutation”), one or more nucleotide additions or one or more nucleotide deletions, as well as the changes in amino acid sequence, if any, resulting from these nucleotide alterations.

**[0063]** As used herein, a “derivative” of a bacterial strain is any genetically distinct version of that strain.

**[0064]** As used herein, the term “subject”, when used in reference to an individual, refers to an individual animal, e.g., a human.

**[0065]** As used herein, the terms “therapy,” “treatment,” “therapeutic intervention” and “amelioration” refer to any activity resulting in a reduction in the severity of symptoms. The terms “treat” and “prevent” are not intended to be absolute terms. Treatment and prevention can refer to any delay in onset, amelioration of symptoms, mitigation of severity of symptoms, improvement in patient survival, increase in survival time or rate, etc. The effect of treatment can be compared to an individual or pool of individuals not receiving the treatment, or to the same patient prior to treatment or at a different time during treatment. In some aspects, the severity of disease is reduced by at least 10%, as compared, e.g., to the individual before administration or to a control individual not undergoing treatment. In some aspects, the severity of the condition is reduced by at least 25%, 50%, 75%, 80%, or 90%, or in some cases, no longer detectable using standard diagnostic techniques.

**[0066]** As used herein, the term “probiotic composition” refers to a composition comprising probiotic microorganisms and a physiologically acceptable carrier. Typically, a probiotic composition confers a health or wellness benefit on the host subject to whom it is administered.

**[0067]** As used herein, the term “physiologically acceptable” refers to a carrier that is compatible with the other ingredients of a composition and can be safely administered to a

subject. Probiotic compositions and techniques for their preparation and use are known to those of skill in the art in light of the present disclosure. For a detailed listing of suitable pharmacological compositions and techniques for their administration one may refer to texts such as Remington's Pharmaceutical Sciences, 17th ed. 1985; Brunton et al., "Goodman and Gilman's The Pharmacological Basis of Therapeutics," McGraw-Hill, 2005; University of the Sciences in Philadelphia (eds.), "Remington: The Science and Practice of Pharmacy," Lippincott Williams & Wilkins, 2005; and University of the Sciences in Philadelphia (eds.), "Remington: The Principles of Pharmacy Practice," Lippincott Williams & Wilkins, 2008.

**[0068]** The probiotic composition may be a liquid formulation or a solid formulation. When the probiotic composition is a solid formulation it may be formulated as a tablet, a sucking tablet, a chewing tablet, a chewing gum, a capsule, a sachet, a powder, a granule, a coated particle, a coated tablet, an enterocoated tablet, an enterocoated capsule, a melting strip or a film. When the probiotic composition is a liquid formulation it may be formulated as an oral solution, a suspension, an emulsion or syrup. Said composition may further comprise a carrier material independently selected from, but not limited to, the group consisting of lactic acid fermented foods, fermented dairy products, resistant starch, dietary fibers, carbohydrates, proteins, and glycosylated proteins.

**[0069]** As used herein, the probiotic composition could be formulated as a food composition, a dietary supplement, a functional food, a medical food or a nutritional product as long as the required effect is achieved, e.g. treatment or prevention of an alcohol hangover. Said food composition may be chosen from the group consisting of beverages, yogurts, juices, ice creams, breads, biscuits, crackers, cereals, health bars, spreads, gummies and nutritional products. The food composition may further comprise a carrier material, wherein said carrier material is chosen from the group consisting of lactic acid fermented foods, fermented dairy products, resistant starch, dietary fibers, carbohydrates, proteins and glycosylated proteins.

**[0070]** The terms "effective amount" and "effective dose" refer to that amount of an agent effective to ameliorate a disorder or condition. For example, for the given parameter, an effective amount will show an increase or decrease of therapeutic effect at least any of 5%, 10%, 15%, 20%, 25%, 40%, 50%, 60%, 75%, 80%, 90%, or at least 100%. Efficacy can also be expressed as "-fold" increase or decrease. For example, an effective amount can have at least any of a 1.2-fold, 1.5-fold, 2-fold, 5-fold, or more effect over a control.

**[0071]** An effective amount of the agent is determined based on the intended goal. The term "unit dose" refers to a physically discrete unit suitable for use in a subject, each unit containing a predetermined quantity of the composition calculated to produce the desired response in association with its administration, i.e., the appropriate route and treatment regimen. The quantity to be administered, both according to number of treatments and unit

dose, depends on the subject to be treated, the state of the subject and the protection desired. Precise amounts of the composition also depend on the judgment of the practitioner and are peculiar to each individual.

**[0072]** The terms “dose” and “dosage” are used interchangeably herein. A dose refers to the amount of active ingredient given to an individual at each administration. For the present invention, the dose can refer to the number of bacteria, e.g., the amount of colony forming units (CFU). The dose will vary depending on a number of factors, including frequency of administration; size and tolerance of the individual; severity of the condition; risk of side effects; the route of administration. One of skill in the art will recognize that the dose can be modified depending on the above factors or based on therapeutic progress. The term “dosage form” refers to the particular format of the pharmaceutical, and depends on the route of administration. For example, a dosage form can be in a liquid, e.g., a beverage for oral consumption.

### **III. Expression Constructs**

**[0073]** Contemplated here are methods that reduce either or both of inhibition of transcription from the flagellin promoter and inhibition of translation of an mRNA transcript from a flagellin promoter.

**[0074]** Certain polynucleotides disclosed herein comprise an expression construct comprising a flagellin gene transcription regulatory sequence, e.g., a promoter, operatively linked with a heterologous nucleotide sequence encoding a subject polypeptide, that is, a polypeptide subject to expression. In certain embodiments, the *hag* promoter comprises genetic alterations such that, upon transcription of an mRNA from the *hag* promoter, CsrA inhibition of mRNA translation is repressed.

#### **A. Flagellin Gene Transcription Regulatory Sequences**

**[0075]** Many bacteria have a flagellin gene homolog. The gene goes by many names, some examples of which are: *hag* in *B. subtilis*; *fliC* in *Escherichia coli*, *Bacillus thuringiensis*, and several *Lactobacillus* species; and *flaA* or *flaB/C/D/E/F/etc.* in *Legionella* species, *Vibrio* species, and *Campylobacter* species.

**[0076]** In one embodiment, the flagellin gene promoter is a *hag* gene promoter. The flagellin homolog is *hag* in *Bacillus*, e.g., *B. subtilis*. In another embodiment, the flagellin gene promoter is native to the microorganism in which the subject protein is to be expressed. For example, this can be the case when the expression construct is located in a bacterial chromosome.

**[0077]** As used herein, the term “Hag” (or “*hag*” or “*hag*”) refers to the protein (or gene encoding such protein) annotated as “Hag” in *B. subtilis* or any homolog in the same or other

genus, species, or strain, which is the structural subunit also known more generically as “flagellin” used to assemble a flagellum. It is known by several other names in other genera, species, and strains. It is defined in *B. subtilis* by the sequence:

atgagaattaaccacaatattgcagcgcttaacacactgaaccgtttgtcttcaacaacagtgcgagcca  
aaagaacatggagaaactttcttcagggtcttcgcatcaaccgtgcgggagatgacgcagcagggtcttgcca  
tctctgaaaaaatgagaggacaaatcagagggtcttgaaatggcttctaaaaactctcaagacggaatctct  
cttatccaaacagctgaggggtgcattaactgaaactcatgcgacatcctcaacgtgttcgtgagctagtgttca  
agctggaacactggaactcaggacaaagcaactgatttgcaatctattcaagatgaaatttcagctttaac  
agatgaaatcgatgggtatttcaaactgtacagaattcaatggtaagaaattgctcgatggcacttacaagtt  
gacacagctactcctgcaaatcaaaagaacttggattccaaatcggagcaaatgctacacagcaaatctc  
tgtaaatattgaggatatgggtgctgacgctcttggaattaaagaagctgatggttcaattgcagctcttcattc  
agttaatgatcttgacgtacaaaaattcgcagataatgcagcagatactgctgatatcggtttcgatgctcaa  
ttgaaagttgttgatgaagcgatcaaccaagtttcttctcaacgtgctaagcttgggtcggtacaaaatcgctc  
agagcacacaattaacaacttaagcgcttctgggtgaaaacttgacagctgctgagctcgtatccgtgacggt  
gacatggctaaagagatgagcgaattcacaagaacaacattcttctcaggcttctcaagctatgcttgctc  
aagcaaaccaacagccgcaaacgtacttcaattattacgttaa [SEQ ID NO: 6].

**[0078]** As used herein, the term “*hag* promoter” refers to a naturally occurring flagellin gene promoter cognate from genus *Bacillus* and promoters having sequences substantially identical thereto or hybridizing specifically thereto. In *B. subtilis*, the *hag* promoter is comprised in a 273 base-pair sequence 5' of the start codon of the *hag* gene. It has the nucleotide sequence:

ggaattgacgccccaaagcatattgatattcacaggaaagaaatttacttgaccattcaggaagaaaataa  
ccgtgcagcagcggtatccagcgatgtgatctccgcattatcctcacaataaaagtgaggattttttattttg  
**tattaa**caaaaatcagagacaat**ccgatat**aatgatgtagccgggaggaggcgcaaaagactcagc  
**cagttacaaaataagggcacaaggacgtgccttaacaacatattcagggagggaacaaaaca**ATG  
[SEQ ID NO: 7] (where “ATG” represents the start codon of *hag*). It is expected that the sequence in bold beginning with TTAA (underlined) through the start codon ATG suffices to promote gene expression.

**[0079]** In particular, the *hag* promoter contains a SigD recognition sequence defined by a “ttaa” sequence (underlined), which is the -35 SigD RNA polymerase binding site and a “tccgatat” sequence (underlined), which is the -10 SigD RNA polymerase binding site. In addition, *hag* has two CsrA binding sites defined by the sequences “gcacaaggacgt” [SEQ ID NO: 8] (high-affinity binding site 1, or “BS1”) (underlined) and “attcagggaggaa” [SEQ ID NO: 9] (low-affinity binding site 2, or “BS2”) (underlined). The *hag* promoter is also defined by a Shine-Dalgarno sequence “agggagga”. The overall structure of the *hag* promoter is illustrated in FIG. 4.

**[0080]** As used herein, the term “CsrA” (“Carbon storage regulator A”) refers to the protein (or gene encoding such protein) annotated as “CsrA” in *B. subtilis* -- or any homolog or ortholog in another genus or species, or paralog in the same species. CsrA is homologously referred to as RsmA in some species. CsrA protein binds to a stem-loop RNA motif having the consensus sequence AGGA in the loop, thereby inhibiting translation into polypeptide of a nucleotide sequence incorporated in an mRNA comprising the consensus sequence. CsrA can inhibit expression of an mRNA transcribed from the *hag* promoter either directly by binding to the RNA and preventing translation or indirectly by binding to another RNA that encodes a protein that otherwise regulates flagellar expression. CsrA is defined in *B. subtilis* by the sequence:

atgctagttttatcgcggaataaaacgaagcgattcaaatagggtgctgatattgaagtaaaagtgattgcg  
gttgaaggggatcaagtgaagcttggaattgacgccccaaagcatattgatattcacaggaaagaaattta  
cttgaccattcaggaagaaaataaccgtgcagcagcggttatccagcgatgtgatctccgcattatcctcaca  
aaaaaagtga [SEQ ID NO: 10].

**[0081]** In certain embodiments, expression constructs of this disclosure comprise genetic modifications in a flagellin gene promoter that, upon transcription from the promoter into a transcript, such as mRNA, repress CsrA inhibition of mRNA translation. This disclosure contemplates several genetic modifications to a flagellin gene promoter, and, in particular, to the *hag* promoter, to achieve this result. In some embodiments, genetic modifications to *hag* to inhibit CsrA repression of translation can comprise an alteration of a stem and loop structure in either or both of BS1 or BS2. In some embodiments, the genetic modification is an insertion or a deletion of one or more nucleotides.

**[0082]** The genetic modification can be one or more point mutations to binding to the CsrA BS1 binding site. BS1 can be modified by altering one or a plurality (e.g., two, three or four) nucleotides in the CsrA recognition sequence, AGGA. For example, the AGGA binding motif of BS1 can be modified to AGAA. Alternatively, the genetic modification can comprise one or more mutations in the 12-base-pair BS1 binding site or in the surrounding bases on either side of the binding site that form the stem of the stem-loop secondary structure of BS1. This includes, for example, modification of one or a plurality (e.g., at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12) of nucleotides in the BS1 binding site, gcacaaggacgt (SEQ ID NO: 8). Alternatively, a genetic alteration can disrupt the stem and loop structure of BS1 by eliminating complementarity that allows hydrogen bonding. Such alterations can be made as one or a plurality of mutations in the sequence taagggcacaaggacgtgcctta [SEQ ID NO: 1] involved in hydrogen bonding, for example, to eliminate one, two, three, four or more hydrogen bond pairs. In one embodiment, the modified BS1 has the nucleotide sequence GCACAAGAACGT [SEQ ID NO: 2].

