

ENGINEERED PROBIOTICS FOR EXPRESSION OF FIBER-SYNTHESIZING ENZYMES IN GUT

PRIORITY APPLICATION

[0001] The present application claims the benefit of U.S. Provisional Patent Application No. 63/089,334, filed on October 8, 2020, the content of which is hereby incorporated by reference herein in its entirety.

BACKGROUND

[0002] The microbiome is a complex collection of microbes that together impact many aspects of organismal health and function. For example, the microbiome can impact immunity, metabolism, and a variety of diseases. Microbes of the microbiome interact with host organisms through a variety of mechanisms, including microbial signaling, metabolite consumption, and small molecule production. Accordingly, the types and proportions of microbes and metabolites in the microbiome can have important impacts on health.

SUMMARY

[0003] Methods and compositions that modify the microbiome, e.g., by introducing new microbial functions or modifying existing microbial functions, can have important impacts on human health. The present disclosure provides solutions to various challenges in health and/or microbiome engineering. For example, the present disclosure recognizes that one challenge in human health is that many individuals consume daily diets that include carbohydrate such as monosaccharides and/or disaccharides, e.g., in excess of individual metabolic need. These consumed monosaccharides and/or disaccharides can have deleterious health effects in the gut (e.g., on the microbiome) and/or after digestion. To provide another example of a challenge in human health recognized by the present disclosure, many individuals consume daily diets that include too little fiber. Low consumption of fiber and/or low levels of gut fiber can have various deleterious health effects. The present disclosure provides solutions that simultaneously addresses harm caused by consumption of carbohydrates and harm caused by non-consumption of fiber.

[0004] The present disclosure provides, among other things, bacteria engineered for expression of a fiber-synthesizing enzyme, e.g., for expression of the enzyme in the gut (e.g., in intestine, e.g., in the small intestine and/or large intestine). In particular embodiments, the present disclosure provides bacteria engineered for expression of a fiber-synthesizing enzyme that consumes carbohydrate in the process of synthesizing fiber, e.g., in the gut. Those of skill in the art will appreciate that all references to gut herein include at least one or both of the small intestine and/or large intestine, as do references to intestine herein.

[0005] The present disclosure further provides, among other things, formulations for oral delivery of fiber-synthesizing enzymes, e.g., for delivery of fiber-synthesizing enzyme to gut (e.g., to intestine, e.g., to small intestine and/or large intestine). In various embodiments, the present disclosure provides a pill or sachet including isolated fiber-synthesizing enzyme.

[0006] In at least certain aspects, the present disclosure provides a method of decreasing the amount or concentration of a carbohydrate in the gut of a subject comprising administering to the subject a bacterium engineered to express a heterologous nucleic acid sequence encoding a fiber-synthesizing enzyme, wherein the fiber-synthesizing enzyme synthesizes fiber from a carbohydrate substrate.

[0007] In at least certain aspects, the present disclosure provides a method of increasing the amount or concentration of a fiber in the gut of a subject comprising administering to the subject a bacterium engineered to express a heterologous nucleic acid sequence encoding a fiber-synthesizing enzyme, wherein the fiber-synthesizing enzyme synthesizes fiber from a carbohydrate substrate.

[0008] In at least certain aspects, the present disclosure provides a method of treating a subject in need of decreased amount or concentration of a carbohydrate in the gut or increased amount or concentration of a fiber in the gut comprising administering to the subject a bacterium engineered to express a heterologous nucleic acid sequence encoding a fiber-synthesizing enzyme, wherein the fiber-synthesizing enzyme synthesizes fiber from a carbohydrate substrate.

[0009] In various embodiments, the subject is suffering from a condition positively correlated or associated with consumption of carbohydrate. In various embodiments, the subject is suffering from a condition negatively correlated or associated with consumption of fiber. In various embodiments, the subject is suffering from a condition selected from cardiovascular

disease, heart disease, high blood pressure, high blood cholesterol, high blood glucose, diabetes, obesity, dysbiosis of the gut, inflammatory bowel disease, irritable bowel syndrome (IBS), diverticulitis, colorectal cancer, intestinal cancer, bloating, cramping, gas, hemorrhoids, and diarrhea.

[0010] In various embodiments, the fiber is a soluble fiber. In various embodiments, the synthesized fiber comprises glucose, fructose, galacturonic acid, N-acetyl-D-glucosamine, and/or galactose monomers. In various embodiments, the carbohydrate substrate is selected from one or more of UDP-glucose, UDP-galactose, UDP-fucose, ADP- α -D-glucose, UDP-galacturonic acid, UDP-N-acetyl-alpha-D-glucosamine, galactinol, lactose, glucose, and sucrose and/or wherein the carbohydrate decreased in amount or concentration in the gut is selected from one or more of glucose, galactose, sucrose, fructose, and/or lactose. In various embodiments, the synthesized fiber is a fructooligosaccharide and/or is selected from laminaribiose, callose, curdlan, oat beta-glucan, laminarin, pleuran, lentinan, yeast beta glucan, trehalulose, trehalose, inulin, kestose, nystose, levan, raffinose, stachyose, verbascose, globotriose, human milk oligosaccharides (HMOs), cellobiose, cellulose, microcellulose, cotton, maltose, amylose, starch, glycogen, amylopectin, pectin, chitin. In various embodiments, the fiber-synthesizing enzyme is selected from a sucrase, an inulosucrase, a levansucrase, a 1,3-beta-glucan synthase, a 1,3;1,4-beta-D-glucan synthase, a 1,6-beta-glucan synthase, a sucrose isomerase, a 1,6-alpha-galactosyltransferase, a trehalulose synthase, a trehalose-6-phosphate synthase, an α -1,4-galactosyltransferase, an alpha-1,2-fucosyltransferase, a beta-galactosidase, a b-D-Galactosidase, a cellulose synthase, a maltose synthase, a starch synthase, a starch-branching enzyme, a glycogen synthase, a galacturonosyltransferase, and a chitinoligosaccharide synthase. In various embodiments, the fiber-synthesizing enzyme is operatively linked with a secretion polypeptide.

[0011] In various embodiments, the bacterium is a spore-forming bacterium and/or is in a spore form. In various embodiments, the bacterium is a probiotic bacterium. In various embodiments, the bacterium is of a genus selected from *Bacillus*, *Bifidobacterium*, *Enterococcus*, *Escherichia*, *Lactobacillus*, *Lactococcus*, *Leuconostoc*, *Pediococcus*, and *Streptococcus*, optionally wherein the bacterium is of the genus *Bacillus*, optionally wherein the bacterium is of the species *B. subtilis*. In various embodiments, the bacterium is of a strain characterized in that it does not colonize the gut.

[0012] In various embodiments, the method comprises administering the engineered bacterium to a subject that has consumed carbohydrate and/or the carbohydrate substrate within a preceding period, wherein the preceding period is a period of 24 hours 12 hours, 6 hours, 3 hours, or 1 hour. In various embodiments, the method comprises administering the engineered bacterium to a subject that has not consumed carbohydrate and/or the carbohydrate substrate. In various embodiments, the method comprises administering the engineered bacterium to a subject that has not consumed carbohydrate and/or the carbohydrate substrate within a preceding period, wherein the preceding period is a period of 24 hours 12 hours, 6 hours, 3 hours, or 1 hour. In various embodiments, the subject consumes carbohydrate and/or the carbohydrate substrate during a period subsequent to administration of the engineered bacterium, wherein the subsequent period is a period of 24 hours 12 hours, 6 hours, 3 hours, or 1 hour.

[0013] In various embodiments, the method prevents accumulation of sugar in the colon of the subject. In various embodiments, fiber-synthesizing enzymes and/or enzymes encoded by heterologous nucleic acid sequences comprised by the engineered bacterium consist of the fiber-synthesizing enzyme.

[0014] In various embodiments, the administration comprises oral administration of a composition comprising the engineered bacterium. In various embodiments, the administration comprises administration of about 10^4 to about 10^{12} colony forming units of the engineered bacterium.

[0015] In various embodiments, the nucleic acid sequence encoding the expression product is operatively linked with a constitutive promoter. In various embodiments, the nucleic acid sequence encoding the expression product is operatively linked with a flagellin gene promoter. In various embodiments, the flagellin gene promoter comprises a mutation in a CsrA binding site, wherein the mutation in the CsrA binding site inhibits binding of CsrA to mRNA transcripts encoding the fiber-synthesizing enzyme but does not preclude expression of the fiber-synthesizing enzyme. In various embodiments, the engineered bacterium comprises a mutation of an endogenous *flgM* gene that reduces inhibition of a sigma factor by FlgM. In various embodiments, the flagellin gene promoter is a *B. subtilis hag* promoter. In various embodiments, the mutation in the CsrA binding site is a mutation in a *B. subtilis hag* promoter CsrA binding site selected from binding site 1 (BS1) and binding site 2 (BS2). In various embodiments, the

mutation in the CsrA binding site is a mutation in the stem of the stem-loop secondary structure of BS1 or in the stem of the stem-loop secondary structure of BS2. In various embodiments, the mutation in the CsrA binding site is a mutation in the CsrA BS1 recognition sequence having the sequence AGGA. In various embodiments, the mutation in the CsrA binding site is a BS1 mutation according to SEQ ID NO: 29 or a BS2 mutation according to SEQ ID NO: 30. In various embodiments, the mutation in the CsrA binding site does not disrupt the Shine-Dalgarno sequence of the *B. subtilis* *hag* promoter. In various embodiments, the mutation of the endogenous *flgM* gene comprises deletion of all or a portion of the *flgM* gene. In various embodiments, the mutation of the endogenous *flgM* gene comprises a mutation in the sequence encoding the active site of *flgM*. In various embodiments, the mutation of the endogenous *flgM* gene comprises a mutation in the sequence encoding an amino acid that participates in binding of FlgM to SigD. In various embodiments, the mutation of the endogenous *flgM* gene alters a sequence that encodes an amino acid in the 3rd helix or 4th helix of FlgM at the C-terminal end of the FlgM protein. In various embodiments, the bacterium is of the species *B. subtilis* and the mutation of the endogenous *flgM* gene alters the amino acid sequence encoded by SEQ ID NO: 31 at an amino acid selected from I-58, K-62, I-65, G-68, D-73, and A-78. In various embodiments, the bacterium is of the species *B. subtilis* and the mutation of the endogenous *flgM* gene alters the amino acid sequence encoded by SEQ ID NO: 31 at an amino acid 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. In various embodiments, the mutation of the endogenous *flgM* gene reduces or eliminates FlgM biological activity. In various embodiments, the sigma factor is SigD.

[0016] In at least certain aspects, the present disclosure provides a bacterium engineered to express a heterologous nucleic acid sequence encoding a fiber-synthesizing enzyme, wherein the fiber-synthesizing enzyme synthesizes fiber from a carbohydrate substrate. In various embodiments, the fiber is a soluble fiber. In various embodiments, the synthesized fiber comprises glucose, fructose, galacturonic acid, N-acetyl-D-glucosamine, and/or galactose monomers. In various embodiments, the carbohydrate substrate is selected from one or more of UDP-glucose, UDP-galactose, UDP-fucose, ADP- α -D-glucose, UDP-galacturonic acid, UDP-N-acetyl-alpha-D-glucosamine, galactinol, lactose, glucose, and sucrose and/or wherein the

carbohydrate decreased in amount or concentration in the gut is selected from one or more of glucose, galactose, sucrose, fructose, and/or lactose. In various embodiments, the synthesized fiber is a fructooligosaccharide and/or is selected from laminaribiose, callose, curdlan, oat beta-glucan, laminarin, pleuran, lentinan, yeast beta glucan, trehalulose, trehalose, inulin, kestose, nystose, levan, raffinose, stachyose, verbascose, globotriose, human milk oligosaccharides (HMOs), cellobiose, cellulose, microcellulose, cotton, maltose, amylose, starch, glycogen, amylopectin, pectin, chitin. In various embodiments, the fiber-synthesizing enzyme is selected from a sucrase, an inulosucrase, a levansucrase, a 1,3-beta-glucan synthase, a 1,3;1,4-beta-D-glucan synthase, a 1,6-beta-glucan synthase, a sucrose isomerase, a 1,6-alpha-galactosyltransferase, a trehalulose synthase, a trehalose-6-phosphate synthase, an α -1,4-galactosyltransferase, an alpha-1,2-fucosyltransferase, a beta-galactosidase, a b-D-Galactosidase, a cellulose synthase, a maltose synthase, a starch synthase, a starch-branching enzyme, a glycogen synthase, a galacturonosyltransferase, and a chitinoligosaccharide synthase. In various embodiments, the fiber-synthesizing enzyme is operatively linked with a secretion polypeptide. In various embodiments, the bacterium is a spore-forming bacterium and/or is in a spore form. In various embodiments, the bacterium is a probiotic bacterium. In various embodiments, the bacterium is of a genus selected from *Bacillus*, *Bifidobacterium*, *Enterococcus*, *Escherichia*, *Lactobacillus*, *Lactococcus*, *Leuconostoc*, *Pediococcus*, and *Streptococcus*, optionally wherein the bacterium is of the genus *Bacillus*, optionally wherein the bacterium is of the species *B. subtilis*. In various embodiments, the bacterium is of a strain characterized in that it does not colonize the gut. In various embodiments, fiber-synthesizing enzymes and/or enzymes encoded by heterologous nucleic acid sequences comprised by the engineered bacterium consist of the fiber-synthesizing enzyme.

[0017] In various embodiments, the nucleic acid sequence encoding the expression product is operatively linked with a constitutive promoter. In various embodiments, the nucleic acid sequence encoding the expression product is operatively linked with a flagellin gene promoter. In various embodiments, the flagellin gene promoter comprises a mutation in a CsrA binding site, wherein the mutation in the CsrA binding site inhibits binding of CsrA to mRNA transcripts encoding the fiber-synthesizing enzyme but does not preclude expression of the fiber-synthesizing enzyme. In various embodiments, the engineered bacterium comprises a mutation

of an endogenous *flgM* gene that reduces inhibition of a sigma factor by FlgM. In various embodiments, the flagellin gene promoter is a *B. subtilis hag* promoter. In various embodiments, the mutation in the CsrA binding site is a mutation in a *B. subtilis hag* promoter CsrA binding site selected from binding site 1 (BS1) and binding site 2 (BS2). In various embodiments, the mutation in the CsrA binding site is a mutation in the stem of the stem-loop secondary structure of BS1 or in the stem of the stem-loop secondary structure of BS2. In various embodiments, the mutation in the CsrA binding site is a mutation in the CsrA BS1 recognition sequence having the sequence AGGA. In various embodiments, the mutation in the CsrA binding site is a BS1 mutation according to SEQ ID NO: 29 or a BS2 mutation according to SEQ ID NO: 30. In various embodiments, the mutation in the CsrA binding site does not disrupt the Shine-Dalgarno sequence of the *B. subtilis hag* promoter. In various embodiments, the mutation of the endogenous *flgM* gene comprises deletion of all or a portion of the *flgM* gene. In various embodiments, the mutation of the endogenous *flgM* gene comprises a mutation in the sequence encoding the active site of *flgM*. In various embodiments, the mutation of the endogenous *flgM* gene comprises a mutation in the sequence encoding an amino acid that participates in binding of FlgM to SigD. In various embodiments, the mutation of the endogenous *flgM* gene alters a sequence that encodes an amino acid in the 3rd helix or 4th helix of FlgM at the C-terminal end of the FlgM protein. In various embodiments, the bacterium is of the species *B. subtilis* and the mutation of the endogenous *flgM* gene alters the amino acid sequence encoded by SEQ ID NO: 31 at an amino acid selected from I-58, K-62, I-65, G-68, D-73, and A-78. In various embodiments, the bacterium is of the species *B. subtilis* and the mutation of the endogenous *flgM* gene alters the amino acid sequence encoded by SEQ ID NO: 31 at an amino acid 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. In various embodiments, the mutation of the endogenous *flgM* gene reduces or eliminates FlgM biological activity. In various embodiments, the sigma factor is SigD.