**[0083]** The genetic modification can be one or more point mutations to binding to the CsrA BS2 binding site. This includes one or a plurality (e.g., at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13) of mutations in the 13-base-pair BS2 binding site or in the surrounding bases on either side of the binding site that form the stem of the stem-loop secondary structure of BS2. This also includes a genetic alteration that disrupts the stem and loop structure of BS2 by eliminating complementarity that allows hydrogen bonding. For example, the modified BS2 can have the nucleotide sequence ATTTAGGGAGGAA [SEQ ID NO: 3]. In certain embodiments, the modification does not include an alteration of nucleotides in the Shine-Dalgarno sequence agggagga.

**[0084]** It will be recognized that the genetic modification, while inhibiting CsrA binding, is selected to allow the mRNA to retain ribosome binding activity and to permit translation.

**[0085]** The heterologous nucleotide sequence is heterologous to the flagellin promoter. That is, it is a sequence not normally under the control of the *hag* promoter in nature and, typically, encodes a polypeptide whose expression is not under the control of the flagellin promoter. The subject polypeptide can be any polypeptide desired to be expressed. In certain embodiments, the polypeptide is an aldehyde dehydrogenase.

#### **B. Subject Polypeptide**

**[0086]** The heterologous nucleotide sequence can be placed under operative linkage with the flagellin promoter by any method known in the art. For example, the heterologous nucleotide sequence can be integrated into the bacterial chromosome. Alternatively, the heterologous nucleotide sequence can be attached to a flagellin promoter in a plasmid that is introduced into the microorganism. A heterologous nucleotide sequence can be targeted to the *hag* promoter by, for example, homologous recombination, as described, for example, in PMID: 4994568. Another useful method involves transposon technology. Transposons can target specific sequences in a chromosome and insert and attached a nucleotide sequence there. Materials for transposon integration are available from, e.g., Lucigen (Middleton, WI), which commercializes the EZ-Tn5 Transposase.

**[0087]** The polypeptide to be the subject of expression ("subject polypeptide") can be any polypeptide desired. Such polypeptides can be those desired to be expressed constitutively for the purpose of obtaining high amounts of the polypeptide. The polypeptide can be a biologic drug, a food protein, such as albumin, or an enzyme in a biochemical pathway, such as alpha-galactosidase A. In certain embodiments, the polypeptide is an enzyme that metabolizes analytes likely to be found in the gut of an animal. Such analytes can be found there through ingestion, through production by the microbiome, through production by the subject, or through diffusion or active transport across the membranes of the gut. Such

polypeptides include for example: peptidases like dipeptidyl peptidase IV; and lipases like human pancreatic lipase.

**[0088]** In one embodiment, the polypeptide is an aldehyde dehydrogenase. As used herein, the term “aldehyde dehydrogenase” or “ALDH” refers to any enzyme that is known or predicted to use acetaldehyde as a substrate and catalyze it into a distinct product. Most commonly this is an enzyme whose function is known or predicted to catalyze the oxidation of aldehydes. Generally, this refers to enzymes of the classification EC 1.2.1.-, which catalyze the oxidation of an aldehyde to a carboxylic acid using NAD(P) as a co-factor. It includes enzymes which oxidize acetaldehyde to acetate/acetic acid (e.g. EC 1.2.1.3, 1.2.1.4, and 1.2.1.5), and acetylating enzymes which convert coenzyme A to acetyl-coenzyme A (e.g. EC 1.2.1.10). It also refers to enzymes which use other defined acceptors, including (but not limited to): enzymes using a quinone (e.g. of the classification EC 1.2.5.-), or other unknown acceptors (e.g. of the classification EC 1.2.99.-).

**[0089]** In certain embodiments, the aldehyde dehydrogenase is a human aldehyde dehydrogenase (e.g., ALDH2) or a non-human aldehyde dehydrogenase (e.g., AcoD or AldB). The aldehyde dehydrogenase can be a bacterial aldehyde dehydrogenase and the microorganism in which it is expressed is a bacterium. The aldehyde dehydrogenase can be native to the microorganism in which the gene is to be expressed.

**[0090]** The subject polypeptide is expressed by placing a nucleotide sequence encoding the polypeptide under transcriptional regulatory control of the flagellin promoters described herein. Nucleotide sequences encoding many genes can be found in publicly available databases such as NCBI, uniprot, KEGG, BRENDA, etc. Polynucleotides encoding the amino acid sequences of these polypeptides can be linked to the promoter by methods well known in the art of molecular biology. In one embodiment, the subject polypeptide is not lacZ.

#### **IV. Disruption of Repressor of Sigma Factors**

**[0091]** Sigma factors, such as SigD and its homologs, initiate flagellin synthesis. FlgM and its homologs function as repressors of Sigma factor activity. This disclosure provides for de-repression of Sigma factor activity by disruption of Sigma factor repressors such as FlgM.

**[0092]** As used herein, the term “FlgM” refers to the protein (or the gene encoding such protein) annotated as “FlgM” in *B. subtilis*, or any homolog in another genus or species, which inhibits the sigma factor responsible for recruiting RNA polymerase to late flagellar genes for transcription. This inhibited sigma factor is called SigD in *B. subtilis*, FlhA in *E. coli*, or potentially other names such as sigma 28 in other genera and species in which said sigma factor has a homolog. FlgM is defined in *B. subtilis* by the sequence: atgaaaatcaatcaatttgaacacaatccgttaatccatatcaaaaaattatgataagcaagcgggtgcaaaaaactgttgacaaacctcaagataaaattgaaatttcacacaggctaagaaatgcaacatgcatccg

acgcagtcactggttcacgacaggaaaaattgcgagcttaagcgcaaattgaaaacgggtcataca  
aagtagacgcaaatacatattgcgaaaaatatgattaattttataaaaagcaataa [SEQ ID NO: 11].

**[0093]** As used herein, the term “SigD” refers to the sigma factor (or the gene encoding it) in *B. subtilis* responsible for, among other things, recruiting the RNA polymerase to late flagellar genes for transcription. “SigD” furthermore refers to homologs in other species, such as FliA in *E. coli*, or the broader denotation of sigma-28 in several species. SigD is defined in *B. subtilis* by the sequence:  
atgcaatccttgaattatgaagatcaggtgcttggacgcgctggaaagagtggaaagatcctaagccgg  
tgacgacttaatgcgcggttacatgccgctgtcacatatcatgtaggcagaatttctgctggactgccgaaat  
cagtgcataaagacgatcttatgagccttggtatgcttggtttatatgatgcccttgaaaaatttgaccccagc  
cgggacttaaaatttgatactacgcctcgtttagaattcgcggcgcaatcatagacgggcttcgtaaagaa  
gattggctgcccagaacctgcgcgaaaaaacaataaagggtgaagcagcaattgaaaagcttgaaca  
gcggtatcttcggaatgtatcgccgcggaattgcagaggaactcggaatgacggtacaggatgtcgtgt  
caacaatgaatgaagggtttttgcaaatctgctgtcaattgatgaaaagctccatgatcaagatgacggg  
aaaacattcaagtcgatgatcagagatgacaaaaatgttcgcctgaagaaaagattatgaaggatgaact  
gattgcacagcttgcggaaaaaattcacgaactctctgaaaaaacagctggttgcagttgttctacaa  
agaggagttgacactgacagaaatcggaacaagtattaaatcttctacgtcccgcataatctcagatccattca  
aaggcatttttaataaagaatctgctggaaaaagtgatacaataa [SEQ ID NO: 12].

**[0094]** FlgM binds to SigD via several residues, the majority of which are located in the 4th helix at the C-terminal end of the FlgM protein. Targets for inactivation would be mutation of the highly conserved residues in the 3rd and 4th helices corresponding to I-58, K-62, I-65, G-68, D-73, A-78 of the *B. subtilis* FlgM. More broadly, any one or combination of the 26 residues directly involved in binding to SigD (as identified in PMID: 15068809) could be mutated to potentially generate a protein with reduced or null activity. Alternatively, any mutation or combination of mutations that disrupted the secondary or tertiary structure -- in particular the 4 helices that define the secondary structure -- could potentially successfully reduce or attenuate the ability of FlgM to inhibit SigD.

**[0095]** In some embodiments, the recombinant microorganism comprises a genetic modification that reduces the ability FlgM to inhibit SigD activity. In particular, one such genetic modification is partial or complete deletion of the FlgM gene to reduce or eliminate its biological activity. Partial deletion can include deletion of part of the gene encoding at least 25% of the C-terminus of the protein.

**[0096]** Other genetic modifications include, for example, frameshift mutations producing an inactive FlgM protein, or disruption of the FlgM promoter. Alternatively, FlgM could be rendered inactive by a point mutation that renders it functionally inactive or otherwise inhibits its ability to bind to or otherwise repress SigD. Indirectly, FlgM could be disrupted by making

it insensitive to activators such as ComK or DegU, or by making it overly sensitive to repressors such as ScoC or proteases Epr and Wpr.

#### V. Recombinant Microorganisms

**[0097]** Recombinant microorganisms contemplated for use herein comprise an expression system including at least an expression construct comprising a flagellin gene promoter operatively linked with a heterologous nucleotide sequence encoding a subject polynucleotide for expression, which flagellin gene promoter is genetically modified to reduce CsrA inhibition of translation; and, optionally, a FlgM gene genetically modified to reduce suppression of SigD expression.

**[0098]** The recombinant microorganism is typically a probiotic microorganism that is safe for ingestion by an animal, e.g., a human. Probiotic microorganisms include, without limitation, from *Bacillus*, *Bifidobacterium*, *Enterococcus*, *Escherichia coli*, *Lactobacillus*, *Leuconostoc*, *Pediococcus* and *Streptococcus*. In particular, this disclosure contemplates the use of bacteria of genus *Bacillus* and more particularly, *B. subtilis*. Microorganisms can be obtained from any typical source including, for example, the American Type Culture Collection (ATCC) or the Bacillus Genetic Stock Center (BGSC).

**[0099]** In certain embodiments, the microorganism is not a eukaryote. For example, the microorganism is not a eukaryote used for fermentation of alcoholic beverages, such as *Saccharomyces*.

#### VI. Probiotic Compositions

**[0100]** Also provided herein are compositions comprising probiotic microorganisms. The compositions comprise recombinant microorganisms as provided herein in a physiologically acceptable carrier.

**[0101]** In certain embodiments, the probiotic compositions are provided in a unit dose. A unit dose will typically have between  $10^4$  to  $10^{12}$  colony forming units of the microorganism in a suitable carrier. This includes any effective dose as described herein. The total volume of the dose can be between about 0.05 mL and 500 mL, e.g., between about 0.5 mL and about 50 mL or between about 2.5 mL and 15 mL.

#### VII. Methods of Use

**[0102]** Also provided herein are methods of using recombinant microorganisms and probiotic compositions to metabolize target analytes and/or to provide health or wellness benefits on the host.

**[0103]** The host can be any animal. This includes, for example, mammals, e.g., primates, e.g., humans. Mammals can belong to any class, such as, artiodactyla, carnivora,

cetacean, chiroptera, dermoptera, edentata, hyracoidae, hyracoidae, insectivore, lagomorpha, marsupialia, monotremata, perissodactyla, ungulate, pholidata and pinnipedia.

**[0104]** A host can ingest a probiotic composition of this disclosure in an amount effective to produce the intended result. The intended result can be metabolism of a target analyte in the gut of the host. In this case, the subject polypeptide can be an enzyme that processes the analyte. In particular, this disclosure contemplates metabolism of acetaldehyde, a product of ethanol metabolism, which is a cause of alcohol hangover.

**[0105]** As used herein, the term “alcohol hangover” refers to a condition resulting from consumption of alcohol and characterized by one or more of: dehydration, poor sleep, grogginess, nausea, vomiting, headache, malaise, dry mouth, sensitivity to light and sound, sweating, muscle pain, diarrhea, stomach pain, gastrointestinal distress, vertigo, anxiety, depression, irritability, and elevated acetaldehyde concentration in the body.

**[0106]** Prevention or treatment, e.g., mitigation, of alcohol hangover can involve administering a probiotic composition of this disclosure to a subject before, during or after consumption of alcohol. The composition can be ingested up to any of twenty-four hours, eighteen hours, twelve hours, eleven hours, ten hours, nine hours, eight hours, seven hours, six hours, five hours, four hours, three hours, two hours or one hour before alcohol consumption begins. The probiotic composition can be administered during alcohol consumption. The probiotic composition can be administered up to any of twelve hours, eleven hours, ten hours, nine hours, eight hours, seven hours, six hours, five hours, four hours, three hours, two hours or one hour after alcohol consumption or at any time a subject suffers from symptoms of alcohol hangover.

**[0107]** The composition can be administered to the gut (e.g., stomach or intestines) of the individual, either by oral consumption or through the anus. Living, rather than non-living microorganisms can be administered to the individual. The microorganisms can be provided in a different composition than the alcoholic beverage, that is, can be taken separately from the alcoholic beverage. This can be the case, for example, when the composition is administered before consumption of alcohol.