[0018] In at least certain aspects, the present disclosure provides a composition comprising the engineered bacterium of the present disclosure. In certain embodiments, the composition is formulated for oral administration. In some embodiments, the formulation comprises about 10^4 to about 10^{12} colony forming units of the engineered bacterium. In various

embodiments, the composition comprises a physiologically acceptable carrier. In various embodiments, the physiologically acceptable carrier is selected from a lactic acid fermented food, fermented dairy product, resistant starch, dietary fiber, carbohydrate, protein, glycosylated protein, water, capsule filler, and gummy material.

[0019] In at least certain aspects, the present disclosure provides a bacterial cell culture comprising the engineered bacterium of the present disclosure.

[0020] In at least certain aspects, the present disclosure provides a method of decreasing the amount or concentration of a carbohydrate in the gut of a subject comprising administering to the subject a composition comprising a fiber-synthesizing enzyme, wherein the fiber-synthesizing enzyme synthesizes fiber from a carbohydrate substrate. In various embodiments, the fiber synthesizing enzyme is an isolated fiber-synthesizing enzyme.

[0021] In at least certain aspects, the present disclosure provides a method of increasing the amount or concentration of a fiber in the gut of a subject comprising administering to the subject a composition comprising a fiber-synthesizing enzyme, wherein the fiber-synthesizing enzyme synthesizes fiber from a carbohydrate substrate. In various embodiments, the fiber synthesizing enzyme is an isolated fiber-synthesizing enzyme.

[0022] In at least certain aspects, the present disclosure provides a method of treating a subject in need of decreased amount or concentration of a carbohydrate in the gut or increased amount or concentration of a fiber in the gut comprising administering to the subject a composition comprising a fiber-synthesizing enzyme, wherein the fiber-synthesizing enzyme synthesizes fiber from a carbohydrate substrate. In various embodiments, the fiber synthesizing enzyme is an isolated fiber-synthesizing enzyme.

[0023] In various embodiments, the subject is suffering from a condition positively correlated or associated with consumption of carbohydrate. In various embodiments, the subject is suffering from a condition negatively correlated or associated with consumption of fiber. In various embodiments, the subject is suffering from a condition selected from cardiovascular disease, heart disease, high blood pressure, high blood cholesterol, high blood glucose, diabetes, obesity, dysbiosis of the gut, inflammatory bowel disease, irritable bowel syndrome (IBS), diverticulitis, colorectal cancer, intestinal cancer, bloating, cramping, gas, hemorrhoids, and diarrhea.

[0024] In various embodiments, the fiber is a soluble fiber. In various embodiments, the synthesized fiber comprises glucose, fructose, galacturonic acid, N-acetyl-D-glucosamine, and/or galactose monomers. In various embodiments, the carbohydrate substrate is selected from one or more of UDP-glucose, UDP-galactose, UDP-fucose, ADP- α -D-glucose, UDP-galacturonic acid, UDP-N-acetyl- α -D-glucosamine, galactinol, lactose, glucose, and sucrose and/or wherein the carbohydrate decreased in amount or concentration in the gut is selected from one or more of glucose, galactose, sucrose, fructose, and/or lactose. In various embodiments, the synthesized fiber is a fructooligosaccharide and/or is selected from laminaribiose, callose, curdlan, oat beta-glucan, laminarin, pleuran, lentinan, yeast beta glucan, trehalulose, trehalose, inulin, kestose, nystose, levan, raffinose, stachyose, verbascose, globotriose, human milk oligosaccharides (HMOs), cellobiose, cellulose, microcellulose, cotton, maltose, amylose, starch, glycogen, amylopectin, pectin, chitin. In various embodiments, the fiber-synthesizing enzyme is selected from a sucrase, an inulosucrase, a levansucrase, a 1,3-beta-glucan synthase, a 1,3;1,4-beta-D-glucan synthase, a 1,6-beta-glucan synthase, a sucrose isomerase, a 1,6-alpha-galactosyltransferase, a trehalulose synthase, a trehalose-6-phosphate synthase, an α -1,4-galactosyltransferase, an α -1,2-fucosyltransferase, a beta-galactosidase, a b-D-Galactosidase, a cellulose synthase, a maltose synthase, a starch synthase, a starch-branching enzyme, a glycogen synthase, a galacturonosyltransferase, and a chitinoligosaccharide synthase.

[0025] In various embodiments, the method comprises administering the composition to a subject that has consumed carbohydrate and/or the carbohydrate substrate within a preceding period, wherein the preceding period is a period of 24 hours 12 hours, 6 hours, 3 hours, or 1 hour. In various embodiments, the method comprises administering the composition to a subject that has not consumed carbohydrate and/or the carbohydrate substrate. In various embodiments, the method comprises administering the composition to a subject that has not consumed carbohydrate and/or the carbohydrate substrate within a preceding period, wherein the preceding period is a period of 24 hours 12 hours, 6 hours, 3 hours, or 1 hour. In various embodiments, the subject consumes carbohydrate and/or the carbohydrate substrate during a period subsequent to administration of the composition, wherein the subsequent period is a period of 24 hours 12 hours, 6 hours, 3 hours, or 1 hour.

[0026] In various embodiments, the method prevents accumulation of sugar in the colon of the subject. In various embodiments, fiber-synthesizing enzymes comprising the composition and/or enzymes comprising the composition consist of the fiber-synthesizing enzyme. In various embodiments, the administration comprises oral administration of the composition.

[0027] In at least certain aspects, the present disclosure provides a composition comprising a fiber synthesizing enzyme, where the fiber-synthesizing enzyme synthesizes fiber from a carbohydrate substrate and the composition is formulated for oral administration. In various embodiments, the fiber is a soluble fiber. In various embodiments, the synthesized fiber comprises glucose, fructose, galacturonic acid, N-acetyl-D-glucosamine, and/or galactose monomers. In various embodiments, the carbohydrate substrate is selected from one or more of UDP-glucose, UDP-galactose, UDP-fucose, ADP- α -D-glucose, UDP-galacturonic acid, UDP-N-acetyl- α -D-glucosamine, galactinol, lactose, glucose, and sucrose and/or wherein the carbohydrate decreased in amount or concentration in the gut is selected from one or more of glucose, galactose, sucrose, fructose, and/or lactose. In various embodiments, the synthesized fiber is a fructooligosaccharide and/or is selected from laminaribiose, callose, curdlan, oat beta-glucan, laminarin, pleuran, lentinan, yeast beta glucan, trehalulose, trehalose, inulin, kestose, nystose, levan, raffinose, stachyose, verbascose, globotriose, human milk oligosaccharides (HMOs), cellobiose, cellulose, microcellulose, cotton, maltose, amylose, starch, glycogen, amylopectin, pectin, chitin. In various embodiments, the fiber-synthesizing enzyme is selected from a sucrase, an inulosucrase, a levansucrase, a 1,3-beta-glucan synthase, a 1,3;1,4-beta-D-glucan synthase, a 1,6-beta-glucan synthase, a sucrose isomerase, a 1,6-alpha-galactosyltransferase, a trehalulose synthase, a trehalose-6-phosphate synthase, an α -1,4-galactosyltransferase, an α -1,2-fucosyltransferase, a beta-galactosidase, a b-D-Galactosidase, a cellulose synthase, a maltose synthase, a starch synthase, a starch-branching enzyme, a glycogen synthase, a galacturonosyltransferase, and a chitinoligosaccharide synthase. In various embodiments, fiber-synthesizing enzymes comprising the composition and/or enzymes comprising the composition consist of the fiber-synthesizing enzyme. In various embodiments, the composition comprises a physiologically acceptable carrier. In various embodiments, the physiologically acceptable carrier is selected from a lactic acid fermented food, fermented dairy

product, resistant starch, dietary fiber, carbohydrate, protein, glycosylated protein, water, capsule filler, and gummy material.

DEFINITIONS

[0028] *A, An, The:* As used herein, “a”, “an”, and “the” refer to one or to more than one (*i.e.*, to at least one) of the grammatical object of the article. By way of example, “an element” discloses embodiments of exactly one element and embodiments including more than one element.

[0029] *About:* As used herein, term “about”, when used in reference to a value, refers to a value that is similar, in context to the referenced value. In general, those skilled in the art, familiar with the context, will appreciate the relevant degree of variance encompassed by “about” in that context. For example, in some embodiments, the term “about” may encompass a range of values that within 25%, 20%, 19%, 18%, 17%, 16%, 15%, 14%, 13%, 12%, 11%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, or less of the referenced value.

[0030] *Administration:* As used herein, the term “administration” typically refers to administration of a composition to a subject or system to achieve delivery of an agent that is, or is included in, the composition.

[0031] *Associated with:* Two events or entities are “associated” with one another, as that term is used herein, if the presence, level and/or form of one is correlated with that of the other. For example, a particular entity (e.g., polypeptide, genetic signature, metabolite, microbe, etc.) is considered to be associated with a particular disease, disorder, or condition, if its presence, level and/or form correlates with incidence of and/or susceptibility to the disease, disorder, or condition (e.g., across a relevant population). In some embodiments, two or more entities are physically “associated” with one another if they interact, directly or indirectly, so that they are and/or remain in physical proximity with one another. In some embodiments, two or more entities that are physically associated with one another are covalently linked to one another; in some embodiments, two or more entities that are physically associated with one another are not covalently linked to one another but are non-covalently associated, for example by means of hydrogen bonds, van der Waals interaction, hydrophobic interactions, magnetism, and combinations thereof.

[0032] *Between or From:* As used herein, the term “between” refers to content that falls between indicated upper and lower, or first and second, boundaries, inclusive of the boundaries. Similarly, the term “from”, when used in the context of a range of values, indicates that the range includes content that falls between indicated upper and lower, or first and second, boundaries, inclusive of the boundaries.

[0033] *Binding:* As used herein, the term “binding” refers to a non-covalent association between or among two or more agents. “Direct” binding involves physical contact between agents; indirect binding involves physical interaction by way of physical contact with one or more intermediate agents. Binding between two or more agents can occur and/or be assessed in any of a variety of contexts, including where interacting agents are studied in isolation or in the context of more complex systems (e.g., while covalently or otherwise associated with a carrier agents and/or in a biological system or cell).

[0034] *Control expression or activity:* As used herein, a first element (e.g., a protein, such as a transcription factor, or a nucleic acid sequence, such as promoter) “controls” or “drives” expression or activity of a second element (e.g., a protein or a nucleic acid encoding an agent such as a protein) if the expression or activity of the second element is wholly or partially dependent upon status (e.g., presence, absence, conformation, chemical modification, interaction, or other activity) of the first under at least one set of conditions. Control of expression or activity can be substantial control or activity, e.g., in that a change in status of the first element can, under at least one set of conditions, result in a change in expression or activity of the second element of at least 10% (e.g., at least 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 2-fold, 3-fold, 4-fold, 5-fold, 10-fold, 20-fold, 30-fold, 40-fold, 50-fold, 100-fold) as compared to a reference control.

[0035] *Corresponding to:* As used herein, the term “corresponding to” may be used to designate the position/identity of a structural element in a compound or composition through comparison with an appropriate reference compound or composition. For example, in some embodiments, a monomeric residue in a polymer (e.g., an amino acid residue in a polypeptide or a nucleic acid residue in a polynucleotide) may be identified as “corresponding to” a residue in an appropriate reference polymer. For example, those of skill in the art appreciate that residues in a provided polypeptide or polynucleotide sequence are often designated (e.g., numbered or

labeled) according to the scheme of a related reference sequence (even if, e.g., such designation does not reflect literal numbering of the provided sequence). By way of illustration, if a reference sequence includes a particular amino acid motif at positions 100-110, and a second related sequence includes the same motif at positions 110-120, the motif positions of the second related sequence can be said to “correspond to” positions 100-110 of the reference sequence. Those of skill in the art appreciate that corresponding positions can be readily identified, e.g., by alignment of sequences, and that such alignment is commonly accomplished by any of a variety of known tools, strategies, and/or algorithms, including without limitation software programs such as, for example, BLAST, CS-BLAST, CUDASW++, DIAMOND, FASTA, GGSEARCH/GLSEARCH, Genoogle, HMMER, HHpred/HHsearch, IDF, Infernal, KLAST, USEARCH, parasail, PSI-BLAST, PSI-Search, ScalaBLAST, Sequilab, SAM, SSEARCH, SWAPHI, SWAPHI-LS, SWIMM, or SWIPE.

[0036] *Dosage form or unit dosage form:* Those skilled in the art will appreciate that the term “dosage form” may be used to refer to a physically discrete unit of an agent (e.g., a therapeutic or diagnostic agent) for administration to a subject. Typically, each such unit contains a predetermined quantity of agent. In some embodiments, such quantity is a unit dosage amount (or a whole fraction thereof) appropriate for administration in accordance with a dosing regimen that has been determined to correlate with a desired or beneficial outcome when administered to a relevant population (i.e., with a therapeutic dosing regimen). Those of ordinary skill in the art appreciate that the total or free amount of a therapeutic composition or agent administered to a particular subject is determined by one or more attending physicians and may involve administration of multiple dosage forms.

[0037] *Dosing regimen:* As used herein, the term “dosing regimen” can refer to a set of one or more same or different unit doses administered to a subject, typically including a plurality of unit doses administration of each of which is separated from administration of the others by a period of time. In various embodiments, one or more or all unit doses of a dosing regimen may be the same or can vary (e.g., increase over time, decrease over time, or be adjusted in accordance with the subject and/or with a medical practitioner’s determination). In various embodiments, one or more or all of the periods of time between each dose may be the same or can vary (e.g., increase over time, decrease over time, or be adjusted in accordance with the

subject and/or with a medical practitioner's determination). In some embodiments, a given therapeutic agent has a recommended dosing regimen, which can involve one or more doses. Typically, at least one recommended dosing regimen of a marketed drug is known to those of skill in the art. In some embodiments, a dosing regimen is correlated with a desired or beneficial outcome when administered across a relevant population (i.e., is a therapeutic dosing regimen).