**[0108]** The effective dose of the probiotic can depend on the extent of symptoms of alcohol hangover. Generally, the dosage of recombinant bacteria will vary depending upon such factors as the subject's age, weight, height, sex, general medical condition and previous medical history. In specific embodiments, it may be desirable to administer the bacterium in the range of about  $10^4$  to about  $10^{12}$  CFU,  $10^5$  to  $10^{11}$  CFU,  $10^6$  to  $10^{10}$  CFU,  $10^8$  to  $10^{10}$  CFU or  $10^8$  to  $10^{12}$  CFU (“colony forming units”) of the microorganism.

**[0109]** In another embodiment, microorganisms disclosed herein are used for the production of a desired composition. Methods to produce such compositions can involve culturing a recombinant microorganism that constitutively expresses an enzyme that

transforms a target analyte into the desired composition or an intermediate in the production of the desired composition. In this case, the target analyte is added to the microorganism culture and incubated for a time sufficient to produce the enzymatic product.

## EXAMPLES

### EXAMPLE 1

#### Methods:

[0110] Strains, plasmids, and media

[0111] The *B. subtilis* strain PY79 from the Bacillus Genetic Stock Center was used (strain 1A747) for all manipulations. All *B. subtilis* strains used here are derivatives of PY79.

[0112] Unless otherwise noted, bacteria were grown in LB medium (1% tryptone, 0.5% yeast extract, 0.5% sodium chloride, [1.5% agar if solid medium]). For MLS resistance selection, 1ug/mL erythromycin and 25ug/mL lincomycin were used. For transformation experiments, bacteria were grown in modified competence (MC) medium (100mM phosphate buffer, 2% glucose, 3mM trisodium citrate, 22mg/L ferric ammonium citrate, 0.1% casein hydrolysate, 0.16% glutamic acid, 3mM magnesium sulfate).

[0113] Plasmid used to make genetic modifications is pMiniMAD [9].

[0114] Plasmid used for IPTG-induced expression under pHyspank promoter is pDR111 [10]. 1mM IPTG was used for all inductions.

[0115] flgM deletion

[0116] An *in vitro* transcribed double-stranded DNA molecule containing ~800 base pairs 5' and 3' of the *flgM* gene from *B. subtilis* PY79, along with the first and last 15 base pairs of the coding sequence of *flgM* separated by a PstI restriction enzyme recognition sequence (see full sequence in table 1) was constructed. pMiniMAD was then linearized using primers ZP24 and ZP25 (table 1). The resulting PCR product and the delta *flgM* sequence were then ligated together to create pMiniMAD delta-*flgM* using Gibson assembly [11]. The resulting plasmid was transformed into *E. coli* via heat transformation. Plasmid was miniprepmed using a commercially available kit, and that miniprepmed plasmid was used as the DNA source to do the deletion as follows:

- i. Single colony of PY79 inoculated in 2mL of MC medium in 15mL test tube
- ii. Grow culture at 37 °C with shaking at 275rpm for 4.5 hours, or ~1 hour after exiting exponential growth
- iii. Transfer 400uL of culture to new 15mL test tube and add 1ug of miniprepmed plasmid (described above)
- iv. Return culture with DNA to shaker at 37 °C for 1.5 hours
- v. Plate culture on LB agar with 1ug/mL erythromycin and 25ug/mL lincomycin

- vi. Incubate overnight at 37 °C
- vii. Screen isolated colonies for mutant allele via PCR using primers ZP77 and ZP78, and inoculate positive colony in 3mL LB broth without antibiotics
- viii. Grow overnight at room temperature with shaking at 275rpm
- ix. Isolation streak 10uL of overnight culture on LB without antibiotics and grow at 37 °C overnight. Because of the lack of antibiotic selection, plasmid can be lost during overnight replication. The observed stability of this plasmid is about 90%, with about 1 colony in 10 losing the plasmid at this stage.
- x. Duplicate streak isolated colonies on LB with and without antibiotics (1ug/mL erythromycin and 25ug/mL lincomycin), and grow overnight at 37 °C
- xi. PCR screen antibiotic sensitive colonies with primers ZP77 and ZP78 to identify strains with the mutant allele

**[0117] Table 1: Primers and constructed genes/alleles**

| Primer | Description                   | Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|--------|-------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| ZP24   | Fwd<br>pMiniMAD for<br>gibson | CTGGCGTTACCCAACTTAATC [SEQ ID NO: 13]                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| ZP25   | Rev<br>pMiniMAD for<br>gibson | CTTGGCGTAATCATGGTCATAG [SEQ ID NO: 14]                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| ZP77   | flgM fwd                      | GAGGAAACAGGTGTGGAAGAAG [SEQ ID NO: 15]                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| ZP78   | flgM rev                      | GGTCATCTTCTGTCTGCGTG [SEQ ID NO: 16]                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|        | delta flgM<br>allele          | CTATGACCATGATTACGCCAAGTGAATAATGAGAAACAGTCAAA<br>GAAAAAGAAAACAGAACGCCTGCTGTCAGAGTGCATTTTTGATA<br>CAAAAATAATTTCAGCAGAAGGTATGAATATCATTTTAATAGAC<br>GATCTTTATACAACAGGCGCCACCTTGCACTTCGCAGCCCGCT<br>GCTTATTAGAAAAAGGAAAAGCCGCTTCAGTGTCTATCTTTTACC<br>TTGATCAGAAGCTAAATGATTCTGTTTTTATGCCGATATAATCAC<br>TAGAAATTGACACAGGCATATTATCTAATAAGGAGAAAAAAGA<br>TGGGAGAACTGGCTAATTGTCCGAAATGCAATGCTTTATTTTAA<br>AAAACAAAGCTGCAAACCGTATGTCAGGCGTGTATTAAGGAAG<br>AAGAAAAATCATTTGAGACTGTCTATAAATTTTAAAGAAAACAGG<br>AAAACCGGCAATCAACTTTGAGCCGGATAACTGAGGAAACAGG<br>TGTGGAAGAAGAGCTGATATTGAAATTCATCAGGCAGAAGCGA<br>ATTCAGATCACTCATCTTCCTAATTTGGCATAACCCTTGTGAAAG<br>GTGCGGGACATCGATTAGAGAAGGCAAGTTCTGCAAGGCTTGC<br>CAGTCTGATATTAAGGATCAAATGGATCATTTGAACACGAGGA<br>TGCTCTGAAAAATCGAGAAAGAAAATAGTAAAAAGACACATACT<br>ATGCCTATAATACCAAAAAACAGCTGATTCCCTAACTAACTGAA<br>AACGCAGTCGATAAAAGGGTTAAGATTGTTTAAAGACTGCAACG<br>GAAAGCGAGAGGAATCCTATGAAAATCAATCAACTGCAGTTTAA |

|  |            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|--|------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |            | <p>TAAAAAGCAATAAAAAAGGAGAAAGCCCATGTCAGCGAAGGCA<br/> ATTATTGAACAATTGAAGCGACTTTGCGTTCTGCATGAGCACCT<br/> GCTCACGCTGTCTGAAGAAAAGACGGAAGCGCTCAAAGCCGG<br/> CAAAACAAAAGAGCTTTCTAACATTTTGACAAAAGAGCAAAAAT<br/> ATATTCAAGCAATCACGCAGACAGAAGATGACCGGATCAAAAC<br/> AACTTCGGCCTTTCTCGGATATAGCGAAAAATAACTATTTCCG<br/> CATGTATCGCCAAAACCTCAGGCAGTGAAGGAAGAGCTGGA<br/> ACAACATAACGAATCTCTTTCTCAAGTTCTCGGACGCTGAAAA<br/> AAGTAAATGAGATGAATAGGCAGCTGACAAGAGACGCGCTGCA<br/> ATTCATCTCTATTTTCGTACGATATGCTGGTTCCTAAGGAAAAATA<br/> CTTCAATTACAGCAAAATCAATTAAGCTGAGCTGCCGAAAAGTA<br/> GCAAAATGAACTGTTTGATTCAAAGCTTAGCAGAAAAGGAATT<br/> CAGAAAATGACATCTACCTTTATGGGGCTTGAACTGCAAGGC<br/> GGGCGTTAAGCGCTCAGCAGGCAGCGTTAAGCACTACTGCAAA<br/> TAACGTGGCAAATGCCAATACTGATGGTTATACAAGACAGCGG<br/> GTCTCATTGGAGGCAACTGACTATTTCCCTGCTGTATCTAAAAA<br/> TGCAGAAAAAACAGCGGGACAAATGGGTACGGGCGTTCAAGG<br/> AAAATCAGTTGAGAGAATAAGAGATATCTTTCTTGACTACCAAT<br/> ACCGTCTTCAAAACCTGGCGTTACCCAACTTAATC [SEQ ID NO:<br/> 17]</p>                                                                                                                                                                                                                                                                                                                                                                                                               |
|  | phag + GFP | <p>CTATGACCATGATTACGCCAAGTGAACAATGATCATTTCATACGA<br/> AGTACCATGGCCAAATGAACATAAAAGAAGAACAAATCATTCTT<br/> TTTGAAAGCGGGATTCCAGGCTTTTTAGAAAGAAAACAGTTCGT<br/> CATACTTCCGCTTTCAGAAGACTCTCCATTCGTGGCACTGCAGT<br/> CCGTCACTTCAGAAAATCTTGCGTTTATCGTCGTAAGTCCGTTT<br/> ATCTTTTTTAAGAATTATGAATTTGATCTTGATGAATCAACTGCT<br/> GAACTTTTGGATATCGATAATATTCAAGACGTAGAAGTCATGAC<br/> AATATTGACTATGGCAGAGCCATTTGAAAAGTCTACTGCGAATT<br/> TATTGGCTCCCATTATTGTGAATCGCAAGAACATGATGGCTAAG<br/> CAAGTCGTTTTACACGACTCCTCATATACGACAAAGCATCCGAT<br/> TGGAGGAGAATCATGCTAGTTTTATCGCGGAAAATAAACGAAG<br/> CGATTCAAATAGGTGCTGATATTGAAGTAAAAGTGATTGCGGTT<br/> GAAGGGGATCAAGTGAAGCTTGGAATTGACGCCCCAAAGCATA<br/> TTGATATTACAGGAAAGAAATTTACTTGACCATTACAGGAAGAA<br/> AATAACCGTGCAGCAGCGTTATCCAGCGATGTGATCTCCGCAT<br/> TATCCTCACAAAAAAGTGAGGATTTTTTTATTTTTGTATTAACA<br/> AAATCAGAGACAATCCGATATTAATGATGTAGCCGGGAGGAGG<br/> CGCAAAAGACTCAGCCAGTTACAAAATAAGGGCACAAGAACGT<br/> GCCTTAACAACATATTCAGGGAGGAACAAAACAATGCGTAAAG<br/> GAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTA<br/> GATGGTGATGTTAATGGGCACAAATTTTCTGTCACTGGAGAGG<br/> GTGAAGGTGATGCAACATACGGAAAACCTACCCTTAAATTTATT<br/> TGCACTACTGGAAAACCTACCTGTTCCATGGCCAACACTTGTCAC<br/> TACTTTTCGGTTATGGTGTTCAATGCTTTGCGAGATACCCAGATC<br/> ATATGAAACGGCATGACTTTTTCAAGAGTGCCATGCCCGAAGG<br/> TTATGTACAGGAAAGAACTATATTTTTCAAAGATGACGGGAACT<br/> ACAAGACACGTGCTGAAGTCAAGTTTGAAGGTGATACCTTGTT</p> |

|  |             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|--|-------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |             | AATAGAATCGAGTTAAAAGGTATTGATTTTAAAGAAGATGGAAA<br>CATTCTTGGACACAAATTGGAATACAACATACTCACACAATG<br>TATACATCATGGCAGACAAACAAAAGAATGGAATCAAAGTTAAC<br>TTCAAAATTAGACACAACATTGAAGATGGAAGCGTTCAACTAGC<br>AGACCATTATCAACAAAAATACTCCAATTGGCGATGGCCCTGTCC<br>TTTTACCAGACAACCATTACCTGTCCACACAATCTGCCCTTTCG<br>AAAGATCCCAACGAAAAGAGAGACCACATGGTCCTTCTTGAGT<br>TTGTAACAGCTGCTGGGATTACACATGGCATGGATGAACATAC<br>AAATAATTTTAAAAAAGACCTTGGCGTTGCCAGGGTCTTTTAATT<br>TAAATTTCTATCTCCTAATCATTCTCATCTGTCACTAACTCAT<br>GATATAATAACCGGATTCTCCACTAACTTTTTATAAATGTATTC<br>CATACAAGAAATCTAAACAGAAGATTTTTTCCAAAAATATGTG<br>TAATCTTATCTCGACTTAGTCGATATAAACGATAGATTGGGGCA<br>TAGGGGATGATCAATTGAACATTGAAAGGCTCACTACGTTACAA<br>CCTGTTTGGGATCGTTATGATACTCAAATACATAATCAGAAAGA<br>TAATGATAACGAGGTTCTGTTTCATCAAGTTTCATATACCAATCT<br>TGCTGAAATGGTGGGGGAAATGAACAAGCTTTTGGAACCTTCG<br>CAAGTTCATCTGAAGTTCGAGCTTCATGACAAGTTAAATGAATA<br>CTATGTAAAGGTAATAGAGGACTCTACAAATGAAGTGATCCGC<br>GAAATTCACCAAAAACGGTGGCTTGATTTTTATGCGGCTATGAC<br>TGAATTTCTTGGGTATTTGTAGATGAAAAAAGTAGAATAGGA<br>GTGGTTTGAGATGGTCACAAGAATAACAGGTCTGGCGTCAGGA<br>ATGGATATAGATGATATCGTATCAAAGCTGATGCAGACAGAAAG<br>AGCGCCGCTTGATAAGCTGACACAAAAAAGCAGACTCTTGAA<br>TGGCAGCGTGACAGCTATCGTGAAGTAACTCAAAAATAAAAG<br>AATTGCAAGATTATATGTCTAAAAATACGTTGACATATCCGAGC<br>ACGTATCAGAGCAACTGGCGTTACCCAACCTTAATC [SEQ ID NO:<br>18] |
|  | phag + AcoD | CTATGACCATGATTACGCCAAGTGAACAATGATCATTTCATACGA<br>AGTACCATGGCCAAATGAACATAAAAGAAGAACAAATCATTCTT<br>TTTGAAAGCGGGATTCCAGGCTTTTTAGAAAGAAAAACAGTTCGT<br>CATACTTCCGCTTTCAGAAGACTCTCCATTTCGTGGCACTGCAGT<br>CCGTCACTTCAGAAAATCTTGCGTTTATCGTCGTAAGTCCGTTT<br>ATCTTTTTTAAGAATTATGAATTTGATCTTGATGAATCAACTGCT<br>GAACTTTTGGATATCGATAATATTCAAGACGTAGAAGTCATGAC<br>AATATTGACTATGGCAGAGCCATTTGAAAAGTCTACTGCGAATT<br>TATTGGCTCCCATTATTGTGAATCGCAAGAACATGATGGCTAAG<br>CAAGTCGTTTTACACGACTCCTCATATACGACAAAGCATCCGAT<br>TGGAGGAGAATCATGCTAGTTTTATCGCGGAAAATAAACGAAG<br>CGATTCAAATAGGTGCTGATATTGAAGTAAAGTGATTGCGGTT<br>GAAGGGGATCAAGTGAAGCTTGGAATTGACGCCCCAAAGCATA<br>TTGATATTCACAGGAAAGAAATTTACTTGACCATTACAGGAAGAA<br>AATAACCGTGACAGCAGCGTTATCCAGCGATGTGATCTCCGCAT<br>TATCCTCACAAAAAAGTGAGGATTTTTTTATTTTGTATTAACA<br>AAATCAGAGACAATCCGATATTAATGATGTAGCCGGGAGGAGG<br>CGCAAAAGACTCAGCCAGTTACAAAATAAGGGCACAAGAACGT<br>GCCTTAACAACATATTTCAGGGAGGAACAAAACAATGAATATGGC                                                                                                                                                                                                                                                                                                                                                                                                |

|  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|--|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  | TGAAATCGCCCAGCTTGGAGTCTCAAACCCGTACAAACAACAG<br>TACGAAAACCTATATTGGCGGAGCTTGGGTTCCGCCGCGCCGGTG<br>GCGAATACTTTGAATCTACAACGCCGATTACGGGAAAACCTTTT<br>ACAAGAGTTCCGCGCTCCGGCCAACAGGATGTGGACGCAGCG<br>TTAGATGCTGCCCATGCAGCGAAAAGCTGCCTGGGCTAGAACAT<br>CAACAACGGAACGCGCCAATATTTTAAACCGCATCGCAGATCG<br>TATTGAAGCGAATTTGAAACTGCTTGCTGTGCGCCGAAAGCATTG<br>ACAACGGAACCTGTAAGAGAAAACAACGGCAGCGGATCTGCC<br>GCTTGCACTGGACCATTTTCGTTATTTTGCAGGTTGCATCAGAG<br>CACAAGAAGGCGGCATTAGCGAAATCGATGCAGACACAATTGC<br>GTACCATTTTCATGAACCTCTGGGTGTTGTGGGCCAGATTATCC<br>CGTGGAATTTTCCTTATTGATGGCGACGTGGAAACTGGCACC<br>GGCGCTTGCTGCCGGAACCTGTGTCGTACTTAAACCTGCAGAA<br>CAAACACCGGCGTCTATCTTAGTTTTGATGGAAGTGATTGGCG<br>ATCTGCTGCCGCCGGGCGTTGTGAATGTCATCAACGTTTTGG<br>CTTAGAAGCTGGCAAACCTTTGGCCTCAAGCCCGCGTATTTCC<br>AAAGTAGCTTTTACGGGTGAAACAACGACAGGCCGGTTAATCA<br>TGCAATATGCATCACAGAATTTGATTCCTGTTACACTGGAACCT<br>GGCGGAAAAAGCCCGAACATTTTCTTTGAAGATGTGTTAGCAG<br>CGGATGACGCATTTTTTCGACAAAGCGCTGGAAGGATTTGCAAT<br>GTTTGCCTTAATCAAGGCGAAGTTTGCACATGTCCTTCACGTG<br>CTCTGATCCAGGAAAGCATTTATGATCGGTTTATGGAAAGAGCC<br>CTGAAACGCGTGGCTGCCATCAGACAAGGACATCCGCTTGACA<br>CGGGAACAATGATTGGTGCTCAAGCCTCTGCAGAACAGTTAGA<br>AAAAATCTTGCTCTACATTGATCTGGGCAGAAAAGAAGGAGCA<br>CAGTGCCTTACGGGTGGCGAACGCAATGTCCTGGATGGCGAC<br>CTTGACGGCGGCTATTACGTCAAACCTACAGTATTTGCGGGAC<br>ATAACAAAATGCGCATCTTTCAAGAAGAAATTTTTGGCCCGGTC<br>GTAAGCGTTACGACATTTAAAGATGAAGAAGAAGCACTGGCTAT<br>CGCCAACGACACGTTATATGGATTGGGTGCGGGCGTTTGGACA<br>AGAGATGGAGCACGTGCGTTTCGGATGGGAAGAGGTATTCAAG<br>CTGGCCGCGTGTGGACGAATTGTTATCATGCTTACCCGCCCCA<br>TGCAGCGTTTGGCGGATATAAACAGTCTGGCATCGGACGTGAA<br>AACCATCGGATGATGTTGGATCATTACCAACAGACAAAAAATTT<br>ATTGGTTTCTTACTCCCCGAACGCGTTGGGCTTTTTCTAATTTTA<br>AAAAAGACCTTGGCGTTGCCAGGGTCTTTTAATTTAAATTTCTAT<br>CTCCTAATCATTCTCATCCTGTCACTAACTCATGATATAATAAC<br>CGGATTCTCCACTAACTTTTTATAAATGTATTTCCATACAAGAAA<br>TCTAAAACAGAAGATTTTTTTTCCAAAAATATGTGTAATCTTATCT<br>CGACTTAGTCGATATAAACGATAGATTGGGGCATAGGGGATGA<br>TCAATTGAACATTGAAAGGCTCACTACGTTACAACCTGTTTGGG<br>ATCGTTATGATACTCAAATACATAATCAGAAAGATAATGATAACG<br>AGGTTCTGTTCATCAAGTTTCATATACCAATCTTGCTGAAATG<br>GTGGGGGAAATGAACAAGCTTTTGAACCTTCGCAAGTTCATC<br>TGAAGTTCGAGCTTCATGACAAGTTAAATGAATACTATGTAAAG<br>GTAATAGAGGACTCTACAAATGAAGTGATCCGCGAAATTCAC<br>CAAAACGGTGGCTTGATTTTTATGCGGCTATGACTGAATTTCTT |
|--|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|  |             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|--|-------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |             | GGGTTATTTGTAGATGAAAAAAGTAGAATAGGAGTGGTTTTGAG<br>ATGGTCACAAGAATAACAGGTCTGGCGTCAGGAATGGATATAG<br>ATGATATCGTATCAAAGCTGATGCAGACAGAAAGAGCGCCGCT<br>TGATAAGCTGACACAAAAAAGCAGACTCTTGAATGGCAGCGT<br>GACAGCTATCGTGAAGTAAACTCAAAAATAAAGAATTGCAAGA<br>TTATATGTCTAAAAATACGTTGACATATCCGAGCACGTATCAGA<br>GCAACTGGCGTTACCCAACCTTAATC [SEQ ID NO: 19]                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|  | phag + AldB | CTATGACCATGATTACGCCAAGTGAACAATGATCATTTCATACGA<br>AGTACCATGGCCAAATGAACATAAAAGAAGAACAAATCATTCTT<br>TTTGAAAGCGGGATTCCAGGCTTTTTAGAAAGAAAAACAGTTCGT<br>CATACTTCCGCTTTCAGAAGACTCTCCATTCGTGGCACTGCAGT<br>CCGTCACTTCAGAAAATCTTGCGTTTATCGTCGTAAGTCCGTTT<br>ATCTTTTTTAAGAATTATGAATTTGATCTTGATGAATCAACTGCT<br>GAACTTTTGGATATCGATAATATTCAAGACGTAGAAGTCATGAC<br>AATATTGACTATGGCAGAGCCATTTGAAAAGTCTACTGCGAATT<br>TATTGGCTCCCATTATTGTGAATCGCAAGAACATGATGGCTAAG<br>CAAGTCGTTTTACACGACTCCTCATATACGACAAAGCATCCGAT<br>TGGAGGAGAATCATGCTAGTTTTATCGCGGAAAATAAACGAAG<br>CGATTCAAATAGGTGCTGATATTGAAGTAAAAGTGATTGCGGTT<br>GAAGGGGATCAAGTGAAGCTTGGAATTGACGCCCCAAAGCATA<br>TTGATATTCACAGGAAAGAAATTTACTTGACCATTACAGGAAGAA<br>AATAACCGTGCAGCAGCGTTATCCAGCGATGTGATCTCCGCAT<br>TATCCTCACAAAAAAGTGAGGATTTTTTTATTTTTGTATTAACA<br>AAATCAGAGACAATCCGATATTAATGATGTAGCCGGGAGGAGG<br>CGCAAAAGACTCAGCCAGTTACAAAATAAGGGCACAAGAACGT<br>GCCTTAACAACATATTCAGGGAGGAACAAAACAATGACCAATAA<br>TCCCCCTTCAGCACAGATTAAGCCCGGCGAGTATGGTTTCCCC<br>CTCAAGTTAAAAGCCCGCTATGACAACCTTTATTGGCGGCGAAT<br>GGGTAGCCCCTGCCGACGGCGAGTATTACCAGAATCTGACGC<br>CGGTGACCGGGCAGCTGCTGTGCGAAGTGGCGTCTTCGGGCA<br>AACGAGACATCGATCTGGCGCTGGATGCTGCGCACAAAGTGAA<br>AGATAAATGGGCGCACACCTCGGTGCAGGATCGTGCGGCGAT<br>TCTGTTTAAGATTGCCGATCGAATGGAACAAAACCTCGAGCTGT<br>TAGCGACAGCTGAAACCTGGGATAACGGCAAAACCCATTTCGCGA<br>AACCAGTGCTGCGGATGTACCGCTGGCGATTGACCATTTCGCGC<br>TATTCGCCTCGTGTATTTCGGGCGCAGGAAGGTGGGATCAGTG<br>AAGTTGATAGCGAAACCGTGGCCTATCATTCCATGAACCGTTA<br>GGCGTGGTGGGGCAGATTATCCCGTGGAACCTCCCGCTGCTG<br>ATGGCGAGCTGGAAAATGGCTCCCGCGCTGGCGGCGGGCAAC<br>TGTGTGGTGTGAAACCCGCACGTCTTACCCCGCTTTCTGTAC<br>TGCTGCTAATGGAAATTGTCGGTGATTTACTGCCGCCGGGCGT<br>GGTGAACGTGGTCAATGGCGCAGGTGGGGTAATTGGCGAATA<br>TCTGGCGACCTCGAAACGCATCGCCAAAGTGGCGTTTACCGGC<br>TCAACGGAAGTGGGCCAACAAATTATGCAATACGCAACGCAAA<br>ACATTATTCGGTGACGCTGGAGTTGGGCGGTAAAGTCGCCAAA<br>TATCTTCTTTGCTGATGTGATGGATGAAGAAGATGCCTTTTTCG<br>ATAAAGCGCTGGAAGGCTTTGCACTGTTTGCCTTTAACAGGG |

```

CGAAGTTTGCACCTGTCCGAGTCGTGCTTTAGTGCAGGAATCT
ATCTACGAACGCTTTATGGAACGCGCCATCCGCCGTGTCGAAA
GCATTGCGTAGCGGTAACCCGCTCGACAGCGTGACGCAAATGG
GCGCGCAGGTTTCTCACGGGCAACTGGAACCATCCTCAACTA
CATTGATATCGGTAAAAAAGAGGGCGCTGACGTGCTCACAGGC
GGGCGGCGCAAGCTGCTGGAAGGTGAACTGAAAGACGGCTAC
TACCTCGAACCGACGATTCTGTTTGGTCAGAACAAATATGCGGG
TGTTCCAGGAGGAGATTTTTGGCCCGGTGCTGGCGGTGACCA
CCTTCAAAACGATGGAAGAAGCGCTGGAGCTGGCGAACGATAC
GCAATATGGCCTGGGCGCGGGCGTCTGGAGCCGCAACGGTAA
TCTGGCCTATAAGATGGGGCGCGGCATACAGGCTGGGCGCGT
GTGGACCAACTGTTATCACGCTTACCCGGCACATGCGGCGTTT
GGTGGCTACAAACAATCAGGTATCGGTGCGGAAACCCACAAGA
TGATGCTGGAGCATTACCAGCAAACCAAGTGCCTGCTGGTGAG
CTACTCGGATAAACCGTTGGGGCTGTTCTAATTTAAAAAAGAC
CTTGGCGTTGCCAGGGTCTTTAATTTAAATTTCTATCTCCTAAT
CATTCTCATCCTGTCACTAACTCATGATATAATAACCGGATTCT
CCACTAACTTTTTATAAATGTATTTCCATACAAGAAATCTAAAAC
AGAAGATTTTTTCCAAAAATATGTGTAATCTTATCTCGACTTAG
TCGATATAAACGATAGATTGGGGCATAGGGGATGATCAATTGA
ACATTGAAAGGCTCACTACGTTACAACCTGTTTGGGATCGTTAT
GATACTCAAATACATAATCAGAAAGATAATGATAACGAGGTTCC
TGTTTCATCAAGTTTCATATACCAATCTTGCTGAAATGGTGGGGG
AAATGAACAAGCTTTTGAACCTTCGCAAGTTCATCTGAAGTTC
GAGCTTCATGACAAGTTAAATGAATACTATGTAAAGGTAATAGA
GGACTCTACAAATGAAGTGATCCGCGAAATTCACCAAAAACGG
TGGCTTGATTTTTATGCGGCTATGACTGAATTTCTTGGGTATT
GTAGATGAAAAAAGTAGAATAGGAGTGGTTTGAGATGGTCAC
AAGAATAACAGGTCTGGCGTCAGGAATGGATATAGATGATATC
GTATCAAAGCTGATGCAGACAGAAAGAGCGCCGCTTGATAAGC
TGACACAAAAAAGCAGACTCTTGAATGGCAGCGTGACAGCTA
TCGTGAAGTAACTCAAAAATAAAAGAATTGCAAGATTATATGT
CTAAAAATACGTTGACATATCCGAGCACGTATCAGAGCAACTG
GCGTTACCCAACTTAATC [SEQ ID NO: 20]