[0038] *Engineered:* As used herein, the term “engineered” refers to the aspect of having been manipulated by the hand of man. For example, a polynucleotide is considered to be “engineered” when two or more sequences, that are not linked together in that order in nature, are manipulated by the hand of man to be linked to one another in the engineered polynucleotide. Those of skill in the art will appreciate that an “engineered” nucleic acid or amino acid sequence can be a recombinant nucleic acid or amino acid sequence. In some embodiments, an engineered polynucleotide includes a coding sequence and/or a regulatory sequence that is found in nature operably linked with a first sequence but is not found in nature operably linked with a second sequence, which is in the engineered polynucleotide and operably linked in with the second sequence by the hand of man. In some embodiments, a cell or organism is considered to be “engineered” if it has been manipulated so that its genetic information is altered (e.g., new genetic material not previously present has been introduced, for example by transformation, mating, somatic hybridization, transfection, transduction, or other mechanism, or previously present genetic material is altered or removed, for example by substitution, deletion, or mating). As is common practice and is understood by those of skill in the art, progeny or copies, perfect or imperfect, of an engineered polynucleotide or cell are typically still referred to as “engineered” even though the direct manipulation was of a prior entity.

[0039] *Excipient:* As used herein, “excipient” refers to a non-therapeutic agent that may be included in a pharmaceutical composition, for example to provide or contribute to a desired consistency or stabilizing effect. In some embodiments, suitable pharmaceutical excipients may include, for example, starch, glucose, lactose, sucrose, gelatin, malt, rice, flour, chalk, silica gel, sodium stearate, glycerol monostearate, talc, sodium chloride, dried skim milk, glycerol, propylene, glycol, water, ethanol, or the like.

[0040] *Expression:* As used herein, “expression” refers individually and/or cumulatively to one or more biological process that result in production from a nucleic acid sequence of an

encoded agent, such as a polypeptide. Expression specifically includes either or both of transcription and translation.

[0041] *Fragment:* As used herein, “fragment” refers a structure that is or includes a discrete portion of a reference agent (sometimes referred to as the “parent” agent). In some embodiments, a fragment lacks one or more moieties found in the reference agent. In some embodiments, a fragment is or includes one or more moieties found in the reference agent. In some embodiments, the reference agent is a polymer such as a polynucleotide or polypeptide. In some embodiments, a fragment of a polymer is or includes at least 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500 or more monomeric units (e.g., residues) of the reference polymer. In some embodiments, a fragment of a polymer is or includes at least about 5%, 10%, 15%, 20%, 25%, 30%, 25%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more of the monomeric units (e.g., residues) found in the reference polymer. A fragment of a reference polymer is not necessarily identical to a corresponding portion of the reference polymer. For example, a fragment of a reference polymer can be a polymer having a sequence of residues having at least about 5%, 10%, 15%, 20%, 25%, 30%, 25%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identity to the reference polymer. A fragment may, or may not, be generated by physical fragmentation of a reference agent. In some instances a fragment is generated by physical fragmentation of a reference agent. In some instances, a fragment is not generated by physical fragmentation of a reference agent and can be instead, for example, produced by *de novo* synthesis or other means.

[0042] *Gene or Transgene:* As used herein, the term “gene” refers to a DNA sequence that is or includes coding sequence (i.e., a DNA sequence that encodes an expression product, such as an RNA product and/or a polypeptide product), optionally together with some or all of regulatory sequences that control expression of the coding sequence. In some embodiments, a gene includes non-coding sequence such as, without limitation, introns. In some embodiments, a gene may include both coding (e.g., exonic) and non-coding (e.g., intronic) sequences. In some embodiments, a gene includes a regulatory sequence that is a promoter. In some embodiments, a gene includes one or both of a (i) DNA nucleotides extending a predetermined number of

nucleotides upstream of the coding sequence in a reference context, such as a source genome, and (ii) DNA nucleotides extending a predetermined number of nucleotides downstream of the coding sequence in a reference context, such as a source genome. In various embodiments, the predetermined number of nucleotides can be 500 bp, 1 kb, 2 kb, 3 kb, 4 kb, 5 kb, 10 kb, 20 kb, 30 kb, 40 kb, 50 kb, 75 kb, or 100 kb. As used herein, a “transgene” refers to a gene that is not endogenous or native to a reference context in which the gene is present or into which the gene may be placed by engineering.

[0043] *Gene product or expression product:* As used herein, the term “gene product” or “expression product” generally refers to an RNA transcribed from the gene (pre-and/or post-processing) or a polypeptide (pre- and/or post-modification) encoded by an RNA transcribed from the gene.

[0044] *Heterologous:* As used herein, a first nucleic acid sequence is “heterologous” to a second nucleic acid sequence if the first nucleic acid sequence is not operatively linked with the second nucleic acid sequence in nature. By extension, a polypeptide is “heterologous” to an expression control sequence if it is encoded by nucleic acid sequence heterologous the promoter.

[0045] *Identity:* As used herein, the term “identity” refers to the overall relatedness between polymeric molecules, *e.g.*, between nucleic acid molecules (*e.g.*, DNA molecules and/or RNA molecules) and/or between polypeptide molecules. Methods for the calculation of a percent identity as between two provided sequences are known in the art. Calculation of the percent identity of two nucleic acid or polypeptide sequences, for example, can be performed by aligning the two sequences (or the complement of one or both sequences) for optimal comparison purposes (*e.g.*, gaps can be introduced in one or both of a first and a second sequences for optimal alignment and non-identical sequences can be disregarded for comparison purposes). The nucleotides or amino acids at corresponding positions are then compared. When a position in the first sequence is occupied by the same residue (*e.g.*, nucleotide or amino acid) as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, optionally taking into account the number of gaps, and the length of each gap, which may need to be introduced for optimal alignment of the two sequences. The comparison of sequences and determination of percent identity between two sequences can

be accomplished using a computational algorithm, such as BLAST (basic local alignment search tool).

[0046] *“Improve,” “increase,” “inhibit,” or “reduce”*: As used herein, the terms “improve”, “increase”, “inhibit”, and “reduce”, and grammatical equivalents thereof, indicate qualitative or quantitative difference from a reference.

[0047] *Isolated*: As used herein, “isolated” or “purified” can refer to a substance and/or entity that has been (a) separated from at least some of the components with which it was associated when initially produced (whether in nature and/or in an experimental setting), and/or (b) designed, produced, prepared, and/or manufactured by the hand of man. Isolated substances and/or entities may be separated from at least about 10%, about 20%, about 30%, about 40%, about 50%, about 60%, about 70%, about 80%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, about 99%, or more than about 99% of the other components with which they were initially associated. In some embodiments, isolated substances and/or entities are at least about 80%, about 85%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, about 99%, or more than about 99% pure. As used herein, a substance and/or entity is “pure” if it is substantially free of other components. In some embodiments, as will be understood by those skilled in the art, a substance and/or entity may still be considered “isolated” or “pure” after having been combined with certain other components such as, for example, one or more carriers or excipients (e.g., buffer, solvent, water, etc.); in such embodiments, percent isolation or purity of the substance and/or entity is calculated without including such carriers or excipients. To give but one example, in some embodiments, a biological polymer such as a polypeptide or polynucleotide that occurs in nature can be referred to as “isolated” when, (a) by virtue of its origin or source of derivation is not associated with some or all of the components with which it was associated in its native state in nature; (b) it is substantially free of other polypeptides or nucleic acids of the same species from the species that produces it in nature; (c) is expressed by or is otherwise in association with components from a cell or other expression system that is not of the species that produces it in nature. Thus, for instance, in some embodiments, a polypeptide that is chemically synthesized or is synthesized in a cellular system different from that which produces it in nature is considered to be an “*isolated*” polypeptide. Alternatively or additionally,

in some embodiments, a polypeptide that has been subjected to one or more purification techniques may be considered to be an “*isolated*” polypeptide to the extent that it has been separated from other components (a) with which it is associated in nature; and/or (b) with which it was associated when initially produced.

[0048] *Nucleic acid:* As used herein, in its broadest sense, the term “nucleic acid” refers to any compound and/or substance that is or can be incorporated into an oligonucleotide chain. In some embodiments, a nucleic acid is a compound and/or substance that is or can be incorporated into an oligonucleotide chain via a phosphodiester linkage. As will be clear from context, in some embodiments, the term nucleic acid refers to an individual nucleic acid residue (e.g., a nucleotide and/or nucleoside), and in some embodiments refers to an polynucleotide chain including a plurality of individual nucleic acid residues. A nucleic acid can be or include DNA, RNA, or a combinations thereof. A nucleic acid can include natural nucleic acid residues, nucleic acid analogs, and/or synthetic residues. In some embodiments, a nucleic acid includes natural nucleotides (e.g., adenosine, thymidine, guanosine, cytidine, uridine, deoxyadenosine, deoxythymidine, deoxy guanosine, and deoxycytidine). In some embodiments, a nucleic acid is or includes of one or more nucleotide analogs (e.g., 2-aminoadenosine, 2-thiothymidine, inosine, pyrrolo-pyrimidine, 3 -methyl adenosine, 5-methylcytidine, C-5 propynyl-cytidine, C-5 propynyl-uridine, 2-aminoadenosine, C5-bromouridine, C5-fluorouridine, C5-iodouridine, C5-propynyl-uridine, C5 -propynyl-cytidine, C5-methylcytidine, 2-aminoadenosine, 7-deazaadenosine, 7-deazaguanosine, 8-oxoadenosine, 8-oxoguanosine, 0(6)-methylguanine, 2-thiocytidine, methylated bases, intercalated bases, and combinations thereof). In some embodiments, a nucleic acid has a nucleotide sequence that encodes a functional gene product such as an RNA or protein. In some embodiments, a nucleic acid includes one or more introns. In some embodiments, a nucleic acid includes one or more genes. In some embodiments, nucleic acids are prepared by one or more of isolation from a natural source, enzymatic synthesis by polymerization based on a complementary template (*in vivo* or *in vitro*), reproduction in a recombinant cell or system, and chemical synthesis. In some embodiments, a nucleic acid analog differs from a nucleic acid in that it does not utilize a phosphodiester backbone. For example, in some embodiments, a nucleic acid can include one or more peptide nucleic acids, which are known in the art and have peptide bonds instead of phosphodiester bonds in the

backbone. Alternatively or additionally, in some embodiments, a nucleic acid has one or more phosphorothioate and/or 5'-N-phosphoramidite linkages rather than phosphodiester bonds. In some embodiments, a nucleic acid includes one or more modified sugars (e.g., 2'-fluororibose, ribose, 2'-deoxyribose, arabinose, and hexose) as compared with those in natural nucleic acids. In some embodiments, a nucleic acid is or includes at least 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 600, 700, 800, 900, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000 or more residues. In some embodiments, a nucleic acid is partly or wholly single stranded, or partly or wholly double stranded. In some embodiments a nucleic acid has a sequence including at least one element that encodes, or is the complement of a sequence that encodes, a polypeptide. In some embodiments, a nucleic acid has enzymatic activity.

[0049] *Operably linked:* As used herein, “operably linked” refers to the association of at least a first element and a second element such that the component elements are in a relationship permitting them to function in their intended manner. For example, a nucleic acid regulatory sequence is “*operably linked*” to a nucleic acid coding sequence if the regulatory sequence and coding sequence are associated in a manner that permits control of expression of the coding sequence by the regulatory sequence. In some embodiments, an “*operably linked*” regulatory sequence is directly or indirectly covalently associated with a coding sequence (e.g., in a single nucleic acid). In some embodiments, a regulatory sequence controls expression of a coding sequence in *trans* and inclusion of the regulatory sequence in the same nucleic acid as the coding sequence is not a requirement of operable linkage.

[0050] *Pharmaceutically acceptable:* As used herein, the term “pharmaceutically acceptable,” as applied to one or more, or all, component(s) for formulation of a composition as disclosed herein, means that each component must be compatible with the other ingredients of the composition and not deleterious to the recipient thereof.

[0051] *Pharmaceutically acceptable carrier:* As used herein, the term “pharmaceutically acceptable carrier” refers to a pharmaceutically-acceptable material, composition, or vehicle, such as a liquid or solid filler, diluent, excipient, or solvent encapsulating material, that facilitates formulation of an agent (e.g., a pharmaceutical agent), modifies bioavailability of an agent, or

facilitates transport of an agent from one organ or portion of a subject to another. Some examples of materials which can serve as pharmaceutically-acceptable carriers include: sugars, such as lactose, glucose and sucrose; starches, such as corn starch and potato starch; cellulose, and its derivatives, such as sodium carboxymethyl cellulose, ethyl cellulose and cellulose acetate; powdered tragacanth; malt; gelatin; talc; excipients, such as cocoa butter and suppository waxes; oils, such as peanut oil, cottonseed oil, safflower oil, sesame oil, olive oil, corn oil and soybean oil; glycols, such as propylene glycol; polyols, such as glycerin, sorbitol, mannitol and polyethylene glycol; esters, such as ethyl oleate and ethyl laurate; agar; buffering agents, such as magnesium hydroxide and aluminum hydroxide; alginic acid; pyrogen-free water; isotonic saline; Ringer's solution; ethyl alcohol; pH buffered solutions; polyesters, polycarbonates and/or polyanhydrides; and other non-toxic compatible substances employed in pharmaceutical formulations.

[0052] *Pharmaceutical composition:* As used herein, the term “pharmaceutical composition” refers to a composition in which a therapeutic agent is formulated together with one or more pharmaceutically acceptable carriers.

[0053] *Polypeptide:* As used herein, “polypeptide” refers to any polymeric chain of amino acids. In some embodiments, a polypeptide has an amino acid sequence that occurs in nature. In some embodiments, a polypeptide has an amino acid sequence that does not occur in nature. In some embodiments, a polypeptide has an amino acid sequence that is engineered in that it is designed and/or produced through action of the hand of man. In some embodiments, a polypeptide may be or include of natural amino acids, non-natural amino acids, or both. In some embodiments, a polypeptide may be or include only natural amino acids or only non-natural amino acids. In some embodiments, a polypeptide can include D-amino acids, L-amino acids, or both. In some embodiments, a polypeptide may include only L-amino acids. In some embodiments, a polypeptide may include one or more pendant groups or other modifications, e.g., one or more amino acid side chains, e.g., at the polypeptide's N-terminus, at the polypeptide's C-terminus, at non-terminal amino acids, or at any combination thereof. In some embodiments, such pendant groups or modifications may be selected from acetylation, amidation, lipidation, methylation, phosphorylation, glycosylation, glycation, sulfation, mannosylation, nitrosylation, acylation, palmitoylation, prenylation, pegylation, etc., including

combinations thereof. In some embodiments, a polypeptide may be cyclic, and/or may include a cyclic portion.

[0054] In some embodiments, the term “polypeptide” may be appended to a name of a reference polypeptide, activity, or structure to indicate a class of polypeptides that share a relevant activity or structure. For such classes, the present specification provides and/or those skilled in the art will be aware of exemplary polypeptides within the class whose amino acid sequences and/or functions are known. In some embodiments, a member of a polypeptide class or family shows significant sequence homology or identity with, shares a common sequence motif (e.g., a characteristic sequence element) with, and/or shares a common activity (in some embodiments at a comparable level or within a designated range) with a reference polypeptide of the class. For example, in some embodiments, a member polypeptide shows an overall degree of sequence homology or identity with a reference polypeptide that is at least about 30-40%, and is often greater than about 50%, 60%, 70%, 80%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more and/or includes at least one region (e.g., a conserved region that can in some embodiments be or include a characteristic sequence element) that shows very high sequence identity, often greater than 90% or even 95%, 96%, 97%, 98%, or 99%. Such a conserved region usually encompasses at least 3-4 and in some instances up to 20 or more amino acids; in some embodiments, a conserved region encompasses at least one stretch of at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 or more contiguous amino acids. In some embodiments, a relevant polypeptide can be or include a fragment of a parent polypeptide. In some embodiments, a useful polypeptide may be or include a plurality of fragments, each of which is found in the same parent polypeptide in a different spatial arrangement relative to one another than is found in the polypeptide of interest (e.g., fragments that are directly linked in the parent may be spatially separated in the polypeptide of interest or vice versa, and/or fragments may be present in a different order in the polypeptide of interest than in the parent), so that the polypeptide of interest is a derivative of its parent polypeptide.

[0055] *Prevent or prevention:* The terms “prevent” and “prevention,” as used herein in connection with the occurrence of a disease, disorder, or condition, refers to reducing the risk of developing the disease, disorder, or condition; delaying onset of the disease, disorder, or condition; delaying onset of one or more characteristics or symptoms of the disease, disorder, or

condition; and/or to reducing the frequency and/or severity of one or more characteristics or symptoms of the disease, disorder, or condition. Prevention can refer to prevention in a particular subject or to a statistical impact on a population of subjects. Prevention can be considered to have occurred when onset of a disease, disorder, or condition has been delayed for a period of time that is predefined or understood by those of skill in the art.

[0056] *Promoter:* As used herein, a “promoter” or “promoter sequence” can be a DNA regulatory region that directly or indirectly (e.g., through promoter-bound proteins or substances) participates in initiation and/or processivity of transcription of a coding sequence. A promoter may, under suitable conditions, initiate transcription of a coding sequence upon binding of one or more transcription factors and/or regulatory moieties with the promoter. A promoter that participates in initiation of transcription of a coding sequence can be “operably linked” to the coding sequence. In certain instances, a promoter can be or include a DNA regulatory region that extends from a transcription initiation site (at its 3’ terminus) to an upstream (5’ direction) position such that the sequence so designated includes one or both of a minimum number of bases or elements necessary to initiate a transcription event. A promoter may be, include, or be operably associated with or operably linked to, expression control sequences such as enhancer and repressor sequences.

[0057] *Reference:* As used herein, “reference” refers to a standard or control relative to which a comparison is performed. For example, in some embodiments, an agent, sample, sequence, subject, animal, or individual, or population thereof, or a measure or characteristic representative thereof, is compared with a reference, an agent, sample, sequence, subject, animal, or individual, or population thereof, or a measure or characteristic representative thereof. In some embodiments, a reference is a measured value. In some embodiments, a reference is an established standard or expected value. In some embodiments, a reference is a historical reference. A reference can be quantitative or qualitative. Typically, as would be understood by those of skill in the art, a reference and the value to which it is compared represents measure under comparable conditions. Those of skill in the art will appreciate when sufficient similarities are present to justify reliance on and/or comparison. In some embodiments, an appropriate reference may be an agent, sample, sequence, subject, animal, or individual, or population thereof, under conditions those of skill in the art will recognize as comparable, e.g., for the

purpose of assessing one or more particular variables (e.g., presence or absence of an agent or condition), or a measure or characteristic representative thereof.

[0058] *Regulatory Sequence:* As used herein in the context of expression of a nucleic acid coding sequence, a regulatory sequence is a nucleic acid sequence that controls expression of a coding sequence. In some embodiments, a regulatory sequence can control or impact one or more aspects of gene expression (e.g., cell-type-specific expression, inducible expression, etc.).

[0059] *Risk:* As used herein with respect to a disease, disorder, or condition, the term “risk” refers to the qualitative or quantitative probability (whether expressed as a percentage or otherwise) that a particular individual will develop the disease, disorder, or condition. In some embodiments, risk is expressed as a percentage. In some embodiments, a risk is a qualitative or quantitative probability that is equal to or greater than 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, or 100%. In some embodiments risk is expressed as a qualitative or quantitative level of risk relative to a reference risk or level or the risk of the same outcome attributed to a reference. In some embodiments, relative risk is increased or decreased in comparison to the reference sample by a factor of 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more.

[0060] *Subject:* As used herein, the term “subject” refers to an organism, typically a mammal (e.g., a human, rat, or mouse). In some embodiments, a subject is suffering from a disease, disorder or condition. In some embodiments, a subject is susceptible to a disease, disorder, or condition. In some embodiments, a subject displays one or more symptoms or characteristics of a disease, disorder or condition. In some embodiments, a subject is not suffering from a disease, disorder or condition. In some embodiments, a subject does not display any symptom or characteristic of a disease, disorder, or condition. In some embodiments, a subject has one or more features characteristic of susceptibility to or risk of a disease, disorder, or condition. In some embodiments, a subject is a subject that has been tested for a disease, disorder, or condition, and/or to whom therapy has been administered. In some instances, a human subject can be interchangeably referred to as a “patient” or “individual.”

[0061] *Susceptible to:* An individual who is “susceptible to” a disease, disorder, or condition is at risk for developing the disease, disorder, or condition. In some embodiments, an individual who is susceptible to a disease, disorder, or condition does not display any symptoms

of the disease, disorder, or condition. In some embodiments, an individual who is susceptible to a disease, disorder, or condition has not been diagnosed with the disease, disorder, and/or condition. In some embodiments, an individual who is susceptible to a disease, disorder, or condition is an individual who has been exposed to conditions associated with, or presents a biomarker status (e.g., a methylation status) associated with, development of the disease, disorder, or condition. In some embodiments, a risk of developing a disease, disorder, and/or condition is a population-based risk (e.g., family members of individuals suffering from the disease, disorder, or condition).

[0062] *Symptoms are reduced:* According to the present invention, “symptoms are reduced” when one or more symptoms of a particular disease, disorder or condition is reduced in magnitude (*e.g.*, intensity, severity, *etc.*) and/or frequency. For the avoidance of doubt, a delay in the onset of a particular symptom is considered one form of reducing the frequency of that symptom.

[0063] *Therapeutically effective amount:* As used herein, “therapeutically effective amount” refers to an amount that produces the desired effect for which it is administered. In some embodiments, the term refers to an amount that is sufficient, when administered to a population suffering from or susceptible to a disease, disorder, and/or condition in accordance with a therapeutic dosing regimen, to treat the disease, disorder, and/or condition. In some embodiments, a therapeutically effective amount is one that reduces the incidence and/or severity of, and/or delays onset of, one or more symptoms of the disease, disorder, and/or condition. Those of ordinary skill in the art will appreciate that a therapeutically effective amount does not necessarily achieve successful treatment in every particular treated individual. Rather, a therapeutically effective amount may be that amount that provides a particular desired pharmacological response in a significant number of subjects when administered to patients in need of such treatment. In some embodiments, reference to a therapeutically effective amount may be a reference to an amount as measured in one or more specific tissues (e.g., a tissue affected by the disease, disorder or condition) or fluids (e.g., blood, saliva, serum, sweat, tears, urine, *etc.*). Those of ordinary skill in the art will appreciate that, in some embodiments, a therapeutically effective amount of a particular agent or therapy may be formulated and/or

administered in a single dose. In some embodiments, a therapeutically effective agent may be formulated and/or administered in a plurality of doses, for example, as part of a dosing regimen.

[0064] *Treatment:* As used herein, the term “treatment” (also “treat” or “treating”) refers to administration of a therapy that partially or completely alleviates, ameliorates, relieves, inhibits, delays onset of, reduces severity of, and/or reduces incidence of one or more symptoms, features, and/or causes of a particular disease, disorder, or condition, or is administered for the purpose of achieving any such result. In some embodiments, such treatment can be of a subject who does not exhibit signs of the relevant disease, disorder, or condition and/or of a subject who exhibits only early signs of the disease, disorder, or condition. Alternatively or additionally, such treatment can be of a subject who exhibits one or more established signs of the relevant disease, disorder and/or condition. In some embodiments, treatment can be of a subject who has been diagnosed as suffering from the relevant disease, disorder, and/or condition. In some embodiments, treatment can be of a subject known to have one or more susceptibility factors that are statistically correlated with increased risk of development of the relevant disease, disorder, or condition.

[0065] *Unit dose:* As used herein, the term “unit dose” refers to an amount administered as a single dose and/or in a physically discrete unit of a pharmaceutical composition. In many embodiments, a unit dose contains a predetermined quantity of an therapeutic agent. In some embodiments, a unit dose contains an entire single dose of the agent. In some embodiments, more than one unit dose is administered to achieve a total single dose. In some embodiments, administration of multiple unit doses is required, or expected to be required, in order to achieve an intended effect. A unit dose can be, for example, a volume of liquid (e.g., an acceptable carrier) containing a predetermined quantity of one or more therapeutic agents, a predetermined amount of one or more therapeutic agents in solid form, a sustained release formulation or drug delivery device containing a predetermined amount of one or more therapeutic agents, etc. It will be appreciated that a unit dose can be present in a formulation that includes any of a variety of components in addition to the therapeutic agent(s). For example, acceptable carriers (e.g., pharmaceutically acceptable carriers), diluents, stabilizers, buffers, preservatives, etc., can be included. It will be appreciated by those skilled in the art, in many embodiments, a total appropriate daily dosage of a particular therapeutic agent can include a portion, or a plurality, of

unit doses, and can be decided, for example, by a medical practitioner within the scope of sound medical judgment. In some embodiments, the specific effective dose level for any particular subject or organism can depend upon a variety of factors including the disorder being treated and the severity of the disorder; activity of specific therapeutic agent employed; specific composition employed; age, body weight, general health, sex and diet of the subject; time of administration, and rate of excretion of the specific therapeutic agent employed; duration of the treatment; drugs and/or additional therapies used in combination or coincidental with specific compound(s) employed, and like factors well known in the medical arts.

BRIEF DESCRIPTION OF THE DRAWING

[0066] Fig. 1. The rate of production of levan (mg/dL over time (minutes)) in a culture of bacteria (strain ZB423) engineered to express levansucrase. Levan is calculated from glucose release and verified in batch culture by HPAEC and NMR (see Figures 3-7).

[0067] Fig. 2 is an SDS-PAGE gel showing accumulation of secreted levansucrase enzyme in supernatant. Lanes 1 and 2 are cell pellet lysates from strains not encoding SacB (PY79) or engineered to encode SacB (ZB423) respectively. Lanes 6 and 7 are cell culture supernatant from strains not encoding SacB (PY79) or engineered to encode SacB (ZB423) respectively. Lanes 3, 4, and 5 are supernatant from other *Bacillus subtilis* strains not relevant to this study. Lane 8 contains cell culture media only as a control reference. Lane 9 contains the protein ladder with bands corresponding to 10, 15, 20, 25, 37, 50, 75, 100, 150, 250 kDa from bottom to top. The supernatant fraction of ZB423 clearly shows a band adjacent to the 50 kDa ladder marker, matching the expected size of 52.99 kDa for SacB protein.

[0068] Fig. 3 High Performance Anion Exchange Chromatography (HPAEC) data for Levan isolated from ZB423 culture. A culture of ZB423 in production media containing sucrose was grown for 72 hours. The cell culture was then pelleted to remove cells and cold ethanol (1:1.5) was added to the supernatant. A white precipitate was observed and further purified by twice repeating dissolution in water and precipitation in cold ethanol followed by centrifugation to collect the pellet. The pellet was dried under vacuum for 24 hours before analysis. A separate industry standard for inulin was analyzed using the same HPAEC method. A strain of *Bacillus subtilis* PY79 that was not engineered to express SacB did not produce a precipitate or pellet

from the supernatant and could not be analyzed. A) A chromatogram of 10g of sample from the ZB423 pellet showing a degree of polymerization (DP) of around 50 B) An inulin standard run with the same HPAEC method and column for comparison showing a DP of around 14 consistent with plant inulin extracts.

[0069] Fig. 4 ¹H-NMR spectra of Levan isolated from ZB423 culture. The spectra shows a solvent peak at 4.78 ppm; a doublet at 4.18, 4.16 ppm (integral: 1); a triplet at 4.09, 4.07, 4.06 ppm (integral: 1), two peaks with amorphous splitting at 3.93, 3.91, 3.88, 3.87 ppm (integral: 2), a doublet at 3.76, 3.74 ppm; a doublet at 3.67, 3.64 ppm; and a triplet at 3.55, 3.53, 3.52 ppm. Solvent was D₂O.

[0070] Fig. 5 ¹³C-NMR Spectra of Levan isolated from ZB423 culture. The spectra shows peaks at 59.96, 104.25, 76.35, 75.24, 80.31, 63.42 ppm.

[0071] Fig. 6 HSQC-NMR Spectra of Levan isolated from ZB423 culture. This spectra shows the associations between ¹³C atoms and ¹H NMR peaks. The result matches that analyzed previously for Levan from plant sources.

[0072] Fig. 7 Table of ¹³C-NMR chemical shifts for Levan isolated from ZB423 culture and comparison to literature.

[0073] Fig. 8 Lactose degradation during transgalactosylation reactions in a culture of *B. subtilis* engineered with beta-galactosidase LacZ. Lactose is calculated from released glucose. The direct correlation between glucose accumulation and Lactose degradation is confirmed in control samples using purified lactase to effect complete Lactose degradation and correlating to released glucose.

[0074] Fig. 9 Trehalulose production from cell supernatant (testing activity of secreted enzyme) isolated from culture broth of engineered strains of *Bacillus subtilis* PY79 with mutB (ZS459) or without mutB (ZS161) operably linked to the *hag* promoter in the expression locus. Trehalulose is calculated as the reduction in total sucrose. The theoretical limit is equal to the amount of sucrose added to the cell supernatant during the assay.

[0075] Fig. 10 Trehalulose production from whole cell broth (testing combined activity of intracellular and secreted enzyme) containing engineered strains of *Bacillus subtilis* PY79 with mutB (ZS459) or without mutB (ZS161) operably linked to the *hag* promoter in the

expression locus. Trehalulose is calculated as the reduction in total sucrose. The theoretical limit is equal to the amount of sucrose added to the cell supernatant during the assay.