```

#### Replacement of *hag* with *acoD*, *aldB* or GFP gene

**[0118]** Double-stranded DNA encoding the sequence of GFP, or the aldehyde dehydrogenases AcoD (from *C. necator*) and AldB (from *E. coli*) flanked by 800 base pairs 5' and 3' of the *hag* gene were constructed (see table 1 for sequences). In addition, the CsrA-binding site was disrupted by specifying that each construct have a single point mutation in the 5' flanking *hag* homology precisely 38 base pairs 5' of the GFP/AcoD/AldB coding sequence, changing the naturally occurring "G" to an "A" ("A" mutation is highlighted in these sequences in table 2). These were ligated into pMiniMAD plasmid linearized with ZP24 and ZP25 via Gibson assembly. The resulting plasmid was transformed into *E. coli* via heat

transformation. Plasmid was miniprepmed using a commercially available kit, and that miniprepmed plasmid was used as the DNA source to do the chromosomal modification following the same protocol as for the *flgM* deletion.

GFP assay to quantify promoter strength

**[0119]** Strains to be tested were streaked out on LB plates and grown overnight at 37°C. Single colonies from overnight growth were inoculated into 3mL of LB and 1mM IPTG was added to induce strain of *B. subtilis* with GFP under IPTG-inducible pHyspank promoter. Strains were grown at 37°C with shaking at 275rpm for 7 hours. Timepoints were taken at 2.5 hours, 4 hours, 5.5 hours, and 7 hours. Timepoints were taken by pulling 200uL of culture from each tube and spinning down at 13,000xg for 2 minutes, and then resuspending the pellet in 200uL of PBS. These resuspensions were then aliquoted into wells of a costar black 96-well plate and read on a fluorometer (excitation 485nm; emission 535nm). After fluorescence reading, 100uL of each well was taken and read for absorbance at 600nm on a spectrophotometer to normalize the fluorescence to optical density of the culture.

**[0120]** For plate visualization, frozen stocks were streaked out in patches on LB plates and grown overnight at 37°C. Fluorescence was visualized with a blue LED flashlight (480nm wavelength) and an orange filter.

HPLC assay to quantify ALDH activity

**[0121]** Sample preparation: Single colonies from overnight growth at 37°C on LB plates were inoculated into 4mL of LB medium and grown at 37°C with shaking at 275rpm for 1.75 hours to OD<sub>600</sub> ~0.2 and then split into two tubes of 2mL each. One tube for each strain was induced with 1mM IPTG and the other tube was left uninduced as a control. 5 hours after induction (6.75 hours of total growth), the OD<sub>600</sub> was assessed (all strains were at OD<sub>600</sub> ~4), and 1.5mL of the culture was spun down at 6000xG for 5 minutes and the pellet was resuspended in 200uL of PBS. 22uL of 2M acetaldehyde was added to each tube of cells to give a final concentration of 200mM acetaldehyde. The tubes were then incubated at 37°C for 30 minutes. The cells were then spun down at 20,000xG for 10 minutes at 4°C, and 150uL of the resulting supernatant was added to 1350uL of PBS (thus, if no acetaldehyde were lost, the final concentration should now be 20mM). These diluted supernatants were then frozen at -20°C in 1.5mL microcentrifuge tubes.

**[0122]** Standard preparation: 168uL of acetaldehyde was added to 1332uL of PBS to make 2M stock and frozen at -20°C. Samples to make the standard curve for HPLC were prepared fresh the day of the experiment by diluting from this 2M stock into PBS for final concentrations of 1mM, 2.5mM, 5mM, 10mM, and 20mM standards. Each standard was made in duplicate, with one set run before the samples and one set run after the samples. All standards and samples were then loaded into HPLC vials filled up to the neck to limit

head space. Samples, but not standards, were filtered through a 0.2µm filter to remove any residual bacteria when loaded into the HPLC vials.

**[0123]** Samples were run on a Shimadzu system with an Aminex HPX-87H column. Samples were run at 0.6mL/min flow rate and isocratic elution with 0.005M sulfuric acid mobile phase at 50° C (column temp). Samples were detected via refractive index.

Qualitative analysis of AcoD expression (Coomassie stain)

**[0124]** Single colonies from overnight growth at 37°C on LB plates were inoculated into 2.5mL of LB medium (+1mM IPTG for strains with pHyspank promoter) and grown at 37°C with shaking at 275rpm for 4.5 hours. 1.5mL of each culture was spun down at 13,000xG for 2 minutes and the pellet was resuspended in 50uL of lysis buffer (20mM TRIS, 10mM EDTA, 0.1% lysozyme, 0.01% RNaseA, 0.002% DNaseI, and Roche EDTA-free cOmplete protease inhibitor cocktail tablet diluted per manufacturer specifications). Samples were then incubated in lysis buffer for 30 minutes in 37°C water bath. 50uL of 2x laemmli buffer was then added to samples and the samples were incubated at 100°C on a thermocycler for 5 minutes. Samples were then spun down at 13,000xG for 2 minutes and 10uL were loaded onto an 8% SDS-PAGE gel and run at 140V for 1 hour. The gel was then rinsed with ddH<sub>2</sub>O 6 times with shaking for 5 minutes each. It was then incubated in Coomassie brilliant blue stain overnight at room temperature with shaking.

Results:

Expression of ALDH genes heterologously in *B. subtilis* results in physiologically relevant rates of acetaldehyde breakdown.

**[0125]** To assess the ability of *B. subtilis* to heterologously express a functional acetaldehyde dehydrogenase, a strain encoding an acetaldehyde dehydrogenase gene under the control of a pHyspank promoter, which is IPTG inducible was constructed. Two different acetaldehyde dehydrogenases were selected after review of the literature. The criteria for selection were: characterized enzyme kinetics and high specificity for acetaldehyde; bacterial origin; from a non-pathogenic bacteria; soluble protein product (i.e. not membrane-bound); and amino acid sequence known. After conducting such a search, several candidates were identified, and two were selected as best fits in these criteria. The first was AcoD, encoded by the *acoD* gene from *C. necator* [12], which has an experimentally determined Km for acetaldehyde of ~4µM, and uses NAD as its co-factor for oxidizing acetaldehyde to acetate. The second was AldB, encoded by the *aldB* gene of *E. coli* [13], which has an experimentally determined Km of ~2.5µM and a Vmax of 2U/mg, and uses NADP as its co-factor.

**[0126]** *B. subtilis* with *acoD* under control of the pHyspank promoter and *B. subtilis* with *aldB* under control of the pHyspank promoter were grown with and with IPTG induction and

then incubated with 200mM IPTG for 30 minutes at room temperature. The samples were then filtered to remove all bacteria and frozen at -20°C.

**[0127]** Samples were thawed and diluted 1:10 in PBS and standards were prepared freshly the day of the HPLC assay. All samples were then run on an HPLC to quantify the amount of acetaldehyde. In addition to running two sets of standards, one at the beginning and one at the end of the run, a single preparation of a 10mM standard of acetaldehyde was run at the beginning and the end of the run (i.e. the same tube run twice) to control for any potential evaporation or other stochastic loss of acetaldehyde during the HPLC run, since the tubes would be on the machine for several hours and it was desired to ensure that samples run later in the order would not erroneously appear lower. A standard curve was generated based on area under the curve for each standard sample, and the best-fit line had an  $R^2$  of >0.99. The equation for this line was then used to generate concentrations for each of the samples. The results of this assay are presented in FIG. 1. There was essentially no difference in the acetaldehyde standard run twice, indicating no significant stochastic loss of acetaldehyde from the samples during the run. In addition, the samples with uninduced bacteria had acetaldehyde levels that corresponded nicely with the 20mM standard, demonstrating that the uninduced bacteria do not effectively remove acetaldehyde by naturally occurring mechanisms.

**[0128]** However, when AldB or AcoD were induced, the samples showed a reduction of ~4mM and ~8mM (respectively) in acetaldehyde concentration, indicating that induction of the protein caused removal of acetaldehyde.

Combination of *flgM* deletion and CsrA-binding site point mutation results in robust and constitutive heterologous protein expression via the *hag* promoter.

**[0129]** Flagellin, encoded by the *hag* gene in *B. subtilis*, is expressed at extremely high levels when the cell determines that motility is advantageous. However, because the manufacture of a flagellum is very nutritionally and energetically expensive for the cell, it has several mechanisms to repress production of Hag at times in the life cycle when motility would be inappropriate.

**[0130]** FlgM is known to bind and repress the sigma factor SigD that activates transcription of *hag*. Thus, deletion of *flgM* should remove transcriptional repression and result in high levels of expression constitutively. *hag* translation into protein is further repressed post-transcriptionally by CsrA, which binds to *hag* transcripts at the ribosome binding site, thus competitively inhibiting ribosome translation. A single point mutation in this binding site precludes CsrA binding, but not ribosome binding, thus allowing for constitutive translation of *hag* transcript.

**[0131]** To test the hypothesis that removal of these two key mechanisms of repression would result in robust and constitutive expression of a heterologous protein, *flgM* was

deleted, a single point mutation was made in the CsrA-binding site, and the *hag* gene was replaced with a reporter of expression, GFP. The fluorescence of this strain on a plate was compared to that of a strain expressing GFP via the de-repressed pHyspank promoter (*lacI* was deleted from this construct to ensure constitutive expression, and the delta *lacI* strain was shown to have similar levels of GFP expression to a strain with *lacI* intact and induced with 1mM IPTG [data not shown]). When struck out side by side, two different isolates of the modified *hag* expression system were qualitatively much brighter than the pHyspank expression system (figure 2a).

**[0132]** However, to see if expression was indeed constitutive, and to quantify the difference between these two expression systems, a broth growth and fluorescence time course was taken. Strains were grown in LB, and the OD<sub>600</sub> and fluorescence were assessed at 2.5 hours of growth, and then every 90 minutes thereafter up to 7 hours (FIG. 2b). Indeed, after 2.5 hours the fluorescence normalized to OD<sub>600</sub> was ~4 times higher in the modified *hag* expression system than in the pHyspank expression system. This increased to ~10 times higher by 4 hours and remained ~10 times higher through the 7 hours assessed.

Utilizing the de-repressed *hag* promoter, qualitatively higher levels of AcoD are produced than by pHyspank.