DETAILED DESCRIPTION

[0076] Metabolism and nutrition are complex areas in which many challenges have been identified. Functions of the gut (e.g., human gut) and gut microbiome (e.g., human gut microbiome) have been increasingly recognized, presenting new challenges in metabolism and nutrition. Certain of these challenges relate to the role of carbohydrates in human metabolism and nutrition. For instance, one function of the gut and/or gut microbiome is to digest carbohydrates, often into monosaccharides such as glucose, fructose, and galactose. Diets high in carbohydrates have been linked to, e.g., high blood pressure, heart disease, obesity, diabetes, high blood glucose and other health problems. Another important metabolic and nutritional challenge relates to soluble fiber. Soluble fiber can improve, e.g., gut health, and improve or reduce the risk of conditions such as high blood cholesterol, heart disease, obesity, diabetes, high blood glucose, and other health problems. Gut carbohydrates and gut soluble fiber can beneficially (e.g., therapeutically) or detrimentally impact the composition of the gut microbiome (e.g., the types and concentrations or amounts of the various microbes of the gut microbiome).

[0077] Those of skill in the art will appreciate that modern diets typically include concentrations and/or amounts of carbohydrate (e.g., monosaccharides and/or disaccharides and/or complex carbohydrate) that are in excess of nutritional requirements. Excess monosaccharides and/or disaccharides can in various instances result from consumption of monosaccharides, disaccharides, or complex carbohydrates, any and/or all of which are common in many modern diets. For instance, according to certain estimates, 5% of consumed sucrose can traverse the small intestine without being degraded and digested. Thus, from a typical carbohydrate source such as a can of soda, containing, e.g., 30g of sucrose, as much as 1.5g of sucrose can pass through the stomach and small intestine to reach the colon). Furthermore, sucrose in the intestine can be cleaved into glucose and fructose. While glucose can be rapidly absorbed through the intestinal lining, fructose is absorbed more slowly, resulting in accumulation of fructose in the colon. The present specification recognizes that agents that can increase certain generally beneficial microbiome processes, activities, and/or types of microbes,

that can decrease certain generally deleterious microbiome processes, activities, and/or types of microbes, or that can promote a beneficial overall microbiome composition, are needed. Those of skill in the art will further appreciate that it is not necessary to articulate individual effects (e.g., on the composition of the gut microbiome) of an agent (e.g., to identify gut microbiome bacteria that are increased or decreased by delivery of the agent to the gut) to demonstrate and/or appreciate that the agent has a beneficial impact on health and/or the gut microbiome.

[0078] Those of skill in the art will further appreciate that modern diets typically include concentrations and/or amounts of fiber (e.g., soluble fiber) that are lower than the concentrations and/or amounts optimal for human health. Consumption of fiber (e.g., a high-fiber diet, e.g., a high-soluble fiber diet) helps maintain gut health. Fiber is commonly classified as soluble fiber, which can be dissolved in water to form a gel-like material, or insoluble fiber, which cannot be dissolved in water. Consumption of fiber (particularly soluble fiber) can help to maintain healthy body weight (e.g., to reduce or stabilize body weight), and/or to decrease risk of developing diabetes, heart disease, and/or some types of cancer (e.g., colorectal cancer and/or other gut or intestinal cancers). Consumption of fiber (particularly soluble fiber) can decrease risk of dying from cardiovascular disease and all cancers, causing and/or permitting increased longevity. Consumption of fiber (particularly soluble fiber) can decrease risk of developing hemorrhoids and/or small pouches in your colon (diverticular disease). Consumption of fiber (particularly soluble fiber) can decrease blood pressure and inflammation. Consumption of fiber (particularly soluble fiber) can slow absorption of glucose into blood and/or cause a decrease in blood glucose levels and/or stabilize blood glucose levels (e.g., decrease variation in blood glucose levels over time), e.g., in diabetic subjects. Soluble fiber can help decrease blood cholesterol (e.g., total blood cholesterol concentration or amount), e.g., by decreasing low-density lipoprotein levels.

[0079] Gut carbohydrates and gut soluble fiber impact the composition of the gut microbiome. The gut microbiome typically includes trillions of microbes that contribute to, e.g., metabolism, nutrient and mineral absorption, synthesis of enzymes, vitamins and amino acids, and production of short-chain fatty acids (SCFAs). Microbes of the gut microbiome have tremendous potential to impact health (e.g., contribute to wellbeing or to disease) and/or physiology. Microbes of the gut microbiome can have metabolic activities that impact health of

a host organism, protect a host organism against pathogens, educate the immune system of a host organism, and/or directly or indirectly affect many physiologic functions. The gut microbiome includes a variety of different types of microbes (e.g., a variety of different bacterial phyla, classes, orders, families, genera, species, and/or strains), each of which can be characterized by different metabolic characteristics. For example, different types of microbes may demonstrate increased or decreased growth under any of a variety of conditions, e.g., in the presence of carbohydrates or in the presence of fiber. Moreover, certain microbes are considered to be beneficial, certain microbes are considered to be harmful, and the overall composition of the gut microbiome (e.g., the relative amounts of different microbes in the gut microbiome) is considered to be an important factor in health.

[0080] Carbohydrate consumption and/or gut carbohydrate (e.g., consumption and/or excess of monosaccharides and/or disaccharides) can adversely impact the gut microbiome and/or cause various conditions associated with excess carbohydrate. Insufficient soluble fiber in the gut can also adversely impact the gut microbiome (e.g., gut microbiome composition and/or activity) and/or cause various conditions associated with insufficient fiber. Soluble fiber can also promote growth of health-promoting bacteria. Accordingly, carbohydrate consumption and/or insufficient fiber consumption can cause a disruption of the balance of microbiome strains (dysbiosis) of the gut microbiome, e.g., of the large intestine, e.g., by increasing the concentration or amount of Proteobacteria, decreasing the concentration or amount of Bacteroidetes, and/or causing an increase the ratio of Proteobacteria to Bacteroidetes. Dysbiosis of gut microbiota is associated with the pathogenesis of intestinal conditions (e.g. inflammatory bowel disease, and irritable bowel syndrome (IBS)), and extra-intestinal conditions (e.g. allergy, asthma, metabolic syndrome, cardiovascular disease, and obesity). Dysbiosis can include and/or present symptoms including bloating, cramping, gas, and diarrhea. Additionally, concentration and/or amount of Proteobacteria and/or Bacteroidetes, and/or the ratio thereof, can contribute to maintenance of immunological homeostasis and/or epithelial integrity in the intestinal mucosa. The genus Bacteroidetes has been associated with numerous health benefits, including the downregulation of inflammatory responses in the gut. Further to these impacts on the composition of the gut microbiome and functions thereof, sucrose and gut cleavage products

glucose and fructose are known to have direct effects in silencing certain health-beneficial functions of gut microbiome bacteria such as bacteroides.

[0081] The present disclosure therefore recognizes a variety of distinct gut conditions (e.g., characterized by concentration or amount of one or more metabolites, e.g., increased or decreased relative to a reference) that benefit health and/or the gut microbiome. These target gut = conditions that benefit health and/or the gut microbiome can include any of one or more of (i) decrease in the concentration or amount of monosaccharides and/or disaccharides in the gut as compared to a reference and/or (ii) increase in the concentration or amount of gut fiber (e.g., gut soluble fiber) as compared to a reference, e.g., where the reference is a concentration or amount in the same subject at an earlier time, a concentration or amount identified as normal or healthy, a concentration or amount identified as abnormal or unhealthy, a concentration or amount representative of typical or healthy subjects, a concentration or amount representative of atypical or unhealthy subjects. The present disclosure further provides that decreasing the concentration or amount monosaccharides and/or disaccharides in the gut can be achieved, e.g., by competitive uptake of monosaccharides and/or disaccharides (e.g., sucrose) in the gut (e.g., in the small intestine and/or colon) by an alternative and/or beneficial process, and/or by removal of fructose produced from sucrose degradation in the gut (e.g., in the small intestine). The present disclosure further provides that the concentration or amount of gut fiber (e.g., gut soluble fiber) can be increased by synthesis in the gut of soluble fiber. The present disclosure further provides that these target gut nutrient conditions of reduced carbohydrate and increased fiber provide synergistic benefits to health and/or the gut microbiome, at least because monosaccharides and/or disaccharides can inhibit beneficial growth of the microbiome and/or of microbes thereof while soluble fiber promotes positive growth of the microbiome and/or of microbes thereof, and separately or additionally in some instances because low carbohydrate and high fiber can together modulate, decrease, and/or stabilize blood glucose levels. Moreover, methods and compositions of the present disclosure surprisingly and synergistically achieve both target gut nutrient conditions of decreased concentration or amount of gut monosaccharides and/or disaccharides and increased concentration or amount of fiber.

[0082] The present disclosure recognizes, among other things, at least two forms of synergy between carbohydrates (such as monosaccharides and/or disaccharides, including

without limitation monosaccharides produced by digestion of complex carbohydrates) and soluble fiber: first, that soluble fiber in the gut can slow the digestion of carbohydrate (e.g., monosaccharides and/or disaccharides), and second that both carbohydrates (e.g., monosaccharides and/or disaccharides) and soluble fiber have profound impacts on the gut microbiome, the present disclosure including that reducing gut monosaccharides and/or disaccharides and increasing gut soluble fiber are independently and synergistically beneficial. The present disclosure provides compositions such as engineered bacteria, and methods of using the same, that benefit health by decreasing concentration and/or amount of gut monosaccharides and/or disaccharides and increasing gut soluble fiber, e.g., by enzymatic synthesis of fiber from substrates including carbohydrate (e.g., monosaccharides and/or disaccharides).

[0083] Various methods and compositions of the present disclosure include an engineered bacterium, such as a probiotic bacterium, that decreases amount and/or concentration of monosaccharides and/or disaccharides in the gut and increases amount and/or concentration of soluble fiber in the gut by enzymatically converting carbohydrate (e.g., monosaccharides and/or disaccharides) to soluble fiber. Enzymatic reactions that perform this function include but are not limited to oligofructose-synthases, which cleave sucrose and seclude fructose into a growing oligo-fructose chain which is not digestible by human enzymes in the intestinal tract. Accordingly, in certain exemplary embodiments, a composition or method of the present disclosure degrades sucrose and incorporates fructose into a fiber molecule. In certain particular methods and composition of the present disclosure, an engineered bacterium processes monosaccharides and/or disaccharides to produce levan fiber (e.g., by incorporating carbohydrate moieties into soluble fiber molecules).

Reducing Carbohydrate and Synthesizing Fiber

[0084] Compositions and methods of the present disclosure that decrease concentration or amount of gut carbohydrate (monosaccharides and/or disaccharides) and increase concentration or amount of gut fiber (e.g., soluble fiber) include enzymes that convert carbohydrate to fiber and further include engineered bacteria that encode and/or express enzymes that convert carbohydrate to fiber. The present disclosure includes the inventors' recognition that certain plant and bacterial enzymes participate in, catalyze, and/or cause reactions in which

soluble indigestible fiber molecules are synthesized (e.g., directly) from substrates including one or more carbohydrates, e.g., by incorporating carbohydrate moieties into fiber molecules. Certain such reactions typically occur in non-therapeutic contexts and/or contexts that are not within a subject, e.g., during fermentation of food with soil bacteria, or during growth of plants in soil. Various bacteria and plants employ conversion of carbohydrates to fiber in producing energy storage molecules, structural molecules, and/or osmotic gradients. To the knowledge of the present inventors, none of these bacteria or plants, nor cells thereof, can, upon administration to a subject, cause incorporation of carbohydrate moieties consumed by the subject into fiber molecules, nor do so in therapeutically effective amounts (e.g., there is no therapeutically effective dosage of these bacteria or plants or cells thereof for synthesis in a subject of fiber molecules from consumed carbohydrates). The present disclosure reveals that introduction of such reactions into the context of the gut by administration to a subject of a bacterium engineered for such a reaction can decrease concentration or amount of gut monosaccharides and/or disaccharides while increasing gut soluble fiber.

Transgenes

[0085] The present disclosure includes engineered nucleic acids (transgenes) that encode and/or express a variety of exemplary enzymes that synthesize fiber from substrates including carbohydrate (fiber-synthesizing enzymes). Various fiber-synthesizing enzymes are known in the art, e.g., from studies of the conversion of carbohydrates to fibers in microorganisms (e.g., for use in a controlled fermentation context). For example, various fiber-synthesizing enzymes have been characterized, e.g., for sequence and/or fiber production characteristics in laboratory settings and systems such as *E. coli*.

[0086] In various embodiments, fiber-synthesizing enzymes of the present disclosure utilize a carbohydrate substrate selected from one or more of UDP-glucose, UDP-galactose, UDP-fucose, ADP- α -D-glucose, UDP-galacturonic acid, UDP-N-acetyl- α -D-glucosamine, galactinol, lactose, glucose, and sucrose and/or wherein the carbohydrate decreased in amount or concentration in the gut is selected from one or more of glucose, galactose, sucrose, fructose, and/or lactose as a substrate in synthesizing a fiber that includes glucose, fructose, galacturonic acid, N-acetyl-D-glucosamine, and/or galactose monomers, e.g., wherein synthesized fibers

include one or more of laminaribiose, callose, curdlan, oat beta-glucan, laminarin, pleuran, lentinan, yeast beta glucan, trehalulose, inulin, kestose, nystose, levan, raffinose, stachyose, and/or verbascose. In various embodiments, fiber-synthesizing enzymes of the present disclosure utilize UDP-glucose as a substrate in synthesizing fiber including glucose monomers, e.g., where synthesized fibers are or include one or more of laminaribiose, callose, curdlan, oat beta-glucan, laminarin, pleuran, lentinan, yeast beta glucan, and/or trehalulose. In various embodiments, fiber-synthesizing enzymes of the present disclosure utilize sucrose as a substrate in synthesizing fiber including fructose monomers, e.g., where synthesized fibers are or include inulin, kestose, nystose, and/or levan. In various embodiments, fiber-synthesizing enzymes of the present disclosure utilize sucrose and UDP-galactose as a substrate in synthesizing fiber including galactose monomers, e.g., where synthesized fibers are or include raffinose, stachyose, and/or verbascose.