**[0133]** To ensure robust and constitutive expression via our modified *hag* promoter was not specific to GFP but could be utilized for other heterologous protein expression, including aldehyde dehydrogenases, two different aldehyde dehydrogenase genes were inserted into the *hag* locus of two different strains: *acoD* from *C. necator* and *aldB* from *E. coli*. These strains were then grown up in broth culture along with strains expressing *acoD* and *aldB* via the pHyspank promoter and a strain expressing GFP via the modified *hag* promoter as a positive control. The bacteria were pelleted and lysed and lysates were run on an SDS-PAGE gel and stained for total protein using Coomassie. As expected, there is an obvious band for GFP (FIG. 3). In addition, there are equally bright bands for his-tagged AcoD, untagged AcoD, and his-tagged AldB when expressed via the modified *hag* promoter (FIG. 3). However, there are only faint bands of the expected size for AcoD and AldB under induced pHyspank control. Despite this apparently low level of expression, it is known from the HPLC experiment that this level is sufficient for removal of acetaldehyde at physiologically relevant rates. Thus, as the Coomassie staining demonstrates that there is qualitatively obviously more protein being produced via the modified *hag* promoter, it is reasonable to expect that this expression system would produce at least as robust a rate of acetaldehyde removal, if not much more.

Discussion:

**[0134]** Enzymatic approaches are rapidly replacing small molecules and becoming the standard in the pharmaceutical world in large part due to their superior efficacy, versatility,

and specificity. Evidence of this can be seen in the fact that the biologics drug market was, in 2017, worth over \$200B and set to double in the next seven years. In addition, 8 of the top 10 best-selling drugs in 2016 were biologics.

**[0135]** However, the current process to make biologic drugs is long, expensive, and generally approached with a mindset similar to that used to manufacture small molecule drugs. A stepwise change in method is necessary in order to make this powerful class of drug available for broader uses. The purpose of this invention is to make a cost-effective, orally deliverable enzymatic therapy for the alcohol hangover.

**[0136]** As demonstrated in FIG. 1, intracellular expression of two different heterologous proteins resulted in marked removal of acetaldehyde. Indeed, if one assumes that a heavy night of drinking causes a colonic acetaldehyde level of ~150uM [14], and there are approximately 15L of extracellular fluid in a 72kg human body, then the same amount of bacteria could remove 150uM of acetaldehyde from 15L in roughly 3 hours. While of course this is a "back-of-the-envelope" calculation, it is at least not unreasonable to assume that these could represent relatively physiologically relevant rates.

**[0137]** In addition, the no-bacterial control demonstrates that the drop was not due to simple evaporation or some other method of removal, and the uninduced controls demonstrate that the removal was specific to expression of the ALDH and not some inherent function of the bacteria. Furthermore, the re-reading of the same 20mM standard tube at the beginning and end of the run demonstrate that any drop was not due to evaporation during the HPLC run. Also, the removal of acetaldehyde was dependent on induction of the aldehyde dehydrogenases, indicating that the bacteria do not naturally remove acetaldehyde on their own, but require the heterologously expressed acetaldehyde dehydrogenase.

**[0138]** By removing FlgM and CsrA repression, rapid and robust expression of GFP constitutively was achieved throughout the growth cycle of our strain. The fact that these levels are 5-10 times higher than the expression levels using the pHyspank promoter -- an academic gold standard for expression in *B. subtilis* -- demonstrates the utility of the expression system. Furthermore, it is also important to note that the fluorescence/OD<sub>600</sub> ratio continued to increase throughout the entire growth curve, indicating that the fluorescence was increasing faster than the OD, meaning that each cell was continuing to produce more GFP. This is consistent with the hypothesis that the *flgM* deletion and CsrA-binding site mutation result in constitutive and robust expression. Had the expression not been constitutive throughout the growth curve the ratio of fluorescence to OD<sub>600</sub> would have eventually plateaued or dropped, even if the absolute fluorescence continued to rise.

**[0139]** In addition, the demonstration of physiologically relevant rates of acetaldehyde removal was done using the pHyspank promoter, indication that our system utilizing the edited pHag promoter generates more protein and should even more sufficiently exceed

necessary rates of acetaldehyde removal. Indeed, qualitative comparisons of band densities of AcoD in a Coomassie gel (fig. 2b) demonstrate a much more robust protein band indicating far superior protein expression.

**[0140]** Taken together, a novel system for acetaldehyde removal from the body has been developed that is distinct from any other strategy currently described or executed. It involves edible bacteria engineered to express heterologous protein constitutively for the purpose of enzymatically removing acetaldehyde directly. It utilizes a heterologous protein expression strategy that takes advantage of the robust flagellin promoter, enhanced by removal of repressors of that promoter.

#### EXAMPLE 2

**[0141]** A subject ingests a water suspension of 1E9 CFU of a spore preparation *B. subtilis* with the following modifications: FlgM deleted, a mutation in BS1 as described herein and the *hag* gene replaced with AcoD as described herein. Immediately following this, the person ingests six alcoholic beverages (12 ounces of 5% ethanol) over the next 3 hours before going to bed. Typically, this amount of alcohol results in this person having an alcohol hangover characterized by: grogginess, severe headache, nausea, malaise, sensitivity to light and sound, vertigo, and irritability. These symptoms typically set in within the 8 hours while the person sleeps and persist for 12 hours thereafter. However, the person awakens 8 hours later with none of the excess acetaldehyde formed by his or her body due to the metabolism of the ethanol consumed the night before. Thus, while he or she has mild grogginess and a slight headache due to dehydration, the person has greatly reduced or none of the other symptoms associated with a hangover.

#### EXEMPLARY EMBODIMENTS

**[0142]** 1. A method for preventing or treating an alcohol hangover, comprising administering to the gut of a subject in need thereof an effective amount of a composition comprising a recombinant bacterium recombinantly engineered to constitutively express an aldehyde dehydrogenase.

**[0143]** 2. The method of embodiment 1, wherein the recombinant bacteria comprises:  
a) a polynucleotide comprising an expression construct comprising a flagellin gene promoter operatively linked with a heterologous nucleotide sequence encoding an aldehyde dehydrogenase, wherein the flagellin gene promoter comprises one or more genetic modifications that reduce CsrA repression of translation of an mRNA transcribed from the flagellin gene promoter; and b) a genetic modification of a sigma factor repressor, wherein the modification reduces inhibition of sigma factor initiation of transcription from the flagellin gene promoter.

- [0144] 3. The method of embodiment 2, wherein the bacterium belongs to a genus selected from *Bacillus*, *Bifidobacterium*, *Enterococcus*, *Escherichia*, *Lactobacillus*, *Leuconostoc*, *Pediococcus*, *Pediococcus* and *Streptococcus*.
- [0145] 4. The method of embodiment 2, wherein (1) the bacterium belongs to genus *Bacillus*, (2) the flagellin gene promoter is a modified *Bacillus* hag promoter comprising one or more genetic modifications of the CsrA BS1 binding site and/or the CsrA BS2 binding site, (3) the sigma factor is SigD and (4) the SigD repressor is FlgM.
- [0146] 5. The method of embodiment 4, wherein the hag promoter comprises one or more genetic modifications in the 12-base-pair BS1 binding site or in the surrounding bases on either side of the binding site that form the stem of the stem-loop secondary structure of BS1.
- [0147] 6. The method of embodiment 4, wherein the hag promoter comprises a modification of the CsrA BS1 recognition sequence, AGGA.
- [0148] 7. The method of embodiment 4, wherein the hag promoter comprises the nucleotide sequence GCACAAGAACGT [SEQ ID NO:2].
- [0149] 8. The method of embodiment 4, wherein the one or more genetic modifications comprise one or more point mutations to the CsrA BS2 binding site.
- [0150] 9. The method of embodiment 4, wherein the hag promoter comprises one or more genetic modifications that disrupt the stem and loop structure of BS2 by eliminating complementarity that allows hydrogen bonding.
- [0151] 10. The method of embodiment 4, wherein the hag promoter comprises a modification of the BS2 binding site, wherein the modification comprises the nucleotide sequence ATTTAGGGAGGAA [SEQ ID NO:3].
- [0152] 11. The method of embodiment 1, wherein the aldehyde dehydrogenase is a bacterial aldehyde dehydrogenase.
- [0153] 12. The method of embodiment 4, wherein the aldehyde dehydrogenase is selected from (1) AcoD from *Cupriavidus necator* and (2) AldB from *E. coli*.
- [0154] 13. The method of embodiment 4, wherein the genetic modification in a flgM gene comprises deletion of all or part of the flgM gene.
- [0155] 14. The method of embodiment 4, wherein the genetic modification in a flgM gene disrupts secondary or tertiary structure.
- [0156] 15. The method of embodiment 4, wherein the expression construct is located in a bacterial chromosome.
- [0157] 16. The method of embodiment 1, comprising administering the composition to the subject during consumption of alcohol.
- [0158] 17. The method of embodiment 1, comprising administering the composition to the subject up to 24 hours before commencement of consumption of alcohol.

- [0159]** 18. The method of embodiment 1, wherein the composition is administered orally.
- [0160]** 19. The method of embodiment 1, wherein the composition further comprises a physiologically acceptable carrier selected from lactic acid fermented foods, fermented dairy products, resistant starch, dietary fibers, carbohydrates, proteins and glycosylated proteins, water, capsule filler, and a gummy material.
- [0161]** 20. The method of embodiment 1, wherein the composition comprises about  $10^4$  to about  $10^{12}$  colony forming units of the recombinant bacteria.
- [0162]** 21. A recombinant microorganism comprising: a) a polynucleotide comprising an expression construct comprising a flagellin gene promoter operatively linked with a heterologous nucleotide sequence encoding a subject polypeptide, wherein the flagellin gene promoter comprises one or more genetic modifications that reduce CsrA repression of translation of an mRNA transcribed from the flagellin gene promoter; and b) a genetic modification of a *flgM* gene that reduces inhibition of SigD initiation of transcription.
- [0163]** 22. The recombinant microorganism of embodiment 21, which constitutively expresses the polypeptide.
- [0164]** 23. The recombinant microorganism of embodiment 21, wherein the microorganism is probiotic.
- [0165]** 24. The recombinant microorganism of embodiment 21, wherein the microorganism belongs to genus selected from *Bacillus*, *Bifidobacterium*, *Enterococcus*, *Escherichia coli*, *Lactobacillus*, *Leuconostoc*, *Pediococcus*, *Pediococcus* and *Streptococcus*.
- [0166]** 25. The recombinant microorganism of embodiment 21, wherein the microorganism belongs to genus *Bacillus*.
- [0167]** 26. The recombinant microorganism of embodiment 21, wherein the microorganism is *B. subtilis*.
- [0168]** 27. The recombinant microorganism of any of the foregoing embodiments, wherein the flagellin gene promoter is a *hag* promoter.
- [0169]** 28. The recombinant microorganism of embodiment 27, wherein the *hag* promoter comprises one or more genetic modifications that reduce CsrA repression of translation of an mRNA transcribed from the flagellin gene promoter, wherein the genetic modification comprises modification of the CsrA BS1 binding site and/or CsrA BS2 binding site (e.g., nucleotide substitution, insertion or deletion).
- [0170]** 29. The recombinant microorganism of embodiment 28, wherein the one or more genetic modifications comprise one or a plurality (e.g., two, three or four) of genetic modifications to the CsrA BS1 recognition sequence, AGGA, e.g., to the sequence AGAA.
- [0171]** 30. The recombinant microorganism of embodiment 28, wherein the one or more genetic modifications comprise one or more genetic modifications in the 12-base-pair BS1

binding site or in the surrounding bases on either side of the binding site that form the stem of the stem-loop secondary structure of BS1.

**[0172]** 31. The recombinant microorganism of embodiment 28, wherein the one or more genetic modifications comprise one or a plurality (e.g., at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12) of genetic modifications in the BS1 binding site, gcacaaggacgt.

**[0173]** 32. The recombinant microorganism of embodiment 28, wherein the one or more genetic modifications disrupt the stem and loop structure of BS1 by eliminating complementarity that allows hydrogen bonding.

**[0174]** 33. The recombinant microorganism of embodiment 28, wherein the one or more genetic modifications comprise one or a plurality of genetic modifications in the sequence taagggcacaaggacgtgcctta [SEQ ID NO: 1] that are involved in hydrogen bonding, for example, to eliminate one, two, three, four or more hydrogen bond pairs.

**[0175]** 34. The recombinant microorganism of embodiment 28, wherein the modified BS1 has the nucleotide sequence GCACAAGAACGT [SEQ ID NO: 2].

**[0176]** 35. The recombinant microorganism of embodiment 28, wherein the one or more genetic modifications comprise one or more point mutations to the CsrA BS2 binding site.

**[0177]** 36. The recombinant microorganism of embodiment 28, wherein the one or more genetic modifications comprise one or a plurality (e.g., at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13) of genetic modification in the 13-base-pair BS2 binding site or in the surrounding bases on either side of the binding site that form the stem of the stem-loop secondary structure of BS2.

**[0178]** 37. The recombinant microorganism of embodiment 28, wherein the one or more genetic modifications disrupt the stem and loop structure of BS2 by eliminating complementarity that allows hydrogen bonding.

**[0179]** 38. The recombinant microorganism of embodiment 28, wherein the modified BS2 has the nucleotide sequence ATTTAGGGAGGAA [SEQ ID NO: 3].

**[0180]** 39. The recombinant microorganism of embodiment 28, wherein the one or more genetic modifications to the BS2 binding site does not include an alteration of nucleotides in the Shine-Dalgarno sequence agggagga.

**[0181]** 40. The recombinant microorganism of any of the foregoing embodiments, wherein the flagellin gene promoter is located in a bacterial chromosome or in a plasmid.

**[0182]** 41. The recombinant microorganism of any of the foregoing embodiments, wherein the subject polypeptide is an aldehyde dehydrogenase.

**[0183]** 42. The recombinant microorganism of embodiment 41, wherein the aldehyde dehydrogenase is AcoD from *Cupriavidus necator* and comprises an amino acid sequence

identical to or substantially identical to:

MNMAEIAQLGVSNPYKQQYENYIGGAWVPPAGGEYFESTTPITGKPFTRVPRSGQQDVDA  
ALDAAHAAKAAWARTSTTERANILNRIADRIEANLKLVAESIDNGKPVRETTAADLPLAVD  
HFRYFAGCIRAQEGGISEIDADTIAYHFHEPLGVVGQIIPWNFPLLMATWKLAPALAAGNCVV  
LKPAEQTPASILVLMEVIGDLLPPGVNVINGFGLEAGKPLASSPRISKVAFTGETTTGRLIM  
QYASQNLIPVTLELGGKSPNIFEDVLAADDAFFDKALEGFAMFALNQGEVCTCPSRALIQE  
SIYDRFMERALKRVAAIRQGHPLDTGTMIGAQASAEQLEKILSYIDLGRKEGAQCLTGGERN  
VLDGDLAGGYVVKPTVFAGHNKMRIFQEEIFGPVWSVTTFKDEEEALAIANDTLYGLGAGV  
WTRDGARAFRMGRGIQAGRVWTNCYHAYPAHAAGGYKQSGIGRENHRMMLDHYQQTK  
NLLVSYSPNALGFF [SEQ ID NO: 4].

**[0184]** 43. The recombinant microorganism of embodiment 41, wherein the aldehyde dehydrogenase is a human aldehyde dehydrogenase, e.g., having an amino acid sequence identical to or substantially identical to:

MLRAAARFGPRLGRRLLSAAATQAVPAPNQPEVFCNQIFINNEWHDAVSRTKFTVNPS  
TGEVICQVAEGDKEDVDKAVKAARAAFQLGSPWRRMDASHRGRLNRLADLIERDRTYLA  
ALETLDNGKPYVISYLVLDLDMVLKCLRYAGWADKYHGKTIPIDGFFSYTRHEPVGVC  
QIIPWNFPLLMQAWKLGPALATGNVVMKVAEQTPLTALYVANLIKEAGFPPGVVNIVPG  
FGPTAGAAIASHEDVDKVAFTGSTEIGRVIQVAAGSSNLKRVTLLELGGKSPNIIMSDADM  
DWAVEQAHFALFFNQGCCAGSRTFVQEDIYDEFVERSVARAKSRVVGPNPFDKTEQG  
P  
QVDETQFKKILGYINTGKQEGAKLLCGGGIAADRGYFIQPTVFGDVQDGMTIAKEEIFGP  
VMQILKFKTIEEVGRANNSTYGLAAAVFTKDLDKANYLSQALQAGTVWVNCYDVFGAQS  
PFGGYKMSGSGRELGEYGLQAYTEVKTVTVPKPKNS [SEQ ID NO: 5].

**[0185]** 44. The recombinant microorganism of embodiment 21, wherein the genetic modification in a flgM gene comprises deletion of all or part of the flgM gene.

**[0186]** 45. The recombinant microorganism of embodiment 21, wherein the genetic modification in a flgM gene comprises a single mutation or series of mutations in the sequence encoding active sites of flgM.

**[0187]** 46. The recombinant microorganism of embodiment 21, wherein the genetic modification in a flgM gene disrupts secondary or tertiary structure, such as in one of the helices that defines FlgM secondary structure.

**[0188]** 47. The recombinant microorganism of embodiment 21, wherein the genetic modification in a flgM gene comprises altering an amino acid in the 3rd or 4th helix at the C-

terminal end of the FlgM protein, e.g., selected from I-58, K-62, I-65, G-68, D-73, A-78 of the *B. subtilis* FlgM.

**[0189]** 48. The recombinant microorganism of embodiment 21, wherein the genetic modification in a flgM gene comprises altering one or more amino acids predicted to participate in FlgM binding to SigD, e.g., selected from I-3, G-7, S-10, V-11, A-40, K-41, M43, I-58, L-61, K-62, I-65, Y-70, K-71, V-72, D-73, A-74, H-76, I-77, A-78, N-80, M-81, I-82, N-83, F-84, Y-85, and K-86 of the *B. subtilis* FlgM.

**[0190]** 49. A recombinant probiotic microorganism that constitutively expresses an aldehyde dehydrogenase.

**[0191]** 50. The recombinant probiotic microorganism of embodiment 49, wherein the microorganism comprises: a) a polynucleotide comprising an expression construct comprising a flagellin gene promoter operatively linked with a heterologous nucleotide sequence encoding a subject polypeptide; and b) a genetic modification in a FlgM gene that reduces inhibition of SigD expression.

**[0192]** 51. The recombinant probiotic microorganism of embodiment 50, wherein the flagellin gene promoter comprises one or more genetic modifications that reduce CsrA repression of translation of an mRNA transcribed from the promoter.

**[0193]** 52. A polynucleotide comprising an expression construct comprising a flagellin gene promoter operatively linked with a heterologous nucleotide sequence encoding a subject polypeptide, wherein the flagellin gene promoter comprises one or more genetic modifications that reduce CsrA repression of translation an mRNA transcribed from the flagellin gene promoter.

**[0194]** 53. A method of making a polypeptide comprising culturing a recombinant microorganism of embodiment 21, or embodiment 49.

**[0195]** 54. The method of embodiment 53, further comprising isolating the polypeptide.

**[0196]** 55. A composition comprising a physiologically acceptable carrier and a recombinant probiotic microorganism, wherein the recombinant probiotic microorganism comprises: a) a polynucleotide comprising an expression construct comprising a flagellin gene promoter operatively linked with a heterologous nucleotide sequence encoding a subject polypeptide; and b) a genetic modification of a FlgM gene that reduces inhibition of SigD initiation of transcription.

**[0197]** 56. The composition of embodiment 55, wherein the flagellin gene promoter comprises one or more genetic modifications that reduce CsrA repression of translation of an mRNA transcribed from the flagellin gene promoter.

**[0198]** 57. The composition of embodiment 55, wherein the subject polypeptide is an aldehyde dehydrogenase.

- [0199]** 58. The composition of embodiment 55, wherein the physiologically acceptable carrier is selected from lactic acid fermented foods, fermented dairy products, resistant starch, dietary fibers, carbohydrates, proteins and glycosylated proteins, water, capsule filler, and a gummy material.
- [0200]** 59. A unit dose of a composition of embodiment 55, comprising about 10<sup>4</sup> to about 10<sup>12</sup> colony forming units of the recombinant probiotic microorganisms.
- [0201]** 60. A method of metabolizing an analyte in a gut or in circulation in a subject comprising administering to the subject an effective amount of a composition comprising microorganisms that constitutively express an enzyme that metabolizes the analyte.
- [0202]** 61. A method of producing a target compound comprising: a) contacting a culture comprising a recombinant microorganism of embodiment 1 with an analyte, wherein the subject polypeptide is an enzyme for which the analyte is a substrate, and b) culturing the microorganism, wherein the enzyme catalyzes the conversion of the analyte into the target compound.
- [0203]** 62. The method of embodiment 63, wherein the enzyme is selected from an amylase, a lipase and a protease.

#### Bibliography

1. Sprince, H., et al., *Protective action of ascorbic acid and sulfur compounds against acetaldehyde toxicity: implications in alcoholism and smoking*. Agents Actions, 1975. **5**(2): p. 164-73.
2. Mukherjee, S. and D.B. Kearns, *The structure and regulation of flagellin Bacillus subtilis*. Annu Rev Genet, 2014. **48**: p. 319-40.
3. Guttenplan, S.B., S. Shaw, and D.B. Kearns, *The cell biology of peritrichous flagellin Bacillus subtilis*. Mol Microbiol, 2013. **87**(1): p. 211-29.
4. Macnab, R.M., *Genetics and biogenesis of bacterial flagella*. Annu Rev Genet, 1992. **26**: p. 131-58.
5. Caramori, T., et al., *Role of FlgM in sigma D-dependent gene expression in Bacillus subtilis*. J Bacteriol, 1996. **178**(11): p. 3113-8.
6. Yakhnin, H., et al., *CsrA of Bacillus subtilis regulates translation initiation of the gene encoding the flagellin protein (hag) by blocking ribosome binding*. Mol Microbiol, 2007. **64**(6): p. 1605-20.
7. Mukherjee, S., et al., *CsrA-FliW interaction governs flagellin homeostasis and a checkpoint on flagellar morphogenesis in Bacillus subtilis*. Mol Microbiol, 2011. **82**(2): p. 447-61.

8. Vakulskas, C.A., et al., *Regulation of bacterial virulence by Csr (Rsm) systems*. Microbiol Mol Biol Rev, 2015. **79**(2): p. 193-224.
9. Chen, R., et al., *Role of the sigmaD-dependent autolysins in Bacillus subtilis population heterogeneity*. J Bacteriol, 2009. **191**(18): p. 5775-84.
10. Ben-Yehuda, S., D.Z. Rudner, and R. Losick, *RacA, a bacterial protein that anchors chromosomes to the cell poles*. Science, 2003. **299**(5606): p. 532-6.
11. Gibson, D.G., et al., *Enzymatic assembly of DNA molecules up to several hundred kilobases*. Nat Methods, 2009. **6**(5): p. 343-5.
12. Jendrossek, D., A. Steinbuchel, and H.G. Schlegel, *Three different proteins exhibiting NAD-dependent acetaldehyde dehydrogenase activity from Alcaligenes eutrophus*. Eur J Biochem, 1987. **167**(3): p. 541-8.
13. Ho, K.K. and H. Weiner, *Isolation and characterization of an aldehyde dehydrogenase encoded by the aldB gene of Escherichia coli*. J Bacteriol, 2005. **187**(3): p. 1067-73.
14. Salaspuro, M., *Microbial metabolism of ethanol and acetaldehyde and clinical consequences*. Addict Biol, 1997. **2**(1): p. 35-46.

**[0204]** As used throughout this application, the word "may" is used in a permissive sense (i.e., meaning having the potential to), rather than the mandatory sense (i.e., meaning must). The words "include", "including", and "includes" and the like mean including, but not limited to. As used throughout this application, the singular forms "a," "an," and "the" include plural referents unless the content explicitly indicates otherwise. Thus, for example, reference to "an element" includes a combination of two or more elements, notwithstanding use of other terms and phrases for one or more elements, such as "one or more." The term "or" is, unless indicated otherwise, non-exclusive, i.e., encompassing both "and" and "or."

**[0205]** While embodiments of the present disclosure have been shown and described herein, it will be obvious to those skilled in the art that such embodiments are provided by way of example only. It is not intended that the disclosure be limited by the specific examples provided within the specification. While the disclosure has been described with reference to the aforementioned specification, the descriptions and illustrations of the embodiments herein are not meant to be construed in a limiting sense. Numerous variations, changes, and substitutions will now occur to those skilled in the art without departing from the disclosure. Furthermore, it shall be understood that all aspects of the disclosure are not limited to the specific depictions, configurations or relative proportions set forth herein which depend upon a variety of conditions and variables. It should be understood that various alternatives to the embodiments of the disclosure described herein may be employed in practicing the disclosure. It is therefore contemplated that the

disclosure shall also cover any such alternatives, modifications, variations or equivalents. It is intended that the following claims define the scope of the disclosure and that methods and structures within the scope of these claims and their equivalents be covered thereby.

**WHAT IS CLAIMED IS:**

1. A method for preventing or treating an alcohol hangover, comprising administering to the gut of a subject in need thereof an effective amount of a composition comprising a recombinant bacterium recombinantly engineered to constitutively express an aldehyde dehydrogenase.
2. The method of claim 1, wherein the recombinant bacteria comprises:
  - a) a polynucleotide comprising an expression construct comprising a flagellin gene promoter operatively linked with a heterologous nucleotide sequence encoding an aldehyde dehydrogenase, wherein the flagellin gene promoter comprises one or more genetic modifications that reduce CsrA repression of translation of an mRNA transcribed from the flagellin gene promoter; and
  - b) a genetic modification of a sigma factor repressor, wherein the modification reduces inhibition of sigma factor initiation of transcription from the flagellin gene promoter.
3. The method of claim 2, wherein the bacterium belongs to a genus selected from *Bacillus*, *Bifidobacterium*, *Enterococcus*, *Escherichia*, *Lactobacillus*, *Leuconostoc*, *Pediococcus*, *Pediococcus* and *Streptococcus*.
4. The method of claim 2, wherein (1) the bacterium belongs to genus *Bacillus*, (2) the flagellin gene promoter is a modified *Bacillus* hag promoter comprising one or more genetic modifications of the CsrA BS1 binding site and/or the CsrA BS2 binding site, (3) the sigma factor is SigD and (4) the SigD repressor is *FlgM*.
5. The method of claim 4, wherein the *hag* promoter comprises one or more genetic modifications in the 12-base-pair BS1 binding site or in the surrounding bases on either side of the binding site that form the stem of the stem-loop secondary structure of BS1.
6. The method of claim 4, wherein the *hag* promoter comprises a modification of the CsrA BS1 recognition sequence, AGGA.
7. The method of claim 4, wherein the *hag* promoter comprises the nucleotide sequence GCACAAGAACGT [SEQ ID NO:2].
8. The method of claim 4, wherein the one or more genetic modifications comprise one or more point mutations to the CsrA BS2 binding site.