[0087] Representative examples of fiber-synthesizing enzymes include a 1,3-beta-glucan synthase (e.g., 1,3-beta-glucan synthase component FKS1 or callose synthase 1), a 1,3;1,4-beta-D-glucan synthase (e.g., probable mixed-linked glucan synthase 6), a 1,6-beta-glucan synthase (e.g., Cell wall synthesis protein KRE9) a sucrose isomerase, a levansucrase, or a 1,6-alpha-galactosyltransferase (e.g., galactinol--sucrose galactosyltransferase). Representative examples of fiber-synthesizing enzymes include Levansucrase (sucrose 6-fructosyltransferase; beta-2,6-fructosyltransferase; beta-2,6-fructan:D-glucose 1-fructosyltransferase; EC 2.4.1.10); inulosucrase (sucrose:2,1-beta-D-fructan 1-beta-D-fructosyltransferase; sucrose 1-fructosyltransferase; EC 2.4.1.9); fructosyltransferase (2,1-fructan:2,1-fructan 1-fructosyltransferase; 1,2-beta-D-fructan 1(F)-fructosyltransferase; 1,2-beta-D-fructan:1,2-beta-D-fructan 1(F)-beta-D-fructosyltransferase; 1,2-beta-fructan 1(F)-fructosyltransferase; FFT; Fructan:fructan fructosyl transferase, EC 2.4.1.100); sucrose fructosyltransferase (SST; Sucrose 1(F)-fructosyltransferase; Sucrose-sucrose 1-fructosyltransferase; Sucrose:sucrose 1(F)-beta-D-fructosyltransferase; Sucrose:sucrose 1-fructosyltransferase; EC 2.4.1.99); Mutansucrase (EC 2.4.1.372); Sucrose--glucan glucosyltransferase (EC 2.4.1.4); Sucrose 6-glucosyltransferase (EC 2.4.1.5); Alpha-(1->2) branching sucrase (EC 2.4.1.373); Alternansucrase (Sucrose-1,6(3)-alpha-glucan 6(3)-alpha-glucosyltransferase; Sucrose:1,6-, 1,3-alpha-D-glucan 3-alpha- and 6-alpha-D-

glucosyltransferase; EC 2.4.1.140); Sucrose--1,6-alpha-glucan 3(6)-alpha-glucosyltransferase (EC 2.4.1.125).

[0088] In one embodiment, a probiotic organism is engineered with a 2,1 beta fructosyltransferase (inulosucrase) to enable the synthesis of fructooligosaccharides (inulin, 1-kestose, 1-nystose) from consumed sucrose in the digestive tract. In another embodiment, a probiotic organism is engineered with a 2,6 beta fructosyltransferase (levanosucrase) to enable the synthesis of fructooligosaccharides (levan, 6-kestose, 6-nystose) from consumed sucrose in the digestive tract. In another embodiment, a probiotic organism is engineered with a galactosyltransferase specific to lactose (such as $\beta(1\rightarrow4)$ galactosyltransferase) to enable the synthesis of lactose terminal galactooligosaccharides (human milk oligosaccharides, galactooligosaccharides) from consumed lactose in the digestive tract. In another embodiment, a probiotic organism is engineered with a galactosyltransferase specific to galactinol (such as $\alpha(1\rightarrow6)$ galactosyltransferase) to enable the synthesis of sucrose terminal galactooligosaccharides (raffinose, stachyose, verbascose) from consumed sucrose and glucose in the digestive tract, optionally with galactinol synthase, galactose isomerase, and membrane transporters to increase precursors and improve synthesis rates. In another embodiment, a probiotic organism is engineered with a $\beta(1\rightarrow3)$ glucan synthase (1,3-Beta-glucan synthase) to enable the synthesis of beta-glucans (laminaribiose, callose, curdlan, oat beta-glucan) from consumed glucose or glucose produced from consumed sucrose in the digestive tract, optionally including fructose isomerase to enable synthesis from consumed fructose, optionally including a 1,4-beta-glucosyltransferase to enable synthesis of crosslinked polymers with improved prebiotic qualities from consumed sucrose and glucose in the digestive tract. In another embodiment, a probiotic organism is engineered with a $\beta(1\rightarrow3)$ glucan synthase (1,3-Beta-glucan synthase) is further engineered with a $\beta(1\rightarrow6)$ glucan synthase (1,6-Beta-glucan synthase) to create rare crosslinked prebiotic fibers (Laminarin/Pleurin/Lentinan/Yeast beta glucan) from consumed sucrose and glucose in the digestive tract. In another embodiment, a probiotic organism is engineered with a $\alpha(1\rightarrow1)$ sucrose isomerase (trehalulose synthase) to enable the synthesis of rare prebiotic disaccharides (Trehalulose) from consumed glucose and fructose in the digestive tract. In another embodiment a probiotic organism is engineered with the pectin biosynthesis

pathway to enable the synthesis of pectin and pectin-like polymers from consumed glucose and sucrose in the digestive tract.

[0089] Representative examples of fiber-synthesizing enzymes are further provided in the below table:

Table 1. Exemplary fiber-synthesizing enzymes

Enzyme(s)	Exemplary Gut Carbohydrate(s) Decreased (Enzyme Substrate(s))	Fiber Monomer(s)	Fiber Monomer Linkage(s)	Product(s) (Common name(s))	Exemplary Enzyme Sequences
1,3-beta-glucan synthase	Glucose Sucrose Fructose (UDP-Glucose)	Glucose	$\beta(1\rightarrow3)$	Laminaribiose Callose Curdlan	<i>Saccharomyces cerevisiae</i> FKS1 (1,3-beta-glucan synthase component)
	Glucose Sucrose Fructose (UDP-Glucose)	Glucose	$\beta(1\rightarrow3)$	Callose	<i>Arabidopsis thaliana</i> (Mouse-ear cress) CALS1 (Callose synthase 1)
1,3;1,4-beta-D-glucan synthase	Glucose Sucrose Fructose (UDP-Glucose)	Glucose	$\beta(1\rightarrow3)$ $\beta(1\rightarrow4)$	Oat beta-glucan	<i>Oryza sativa</i> subsp. japonica CSLF6 (Probable mixed-linked glucan synthase 6)
1,3-beta-glucan synthase, 1,6-beta-glucan synthase	Glucose Sucrose Fructose (UDP-Glucose)	Glucose	$\beta(1\rightarrow3)$ $\beta(1\rightarrow6)$	Laminarin Pleurin Lentinan Yeast beta glucan	<i>Saccharomyces cerevisiae</i> KRE9 (Cell wall synthesis protein KRE9)
Sucrose Isomerase, Trehalulose Synthase	Sucrose (Sucrose)	Glucose Fructose	$\alpha(1\rightarrow1)$	Trehalulose	<i>Pseudomonas mesoacidophila</i> MutB (Sucrose isomerase (Trehalulose synthase))
Trehalose-6-phosphate synthase	Glucose Sucrose Fructose (UDP-Glucose, Glucose)	Glucose	$\alpha(1\rightarrow1)$	Trehalose	<i>Escherichia coli</i> (strain K12) otsA (Trehalose-6-phosphate synthase)
Inulinsucrase	Sucrose (Sucrose)	Fructose	$\beta(2\rightarrow1)$	Inulin Kestose Nystose	<i>Lactobacillus johnsonii</i> InuJ (Inulosucrase)
Levansucrase	Sucrose (Sucrose)	Fructose	$\beta(2\rightarrow6)$	Levan	<i>Bacillus subtilis</i> SacB (Levansucrase)
1,6-alpha-galactosyltransferase	Lactose Glucose Sucrose Fructose	Galactose	$\alpha(1\rightarrow6)$	Raffinose Stachyose Verbascose	<i>Oryza sativa</i> subsp. japonica RFS (Galactinol-sucrose galactosyltransferase)

	(Galactinol)				
α -1,4-galactosyltransferase	Lactose Glucose Sucrose Fructose (UDP-galactose)	Galactose	$\alpha(1\rightarrow4)$	Globotriose, Human milk oligosaccharides (HMOs)	<i>Neisseria meningitidis</i> lgtC (Alpha 1,4 galactosyltransferase)
Alpha-1,2-fucosyltransferase	Lactose Glucose Sucrose Fructose (GDP-fucose)	Fucose	$\alpha(1\rightarrow2)$	Human milk oligosaccharides (HMOs)	<i>Thermosynechococcus elongatus</i> tll0994 (Alpha-1,2-fucosyltransferase)
beta-galactosidase	Lactose (Lactose)	Galactose	$\beta(1\rightarrow3)$	Human milk oligosaccharides (HMOs)	<i>Geobacillus kaustophilus</i> BgaB (Beta-galactosidase bgaB)
b-D-Galactosidase	Lactose (Lactose)	Galactose	$\beta(1\rightarrow4)$	Human milk oligosaccharides (HMOs)	<i>Bacillus circulans</i> Bga (Beta-galactosidase)
cellulose synthase	Glucose Sucrose Fructose (UDP-Glucose)	Glucose	$\beta(1\rightarrow4)$	Cellobiose/Cellulose/microcellulose/Cotton	<i>Komagataeibacter xylinus</i> (Gluconacetobacter xylinus) AcsAB (Cellulose synthase 1)
maltose synthase, starch synthase	Glucose Sucrose Fructose (UDP-Glucose)	Glucose	$\alpha(1\rightarrow4)$	Maltose/Amylose/Starch/Glycogen	<i>Zea mays</i> (Maize) Ss1 (Starch synthase, chloroplastic/amyloplastic)
Starch-branching enzyme	Glucose Sucrose Fructose (UDP-Glucose)	Glucose	$\alpha(1\rightarrow4)$, $\alpha(1\rightarrow6)$	Amylopectin/Starch	<i>Arabidopsis thaliana</i> (Mouse-ear cress) SBE3 (1,4-alpha-glucan-branching enzyme 3, chloroplastic/amyloplastic)
Glycogen synthase	Glucose Sucrose Fructose (ADP- α -D-glucose)	Glucose	$\alpha(1\rightarrow4)$	Microbial Glycogen	<i>Bacillus subtilis</i> (strain 168) glgA (Glycogen synthase)
galacturonosyl transferase	Glucose Sucrose Fructose (UDP- α -D-galacturonate, various)	Galacturonic acid	$\alpha(1\rightarrow4)$	Pectin	<i>Arabidopsis thaliana</i> (Mouse-ear cress) GAUT1 (Polygalacturonate 4-alpha-galacturonosyltransferase)
Chitinoligosaccharide synthase	Glucose Sucrose Fructose (UDP-N-acetyl-alpha-D-glucosamine)	N-acetyl-D-glucosamine	$\beta(1\rightarrow4)$	Chitin	<i>Azorhizobium caulinodans</i> nodC (N-acetylglucosaminyltransferase)

[0090] In some embodiments, an FKS1 fiber-synthesizing enzyme has an amino acid sequence according to SEQ ID NO: 1 (UniProt Accession No. P38631). In various embodiments, a fiber-synthesizing enzyme has at least 80% identity to SEQ ID NO: 1, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 1. The present disclosure further includes nucleic acid sequences encoding an FKS1 fiber-synthesizing enzyme. In certain embodiments, a nucleic acid sequence encoding an FKS1 fiber-synthesizing enzyme is a sequence engineered for expression in *B. subtilis*, e.g., a nucleic acid sequence having at least 80% identity to SEQ ID NO: 2., e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 2.

[0091] In some embodiments, a CALS1 fiber-synthesizing enzyme has an amino acid sequence according to SEQ ID NO: 3 (UniProt Accession No. Q9AUE0). In various embodiments, a fiber-synthesizing enzyme has at least 80% identity to SEQ ID NO: 3, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 3. The present disclosure further includes nucleic acid sequences encoding a CALS1 fiber-synthesizing enzyme. In certain embodiments, a nucleic acid sequence encoding a CALS1 fiber-synthesizing enzyme is a sequence engineered for expression in *B. subtilis*, e.g., a nucleic acid sequence having at least 80% identity to SEQ ID NO: 4., e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 4.

[0092] In some embodiments, a CSLF6 fiber-synthesizing enzyme has an amino acid sequence according to SEQ ID NO: 5 (UniProt Accession No. Q84UP7). In various embodiments, a fiber-synthesizing enzyme has at least 80% identity to SEQ ID NO: 5, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 5. The present disclosure further includes nucleic acid sequences encoding a CSLF6 fiber-synthesizing enzyme. In certain embodiments, a nucleic acid sequence encoding a CSLF6 fiber-synthesizing enzyme is a sequence engineered for expression in *B. subtilis*, e.g., a nucleic acid sequence having at least 80% identity to SEQ ID NO: 6., e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 6.

[0093] In some embodiments, a KRE9 fiber-synthesizing enzyme has an amino acid sequence according to SEQ ID NO: 7 (UniProt Accession No. P39005). In various

embodiments, a fiber-synthesizing enzyme has at least 80% identity to SEQ ID NO: 7, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 7. The present disclosure further includes nucleic acid sequences encoding a KRE9 fiber-synthesizing enzyme. In certain embodiments, a nucleic acid sequence encoding a KRE9 fiber-synthesizing enzyme is a sequence engineered for expression in *B. subtilis*, e.g., a nucleic acid sequence having at least 80% identity to SEQ ID NO: 8, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 8.

[0094] In some embodiments, a MutB fiber-synthesizing enzyme has an amino acid sequence according to SEQ ID NO: 9 (UniProt Accession No. Q2PS28). In various embodiments, a fiber-synthesizing enzyme has at least 80% identity to SEQ ID NO: 9, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 9. The present disclosure further includes nucleic acid sequences encoding a MutB fiber-synthesizing enzyme. In certain embodiments, a nucleic acid sequence encoding a MutB fiber-synthesizing enzyme is a sequence engineered for expression in *B. subtilis*, e.g., a nucleic acid sequence having at least 80% identity to SEQ ID NO: 10, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 10.

[0095] In some embodiments, a otsA fiber-synthesizing enzyme has an amino acid sequence according to SEQ ID NO: 33 (UniProt Accession No. P31677). In various embodiments, a fiber-synthesizing enzyme has at least 80% identity to SEQ ID NO: 33, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 33. The present disclosure further includes nucleic acid sequences encoding a MutB fiber-synthesizing enzyme. In certain embodiments, a nucleic acid sequence encoding a MutB fiber-synthesizing enzyme is a sequence engineered for expression in *B. subtilis*, e.g., a nucleic acid sequence having at least 80% identity to SEQ ID NO: 34, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 34.

[0096] In some embodiments, an InuJ fiber-synthesizing enzyme has an amino acid sequence according to SEQ ID NO: 11 (UniProt Accession No. Q74K42). In various embodiments, a fiber-synthesizing enzyme has at least 80% identity to SEQ ID NO: 11, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 11. The present disclosure further includes nucleic acid sequences encoding an InuJ fiber-synthesizing

enzyme. In certain embodiments, a nucleic acid sequence encoding an InuJ fiber-synthesizing enzyme is a sequence engineered for expression in *B. subtilis*, e.g., a nucleic acid sequence having at least 80% identity to SEQ ID NO: 12, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 12.

[0097] In some embodiments, an RFS fiber-synthesizing enzyme has an amino acid sequence according to SEQ ID NO: 13 (UniProt Accession No. Q5VQG4). In various embodiments, a fiber-synthesizing enzyme has at least 80% identity to SEQ ID NO: 13, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 13. The present disclosure further includes nucleic acid sequences encoding an RFS fiber-synthesizing enzyme. In certain embodiments, a nucleic acid sequence encoding an RFS fiber-synthesizing enzyme is a sequence engineered for expression in *B. subtilis*, e.g., a nucleic acid sequence having at least 80% identity to SEQ ID NO: 14, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 14.

[0098] In some embodiments, a IgtC fiber-synthesizing enzyme has an amino acid sequence according to SEQ ID NO: 35 (UniProt Accession No. Q8KHJ3). In various embodiments, a fiber-synthesizing enzyme has at least 80% identity to SEQ ID NO: 35, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 35. The present disclosure further includes nucleic acid sequences encoding a IgtC fiber-synthesizing enzyme. In certain embodiments, a nucleic acid sequence encoding a IgtC fiber-synthesizing enzyme is a sequence engineered for expression in *B. subtilis*, e.g., a nucleic acid sequence having at least 80% identity to SEQ ID NO: 36, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 36.

[0099] In some embodiments, a tll0994 fiber-synthesizing enzyme has an amino acid sequence according to SEQ ID NO: 37 (UniProt Accession No. Q8DK72). In various embodiments, a fiber-synthesizing enzyme has at least 80% identity to SEQ ID NO: 37, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 37. The present disclosure further includes nucleic acid sequences encoding a tll0994 fiber-synthesizing enzyme. In certain embodiments, a nucleic acid sequence encoding a tll0994 fiber-synthesizing enzyme is a sequence engineered for expression in *B. subtilis*, e.g., a nucleic acid sequence

having at least 80% identity to SEQ ID NO: 38, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 38.

[0100] In some embodiments, a BgaB fiber-synthesizing enzyme has an amino acid sequence according to SEQ ID NO: 39 (UniProt Accession No. P19668). In various embodiments, a fiber-synthesizing enzyme has at least 80% identity to SEQ ID NO: 39, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 39. The present disclosure further includes nucleic acid sequences encoding a BgaB fiber-synthesizing enzyme. In certain embodiments, a nucleic acid sequence encoding a BgaB fiber-synthesizing enzyme is a sequence engineered for expression in *B. subtilis*, e.g., a nucleic acid sequence having at least 80% identity to SEQ ID NO: 40, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 40.

[0101] In some embodiments, a Bga fiber-synthesizing enzyme has an amino acid sequence according to SEQ ID NO: 41 (UniProt Accession No. E5RWQ2). In various embodiments, a fiber-synthesizing enzyme has at least 80% identity to SEQ ID NO: 41, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 41. The present disclosure further includes nucleic acid sequences encoding a Bga fiber-synthesizing enzyme. In certain embodiments, a nucleic acid sequence encoding a Bga fiber-synthesizing enzyme is a sequence engineered for expression in *B. subtilis*, e.g., a nucleic acid sequence having at least 80% identity to SEQ ID NO: 42, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 42.

[0102] In some embodiments, a ScsAB fiber-synthesizing enzyme has an amino acid sequence according to SEQ ID NO: 43 (UniProt Accession No. P0CW87). In various embodiments, a fiber-synthesizing enzyme has at least 80% identity to SEQ ID NO: 43, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 43. The present disclosure further includes nucleic acid sequences encoding a ScsAB fiber-synthesizing enzyme. In certain embodiments, a nucleic acid sequence encoding a ScsAB fiber-synthesizing enzyme is a sequence engineered for expression in *B. subtilis*, e.g., a nucleic acid sequence having at least 80% identity to SEQ ID NO: 44, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 44.

[0103] In some embodiments, a Ss1 fiber-synthesizing enzyme has an amino acid sequence according to SEQ ID NO: 45 (UniProt Accession No. O49064). In various embodiments, a fiber-synthesizing enzyme has at least 80% identity to SEQ ID NO: 45, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 45. The present disclosure further includes nucleic acid sequences encoding a Ss1 fiber-synthesizing enzyme. In certain embodiments, a nucleic acid sequence encoding a Ss1 fiber-synthesizing enzyme is a sequence engineered for expression in *B. subtilis*, e.g., a nucleic acid sequence having at least 80% identity to SEQ ID NO: 46, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 46.

[0104] In some embodiments, a SBE3 fiber-synthesizing enzyme has an amino acid sequence according to SEQ ID NO: 47 (UniProt Accession No. D2WL32). In various embodiments, a fiber-synthesizing enzyme has at least 80% identity to SEQ ID NO: 47, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 47. The present disclosure further includes nucleic acid sequences encoding a SBE3 fiber-synthesizing enzyme. In certain embodiments, a nucleic acid sequence encoding a SBE3 fiber-synthesizing enzyme is a sequence engineered for expression in *B. subtilis*, e.g., a nucleic acid sequence having at least 80% identity to SEQ ID NO: 48, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 48.

[0105] In some embodiments, a HlgA fiber-synthesizing enzyme has an amino acid sequence according to SEQ ID NO: 49 (UniProt Accession No. P39125). In various embodiments, a fiber-synthesizing enzyme has at least 80% identity to SEQ ID NO: 49, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 49. The present disclosure further includes nucleic acid sequences encoding a HlgA fiber-synthesizing enzyme. In certain embodiments, a nucleic acid sequence encoding a HlgA fiber-synthesizing enzyme is a sequence engineered for expression in *B. subtilis*, e.g., a nucleic acid sequence having at least 80% identity to SEQ ID NO: 50, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 50.

[0106] In some embodiments, a GAUT1 fiber-synthesizing enzyme has an amino acid sequence according to SEQ ID NO: 51 (UniProt Accession No. Q9LE59). In various embodiments, a fiber-synthesizing enzyme has at least 80% identity to SEQ ID NO: 51, e.g., at

least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 51. The present disclosure further includes nucleic acid sequences encoding a GAUT1 fiber-synthesizing enzyme. In certain embodiments, a nucleic acid sequence encoding a GAUT1 fiber-synthesizing enzyme is a sequence engineered for expression in *B. subtilis*, e.g., a nucleic acid sequence having at least 80% identity to SEQ ID NO: 52, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 52.

[0107] In some embodiments, a NodC fiber-synthesizing enzyme has an amino acid sequence according to SEQ ID NO: 53 (UniProt Accession No. Q07755). In various embodiments, a fiber-synthesizing enzyme has at least 80% identity to SEQ ID NO: 53, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 53. The present disclosure further includes nucleic acid sequences encoding a NodC fiber-synthesizing enzyme. In certain embodiments, a nucleic acid sequence encoding a NodC fiber-synthesizing enzyme is a sequence engineered for expression in *B. subtilis*, e.g., a nucleic acid sequence having at least 80% identity to SEQ ID NO: 54, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 54.

[0108] In some embodiments, a levansucrase fiber-synthesizing enzyme has an amino acid sequence according to SEQ ID NO: 15 (UniProt Accession No. Q43998), a levansucrase enzyme expressed by *Acetobacter diazotrophicus*, and/or a levansucrase enzyme encoded by *Acetobacter diazotrophicus* lsdA. In various embodiments, a fiber-synthesizing enzyme has at least 80% identity to SEQ ID NO: 15, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 15.

[0109] In some embodiments, a levansucrase fiber-synthesizing enzyme has an amino acid sequence according to SEQ ID NO: 16 (UniProt Accession No. P0DJA3), a levansucrase enzyme expressed by *Zymomonas mobilis* subsp. *Mobilis*, and/or a levansucrase enzyme encoded by *Zymomonas mobilis* subsp. *mobilis* sacB. In various embodiments, a fiber-synthesizing enzyme has at least 80% identity to SEQ ID NO: 16, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 16.

[0110] In some embodiments, a levansucrase fiber-synthesizing enzyme has an amino acid sequence according to SEQ ID NO: 17 (UniProt Accession No. Q97I81), a levansucrase enzyme expressed by *Clostridium acetobutylicum*, and/or a levansucrase enzyme encoded by

Clostridium acetobutylicum sacB. In various embodiments, a fiber-synthesizing enzyme has at least 80% identity to SEQ ID NO: 17, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 17.

[0111] In some embodiments, a levansucrase fiber-synthesizing enzyme has an amino acid sequence according to SEQ ID NO: 18 (UniProt Accession No. P05655), a levansucrase enzyme expressed by *Bacillus Subtilis natto* or *Bacillus Subtilis PY79*, and/or a levansucrase enzyme encoded by *Bacillus Subtilis natto sacB* and/or *Bacillus Subtilis PY79 sacB*. In various embodiments, a fiber-synthesizing enzyme has at least 80% identity to SEQ ID NO: 18, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 18.

[0112] The present disclosure further includes nucleic acid sequences encoding a levansucrase fiber-synthesizing enzyme. In certain embodiments, a nucleic acid sequence encoding a levansucrase fiber-synthesizing enzyme is a sequence engineered for expression in *B. subtilis*, e.g., a nucleic acid sequence having at least 80% identity to SEQ ID NO: 19, e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity to SEQ ID NO: 19. Those of skill in the art will appreciate that fiber-synthesizing enzymes of the present disclosure are effective to produce fiber when provided with enzyme substrate as disclosed herein.

Accordingly, at least because as those of skill in the art will appreciate many bacteria are able to process monosaccharides and/or disaccharides into modified forms, constituent monosaccharides, and/or other substrate carbohydrates those of skill will appreciate that a variety of carbohydrate sources may be decreased in the gut and consumed directly or indirectly by enzymes or by cells of the present disclosure that express such fiber-synthesizing enzymes. In general, it is to be appreciated that any of, for example, glucose, sucrose, or fructose may be modified and/or processed to provide enzyme substrate carbohydrates for the production of any of a wide variety of fibers as disclosed herein, including particular fibers produced by particular enzymes as disclosed herein. Accordingly, in certain embodiments, fiber-synthesizing enzymes of the present disclosure can directly utilize gut carbohydrate and in certain embodiments can utilize modified and/or processed forms.

[0113] In some embodiments, fiber produced by a composition or method of the present disclosure is soluble levan fiber. Levan is a naturally occurring fructan present in certain plant and microorganism species. Rare fibers such as levan can favor a healthy and/or healthily

balanced microbiome or microbiome composition, e.g., by stimulating the growth and/or increasing the concentration and/or amount of various beneficial types of gut microbiome constituent microbes such as Bifidobacteria, Roseburia, and/or Eubacterium rectale. For at least these reasons, increasing levan can increase the concentration or amount of beneficial molecules such as butyrate and propionate. Fructans such as levan can be synthesized from sucrose. Levan is known to be abundantly produced when the bacterium *Bacillus subtilis* sp. *Natto* (*B.s. Natto*) is cultured on soy beans and *B.s. natto* can be used to make the food 'Nattō' (for which the bacterial strain is named). *B.s. natto* likely originated from the soil and is not a commensal of the human gut.

[0114] Levan fiber polymer is made up of fructose, a monosaccharide, conjugated in 2,6 beta glycosidic linkages. In some embodiments of the present disclosure, levan fiber is enzymatically synthesized by a process and/or enzyme that utilizes a fructose molecule liberated by enzymatic digestion of sucrose (which includes two monosaccharide moieties, one glucose and one fructose). In some embodiments, the present disclosure includes a levansucrase enzyme that synthesizes levan from sucrose. In various embodiments, a levansucrase enzyme can cleave sucrose into constituent monosaccharide moieties and incorporate the fructose monosaccharide moiety into levan. Accordingly, without wishing to be bound by theory, a levansucrase enzyme can convert 50% of the carbohydrate biomass and/or monosaccharide moieties, and/or 100% of the fructose biomass and/or fructose moieties, generated by levansucrase cleavage of sucrose to soluble fiber.

[0115] In various embodiments a fiber-synthesizing enzyme of the present disclosure can synthesize fiber from substrates including carbohydrate (e.g., monosaccharides and/or disaccharides) within or across a pH range of (e.g., in an environment characterized by a pH that is in the range of) 5.7 to 8.5. In certain embodiments, a fiber-synthesizing enzyme of the present disclosure can synthesize fiber at pH 5.7, pH 6.0, pH 7.4, and/or pH 8.5. In certain embodiments, a fiber-synthesizing enzyme of the present disclosure can synthesize fiber within or across a pH range of 5.7 to 6.0, 6.0 to 7.4, and/or 7.4 to 8.5, or within or across a range having a lower bound selected from about 5.5, 5.6, 5.7, 5.8, 5.9, 6.0, 6.1, 6.2, 6.3, 6.4, 6.5, 6.6, 6.7, 6.8, 6.9, and 7.0 and an upper bound selected from about 6.0, 6.1, 6.2, 6.3, 6.4, 6.5, 6.6, 6.7, 6.8, 6.9, 7.0, 7.1, 7.2, 7.3, 7.4, 7.5, 7.6, 7.7, 7.8, 7.9, 8.0, 8.1, 8.2, 8.3, 8.4, or 8.5. Without wishing to be

bound by any particularly scientific theory, the human gastrointestinal tract pH typically gradually increases in the small intestine from about pH 6 to about pH 7.4 in the terminal ileum, drops to about 5.7 in the caecum, and again gradually increases through the colon, reaching about pH 6.7 in the rectum. Those of skill in the art will appreciate that various established techniques can be used to modify polypeptide pH and/or peptidase resistance. In various embodiments, a fiber-synthesizing enzyme of the present disclosure is secluded and/or protected within a fiber matrix that prevents larger peptidases from engaging in degradation of the enzyme (while still allowing agents for fiber synthesis, such as carbohydrate, e.g., sucrose, to contact the enzyme). Those of skill in the art will further appreciate that coding sequences encoding a fiber-synthesizing enzyme of the present disclosure can be modified to reflect codon usage of a cell in which the enzyme is expressed.

[0116] In various embodiments, a polypeptide of the present disclosure (e.g., an enzyme, e.g., a levansucrase enzyme) is operably linked with a secretion polypeptide. Identification of secretion polypeptides is common in the art, e.g., by synthesis of a fusion polypeptides including signaling polypeptides of interest and screening the fusion polypeptides for secretion from an exemplary cell. Exemplary signal sequences can include, without limitation, those provided in Table 2.