9. The method of claim 4, wherein the *hag* promoter comprises one or more genetic modifications that disrupt the stem and loop structure of BS2 by eliminating complementarity that allows hydrogen bonding.
10. The method of claim 4, wherein the *hag* promoter comprises a modification of the BS2 binding site, wherein the modification comprises the nucleotide sequence ATTTAGGGAGGAA [SEQ ID NO:3].
11. The method of claim 1, wherein the aldehyde dehydrogenase is a bacterial aldehyde dehydrogenase.
12. The method of claim 4, wherein the aldehyde dehydrogenase is selected from (1) AcoD from *Cupriavidus necator* and (2) AldB from *E. coli*.
13. The method of claim 4, wherein the genetic modification in a *flgM* gene comprises deletion of all or part of the *flgM* gene.
14. The method of claim 4, wherein the genetic modification in a *flgM* gene disrupts secondary or tertiary structure.
15. The method of claim 4, wherein the expression construct is located in a bacterial chromosome.
16. The method of claim 1, comprising administering the composition to the subject during consumption of alcohol.
17. The method of claim 1, comprising administering the composition to the subject up to 24 hours before commencement of consumption of alcohol.
18. The method of claim 1, wherein the composition is administered orally.
19. The method of claim 1, wherein the composition further comprises a physiologically acceptable carrier selected from lactic acid fermented foods, fermented dairy products, resistant starch, dietary fibers, carbohydrates, proteins and glycosylated proteins, water, capsule filler, and a gummy material.
20. The method of claim 1, wherein the composition comprises about  $10^4$  to about  $10^{12}$  colony forming units of the recombinant bacteria.
21. A recombinant microorganism comprising:

a) a polynucleotide comprising an expression construct comprising a flagellin gene promoter operatively linked with a heterologous nucleotide sequence encoding a subject polypeptide, wherein the flagellin gene promoter comprises one or more genetic modifications that reduce CsrA repression of translation of an mRNA transcribed from the flagellin gene promoter; and

b) a genetic modification of a *flgM* gene that reduces inhibition of SigD initiation of transcription.

22. The recombinant microorganism of claim **21**, which constitutively expresses the polypeptide.

23. The recombinant microorganism of claim **21**, wherein the microorganism is probiotic.

24. The recombinant microorganism of claim **21**, wherein the microorganism belongs to genus selected from *Bacillus*, *Bifidobacterium*, *Enterococcus*, *Escherichia coli*, *Lactobacillus*, *Leuconostoc*, *Pediococcus*, *Pediococcus* and *Streptococcus*.

25. The recombinant microorganism of claim **21**, wherein the microorganism belongs to genus *Bacillus*.

26. The recombinant microorganism of claim **21**, wherein the microorganism is *B. subtilis*.

27. The recombinant microorganism of any of the foregoing claims, wherein the flagellin gene promoter is a *hag* promoter.

28. The recombinant microorganism of claim **27**, wherein the *hag* promoter comprises one or more genetic modifications that reduce CsrA repression of translation of an mRNA transcribed from the flagellin gene promoter, wherein the genetic modification comprises modification of the CsrA BS1 binding site and/or CsrA BS2 binding site (e.g., nucleotide substitution, insertion or deletion).

29. The recombinant microorganism of claim **28**, wherein the one or more genetic modifications comprise one or a plurality (e.g., two, three or four) of genetic modifications to the CsrA BS1 recognition sequence, AGGA, e.g., to the sequence AGAA.

30. The recombinant microorganism of claim **28**, wherein the one or more genetic modifications comprise one or more genetic modifications in the 12-base-pair BS1 binding

site or in the surrounding bases on either side of the binding site that form the stem of the stem-loop secondary structure of BS1.

31. The recombinant microorganism of claim **28**, wherein the one or more genetic modifications comprise one or a plurality (e.g., at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12) of genetic modifications in the BS1 binding site, gcacaaggacgt.

32. The recombinant microorganism of claim **28**, wherein the one or more genetic modifications disrupt the stem and loop structure of BS1 by eliminating complementarity that allows hydrogen bonding.

33. The recombinant microorganism of claim **28**, wherein the one or more genetic modifications comprise one or a plurality of genetic modifications in the sequence taagggcacaaggacgtgcctta [SEQ ID NO: 1] that are involved in hydrogen bonding, for example, to eliminate one, two, three, four or more hydrogen bond pairs.

34. The recombinant microorganism of claim **28**, wherein the modified BS1 has the nucleotide sequence GCACAAGAACGT [SEQ ID NO: 2].

35. The recombinant microorganism of claim **28**, wherein the one or more genetic modifications comprise one or more point mutations to the CsrA BS2 binding site.

36. The recombinant microorganism of claim **28**, wherein the one or more genetic modifications comprise one or a plurality (e.g., at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13) of genetic modification in the 13-base-pair BS2 binding site or in the surrounding bases on either side of the binding site that form the stem of the stem-loop secondary structure of BS2.

37. The recombinant microorganism of claim **28**, wherein the one or more genetic modifications disrupt the stem and loop structure of BS2 by eliminating complementarity that allows hydrogen bonding.

38. The recombinant microorganism of claim **28**, wherein the modified BS2 has the nucleotide sequence ATTTAGGGAGGAA [SEQ ID NO: 3].

39. The recombinant microorganism of claim **28**, wherein the one or more genetic modifications to the BS2 binding site does not include an alteration of nucleotides in the Shine-Dalgarno sequence aggagga.

40. The recombinant microorganism of any of the foregoing claims, wherein the flagellin gene promoter is located in a bacterial chromosome or in a plasmid.

41. The recombinant microorganism of any of the foregoing claims, wherein the subject polypeptide is an aldehyde dehydrogenase.

42. The recombinant microorganism of claim **41**, wherein the aldehyde dehydrogenase is AcoD from *Cupriavidus necator* and comprises an amino acid sequence identical to or substantially identical to:  
 MNMAEIAQLGVSNPYKQQYENYIGGAWVPPAGGEYFESTTTPITGKPFTRVPRSGQQDQVDA  
 ALDAAHAAKAAWARTSTTERANILNRIADRIEANLKLVAESIDNGKPVRETTAADLPLAVD  
 HFRYFAGCIRAQEGGISEIDADTIAYHFHEPLGVVGQIIPWNFPLLMATWKLAPALAAGNCVV  
 LKPAEQTPASILVLMVEIGDLLPPGVNVINGFGLGKPLASSPRISKVAFTGETTTGRLIM  
 QYASQNLIPVTLELGGKSPNIFEDVLAADDAFFDKALEGFAMFALNQGEVCTCPSRALIQE  
 SIYDRFMERALKRVAAIRQGHPLDTGTMIGAQAASAEQLEKILSYIDLGRKEGAQCLTGGERN  
 VLDGDLAGGYVVKPTVFAGHNKMRIFQEEIFGPVSVTTTFKDEEEALAIANDTLYGLGAGV  
 WTRDGARAFRMGRGIQAGRVWTNCYHAYPAHAAFGGYKQSGIGRENHRMMLDHYQQTK  
 NLLVSYSPNALGFF [SEQ ID NO: 4].

43. The recombinant microorganism of claim **41**, wherein the aldehyde dehydrogenase is a human aldehyde dehydrogenase, e.g., having an amino acid sequence identical to or substantially identical to  
 MLRAAARFGPRLGRRLLSAAATQAVPAPNQPEVFCNQIFINNEWHDAVSRKTFPTVNPS  
 TGEVICQVAEGDKEDVDKAVKAARAAFQLGSPWRRMDASHRGRLNRLADLIERDRTYLA  
 ALETLDNGKPYVISYLVLDLDMVLKCLRYAGWADKYHGKTIPIDGFFSYTRHEPVGVCG  
 QIIPWNFPLLMQAWKLGPALATGNVVMKVAEQTPLTALYVANLIKEAGFP PGVNVNIVPG  
 FGPTAGAAIASHEDVDKVAFTGSTIEGRVIQVAAGSSNLKRVTELELGGKSPNIIMSDADM  
 DWAVEQAHFALFFNQGCCAGSRTFVQEDIYDEFVERSVARAKSRVGNPFDSKTEQG  
 P  
 QVDETQFKKILGYINTGKQEGAKLLCGGGIAADRGYFIQPTVFGDVQDGMTIAKEEIFGP  
 VMQILKFKTIEEVGRANNSTYGLAAVFTKDLDKANYLSQALQAGTVWVNCYDVFGAQS  
 PFGGYKMSGSGRELGEYGLQAYTEVKTVTVKVPQKNS [SEQ ID NO: 5].

44. The recombinant microorganism of claim **21**, wherein the genetic modification in a *flgM* gene comprises deletion of all or part of the *flgM* gene.

45. The recombinant microorganism of claim **21**, wherein the genetic modification in a *flgM* gene comprises a single mutation or series of mutations in the sequence encoding active sites of *flgM*.

46. The recombinant microorganism of claim **21**, wherein the genetic modification in a *flgM* gene disrupts secondary or tertiary structure, such as in one of the helices that defines FlgM secondary structure.

47. The recombinant microorganism of claim **21**, wherein the genetic modification in a *flgM* gene comprises altering an amino acid in the 3<sup>rd</sup> or 4<sup>th</sup> helix at the C-terminal end of the FlgM protein, e.g., selected from I-58, K-62, I-65, G-68, D-73, A-78 of the *B. subtilis* FlgM.

48. The recombinant microorganism of claim **21**, wherein the genetic modification in a *flgM* gene comprises altering one or more amino acids predicted to participate in FlgM binding to SigD, e.g., selected from I-3, G-7, S-10, V-11, A-40, K-41, M43, I-58, L-61, K-62, I-65, Y-70, K-71, V-72, D-73, A-74, H-76, I-77, A-78, N-80, M-81, I-82, N-83, F-84, Y-85, and K-86 of the *B. subtilis* FlgM.

49. A recombinant probiotic microorganism that constitutively expresses an aldehyde dehydrogenase.

50. The recombinant probiotic microorganism of claim **49**, wherein the microorganism comprises:

- a) a polynucleotide comprising an expression construct comprising a flagellin gene promoter operatively linked with a heterologous nucleotide sequence encoding a subject polypeptide; and
- b) a genetic modification in a FlgM gene that reduces inhibition of SigD expression.

51. The recombinant probiotic microorganism of claim **50**, wherein the flagellin gene promoter comprises one or more genetic modifications that reduce CsrA repression of translation of an mRNA transcribed from the promoter.

52. A polynucleotide comprising an expression construct comprising a flagellin gene promoter operatively linked with a heterologous nucleotide sequence encoding a subject polypeptide, wherein the flagellin gene promoter comprises one or more genetic modifications that reduce CsrA repression of translation an mRNA transcribed from the flagellin gene promoter.

53. A method of making a polypeptide comprising culturing a recombinant microorganism of claim 21, or claim 49.

54. The method of claim 53, further comprising isolating the polypeptide.

55. A composition comprising a physiologically acceptable carrier and a recombinant probiotic microorganism, wherein the recombinant probiotic microorganism comprises:

a) a polynucleotide comprising an expression construct comprising a flagellin gene promoter operatively linked with a heterologous nucleotide sequence encoding a subject polypeptide; and

b) a genetic modification of a FlgM gene that reduces inhibition of SigD initiation of transcription.

56. The composition of claim 55, wherein the flagellin gene promoter comprises one or more genetic modifications that reduce CsrA repression of translation of an mRNA transcribed from the flagellin gene promoter.

57. The composition of claim 55, wherein the subject polypeptide is an aldehyde dehydrogenase.

58. The composition of claim 55, wherein the physiologically acceptable carrier is selected from lactic acid fermented foods, fermented dairy products, resistant starch, dietary fibers, carbohydrates, proteins and glycosylated proteins, water, capsule filler, and a gummy material.

59. A unit dose of a composition of claim 55, comprising about  $10^4$  to about  $10^{12}$  colony forming units of the recombinant probiotic microorganisms.

60. A method of metabolizing an analyte in a gut or in circulation in a subject comprising administering to the subject an effective amount of a composition comprising microorganisms that constitutively express an enzyme that metabolizes the analyte.

61. A method of producing a target compound comprising:

a) contacting a culture comprising a recombinant microorganism of claim 1 with an analyte, wherein the subject polypeptide is an enzyme for which the analyte is a substrate, and

b) culturing the microorganism, wherein the enzyme catalyzes the conversion of the analyte into the target compound.

62. The method of claim 61, wherein the enzyme is selected from an amylase, a lipase and a protease.