Table 2: Secretion polypeptides

Name	Amino Acid Sequence
sacB	MNIKKFAKQATVLTFTTALLAGGATQAFA (SEQ ID NO: 55)
yvcE	MRKSLITLGLASVIGTSSFLIPFTSKTASA (SEQ ID NO: 56)
yoqM	MKLRKVL TGSVLSLGLLV SASPAFA (SEQ ID NO: 57)
yuaB	MKRKLLSSLAISLSLGLLV SAPTASFAAE (SEQ ID NO: 58)
pel	MKKVMLATALFLGLTPAGANA (SEQ ID NO: 59)
pelB	MKRLCLWFTVFSFLVLLPGKALG (SEQ ID NO: 60)
yoaW	MKKMLMLAFTFLLALTIHVGEASA (SEQ ID NO: 61)
yqxI	MFKKLLLATSFSLVPLDGHAKA (SEQ ID NO: 62)
lipA	MKFVKRRIIALVTILMLSVTSLFALQPSAKAA (SEQ ID NO: 63)
lipB	MKKVLMAFIICLSLISVLAAPPSGAKA (SEQ ID NO: 64)
yoqH	MKRFILVLSFLSIIVAYPIQTNA (SEQ ID NO: 65)
ybfO	MKRMIVRMTLPLLIVCLAFSSFSASARA (SEQ ID NO: 66)
sacB	MNIKKFAKQTVLTFTTALLAGGATQAFA (SEQ ID NO: 67)
bglS	MPYLKRVLLLLVTGLFMSLFAVTATASA (SEQ ID NO: 68)
yddT	MRKKRVITCVMAASLTGSLLPAGYASA (SEQ ID NO: 69)
yobB	MKIRKILLSSALSFGMLISAVPALA (SEQ ID NO: 70)

Those of skill in the art will further appreciate that further signal sequences are present in the *Bacillus subtilis* genome and can be included in transgene and enzymes of the present disclosure. Of the three known secretory routes in *B. subtilis*, the Sec pathway directs the majority of secretory proteins into the growth medium. Alternatively, a small number of exoproteins with specific functions are secreted via the Tat pathway or ABC. Export by the Sec pathway can be achieved by inclusion in a polypeptide of a hydrophobic secretion polypeptide at the N-terminus of the secreted polypeptide, which secretion polypeptide can be about 20 amino acids in length and can include, in some examples, 3 regions: a positively charged amino terminal, a

hydrophobic core, and a polar carboxyl-terminal. A commercial kit (Takara Bio) is available to test with limited effort 173 distinct secretion polypeptides. In certain Examples provided herein, a fusion polypeptide includes a levansucrase enzyme operatively linked to a secretion polypeptide derived from a sequence naturally associated with SacB.

[0117] Expression of a fiber-synthesizing enzyme engineered to include a secretion polypeptide (also known in the art as a secretion signal) can result in secretion of the enzyme (e.g., secretion from cells including a transgene encoding the fiber-synthesizing enzyme, e.g., into gut of a subject to which the cells were administered). The present disclosure includes the recognition that, in various embodiments, synthesis of fiber from carbohydrate in the gut can advantageously occur extracellularly with respect to administered engineered cells of the present disclosure. For example, extracellular synthesis of fiber by a fiber-synthesizing enzyme delivered to the gut by engineered cells of the present disclosure can obviate the need to import sucrose across the cell membrane into the cytosol of engineered cells, resulting in higher enzyme activity than would be achieved by intracellular expression of enzyme, and therefore faster and/or greater consumption of sugar and/or synthesis of fiber. Moreover, because comparatively higher sucrose concentrations can cause production of comparatively higher molecular weight fiber molecules, synthesis of fiber in the gut where carbohydrate concentration is maximal can increase average fiber molecular weight.

[0118] Various methods are known in the art for selecting a secretion polypeptide for pairing (e.g., fusing or otherwise operably linking) with a polypeptide or nucleic acid sequence encoding the polypeptide. For example, various commercial kits such as that from Takara Bio (3380 - B. Subtilis Secretory Protein Expression System) are available for rapid screening of secretion polypeptides. Those of skill in the art will further appreciate that laboratory methods can be used to select or evolve (e.g., by random evolution) variants of fusion polypeptides that include a fiber-synthesizing enzyme of the present disclosure and a secretion polypeptide, where the variants can have increased stability and/or expression as compared to a reference fusion polypeptide (e.g., fusion polypeptide prior to selection or evolution). In various embodiments, a secretion polypeptide of the present disclosure can be selected based on the cell type in which an engineered polypeptide (e.g., a fusion polypeptide) is expressed.

[0119] In various embodiments, a fiber-synthesizing enzyme of the present disclosure is not secreted. Accordingly, in some embodiments a fiber-synthesizing enzyme is engineered for secretion such that the enzyme acts directly on substrates present in the gut, and in some embodiments a fiber-synthesizing enzyme is engineered such that it is not secreted and acts on substrates internalized by a cell that expresses a fiber-synthesizing enzyme. Exemplary secreted enzymes can include a sucrose isomerase, trehalulose synthase, inulinsucrase, levanosucrase, beta-galactosidase, or b-D-Galactosidase. Accordingly, certain such secreted fiber-synthesizing enzymes can utilize substrate present in gut and decrease the concentration or amount of that substrate in gut (e.g., the substrate sucrose for sucrose isomerase, trehalulose synthase, inulinsucrase, or levanosucrase, or the substrate lactose for beta-galactosidase, or b-D-Galactosidase). Exemplary fiber-synthesizing enzymes that in various embodiments are not secreted and synthesize fiber at or in cells of the present disclosure (e.g., through cytosolic or membrane activity) can include a 1,3-beta-glucan synthase, 1,3;1,4-beta-D-glucan synthase, 1,3-beta-glucan synthase, 1,6-beta-glucan synthase, Trehalose-6-phosphate synthase, 1,6-alpha-galactosyltransferase, α -1,4-galactosyltransferase, Alpha-1,2-fucosyltransferase, cellulose synthase, maltose synthase, starch synthase, Starch-branching enzyme, Glycogen synthase, galacturonosyltransferase, or Chitinoligosaccharide synthase. Accordingly, various non-secreted fiber-synthesizing enzymes present in the cytosol or membranes of cells of the present disclosure can act on substrates internalized by cells and optionally modified by cells prior to use as a substrate in fiber synthesis by the enzyme. Fiber-synthesizing enzymes such as 1,3-beta-glucan synthase, 1,3;1,4-beta-D-glucan synthase, 1,3-beta-glucan synthase, 1,6-beta-glucan synthase, Trehalose-6-phosphate synthase, cellulose synthase, maltose synthase, starch synthase, and Starch-branching enzyme can utilize UDP-glucose as a substrate where activity of the enzyme decreases the amount or concentration of glucose, sucrose, and/or fructose in gut, in that glucose, sucrose, and/or fructose carbohydrates taken up by cells are naturally converted to UDP-glucose that provide a substrate for the fiber-synthesizing enzyme. Fiber-synthesizing enzymes such as Glycogen synthase can utilize ADP- α -D-glucose as a substrate where activity of the enzyme decreases the amount or concentration of glucose, sucrose, and/or fructose in gut, in that glucose, sucrose, and/or fructose carbohydrates taken up by cells are naturally converted to ADP- α -D-glucose that provides a substrate for the fiber-synthesizing enzyme. Fiber-synthesizing enzymes

such as α -1,4-galactosyltransferase can utilize UDP-galactose as a substrate where activity of the enzyme decreases the amount or concentration of lactose in gut, in that lactose carbohydrates taken up by cells are naturally converted to UDP-galactose that provides a substrate for the fiber-synthesizing enzyme. Fiber-synthesizing enzymes such as Alpha-1,2-fucosyltransferase can utilize UDP-galactose and/or UDP-fucose as a substrate where activity of the enzyme decreases the amount or concentration of lactose in gut, in that lactose carbohydrates taken up by cells are naturally converted to utilize UDP-galactose and/or UDP-fucose that provide a substrate for the fiber-synthesizing enzyme. Fiber-synthesizing enzymes such as 1,6-alpha-galactosyltransferase can utilize lactose and/or galactinol as a substrate where activity of the enzyme decreases the amount or concentration of lactose in gut, in that lactose carbohydrates taken up by cells are naturally converted to lactose and/or galactinol that provide a substrate for the fiber-synthesizing enzyme. Fiber-synthesizing enzymes such as galacturonosyltransferase can utilize various molecules as a substrate where activity of the enzyme decreases the amount or concentration of glucose, sucrose, and/or fructose in gut, in that glucose, sucrose, and/or fructose carbohydrates taken up by cells are naturally converted to substrates of the fiber-synthesizing enzyme. Fiber-synthesizing enzymes such as Chitinoligosaccharide synthase can utilize UDP-N-acetyl-alpha-D-glucosamine as a substrate where activity of the enzyme decreases the amount or concentration of glucose, sucrose, and/or fructose in gut, in that glucose, sucrose, and/or fructose carbohydrates taken up by cells are naturally converted to UDP-N-acetyl-alpha-D-glucosamine that provide a substrate for the fiber-synthesizing enzyme.

[0120] The present disclosure generally includes, among other things, the innovation of administering to a subject an engineered cells that encodes and/or expresses a fiber-synthesizing enzyme of the present disclosure. Those of skill in the art will further appreciate that, in various embodiments, a nucleic acid sequence encoding a fiber-synthesizing enzyme of the present disclosure can be operably linked with a regulatory nucleic acid sequence that controls expression of the enzyme in the cell. Moreover, as those of skill in the art will appreciate, a wide variety of regulatory nucleic acid sequences for operable linkage to coding sequences and/or for expression of encoded polypeptides are known in the art.

[0121] In some embodiments, a regulatory nucleic acid can be a constitutive promoter and/or cause constitutive expression of a polypeptide encoded by an operably linked coding

sequence, e.g., in bacteria (e.g., in *B. subtilis*). In some embodiments, a regulatory nucleic acid can be a conditional promoter and/or cause conditional expression of a polypeptide encoded by an operably linked coding sequence, e.g., in bacteria (e.g., in *B. subtilis*). In various embodiments, a bacterial regulatory nucleic acid sequence can be a promoter selected from T7, T7lac, Sp6, araBAD, trp, lac, Ptac, or pL. In some embodiments a regulatory nucleic acid sequence can be a *B. subtilis* promoter is selected from Pveg, PserA, PymdA, PfbA, Pzwf, PfoleA, Ppgi, PlepA, or PreLA, e.g., as set forth in Guiziou (2016 *Nucl. Acids Res.* 44(15): 7495-7508), which is herein incorporated by reference with respect to *B. subtilis* promoters and in its entirety.

[0122] In various embodiments, an exemplary regulatory nucleic acid can be based on regulator nucleic acid sequence(s) associated with expression of, e.g., flagellar machinery. Such regulatory nucleic acid sequences have been previously adapted for constitutive and/or robust expression of polypeptide encoded by an operably linked coding nucleic acid sequence. Various bacteria (e.g., *B. subtilis*) regulate expression of a flagellin gene locus (in some systems referred to as *hag*) by a regulatory system that includes positive and negative control. Removal and/or reduction of negative regulation of flagellin gene locus expression and/or addition of or increased positive regulation of flagellin gene locus expression can be utilized to increase expression from a flagellin promoter. Typically, without wishing to be bound by any particular scientific theory, expression of the flagellin gene locus is positively controlled by a sigma factor, SigD, which is repressed by the FlgM protein. Deletion of *flgM* and/or reduction of FlgM expression significantly enhances expression and/or activity of SigD, and consequently results in increased and/or constitutive expression of the flagellin gene locus. Typically, without wishing to be bound by any particular scientific theory, expression of flagellar subunits can be negatively, post-transcriptionally controlled by binding of a protein referred to as CsrA to a ribosome binding site present in flagellin gene locus transcripts. Deletion of *csrA* and/or reduction of CsrA expression significantly enhances expression of flagellar subunits. Deletion and/or reduction of one or more CsrA binding sites in flagellin gene regulatory nucleic acid sequence(s) can additionally or alternatively significantly increases expression of flagellar subunits. In various embodiments, certain point mutations in a CsrA binding site can reduce and/or abrogate negative control of flagellar subunit expression by CsrA. Accordingly, in various embodiments,

expression of a nucleic acid operably linked with a flagellin gene regulatory nucleic acid sequence (e.g., a flagellin promoter, e.g., a *hag* promoter) can be increased by either or both of (i) deleting the *flgM* gene and/or (ii) a point mutation in the CsrA-binding site, e.g., in *B. subtilis*. Those of skill in the art will therefore appreciate that, if the coding nucleic acid sequence of the flagellin gene is replaced with a nucleic acid sequence encoding a fiber-synthesizing enzyme (e.g., a levansucrase enzyme) such that the fiber-synthesizing enzyme is operably linked to flagellin gene regulatory sequence(s) (e.g., including the *hag* promoter), the fiber-synthesizing enzyme can be expressed at high and/or constitutive levels. In various embodiments, the fiber-synthesizing enzyme includes a secretion polypeptide.

[0123] In certain embodiments, a transgene of the present disclosure includes a nucleic acid sequence that encodes a levansucrase enzyme operably linked with a flagellin gene regulatory nucleic acid sequence, e.g., a flagellin gene promoter. In certain embodiments, a transgene of the present disclosure can include a nucleic acid sequence that encodes a levansucrase enzyme is operably linked with a *hag* promoter. In various embodiments, the levansucrase enzyme includes a secretion polypeptide.

[0124] Those of skill in the art will appreciate that endogenous *sacB* loci can include regulatory elements that repress *sacB* expression in the absence of sucrose. In the digestive tract, residence time of digesta can be shorter than the time required for induction of expression by consumed carbohydrate (e.g., sucrose), as the residence time can be less than 20 minutes. Moreover, endogenous gut enzymes can also breakdown carbohydrates. The present disclosure therefore includes embodiments in which fiber-synthesizing enzyme is constitutively expressed to increase the overlap of digesta residence in gut and fiber-synthesizing enzyme activity in gut. Constitutive expression of fiber-synthesizing enzyme in gut can result in accumulation of fiber-synthesizing enzyme available for synthesis of fiber from carbohydrate, e.g., digestion during consumption and digestion.

[0125] The present disclosure further includes nucleic acid vectors that include transgenes of the present disclosure. Those of skill in the art will appreciate that a wide variety of vectors, such as plasmids, are available for introduction of heterologous nucleic acids into target cells, e.g., bacterial target cells disclosed herein, e.g., with integration of a heterologous nucleic acid sequence into the genome of target cells. Exemplary vectors can include plasmids

and viral vectors. In various embodiments vectors include a selectable marker and/or counter-selectable marker. In some embodiments, a vector is engineered for integration of heterologous nucleic acid sequence into a target cell genome by homologous recombination. In some embodiments, a vector is engineered for unstable integration of a heterologous nucleic acid sequence into a target cell genome (e.g., in some embodiments, by production of merodiploid cells). In some embodiments, a vector is engineered to integrate into a host cell genome by a single crossover event subsequently resolved by subsequent double-crossover resulting in loss of plasmid DNA, such as those made from pminiMAD or pMAD plasmids.

Engineered Cells

[0126] The present disclosure includes cells (e.g., spores) that include a heterologous nucleic acid sequence that includes a transgene that encodes and/or expresses a fiber-synthesizing enzyme of the present disclosure (such cells referred to herein as engineered cells). Those of skill in the art will appreciate that any of a wide variety of cells and/or spores can be engineered to produce, and/or used in, a composition, e.g., for administration to a subject, e.g., a human subject.

[0127] In various embodiments, engineered cells of the present disclosure can be vegetative cells and/or non-spore forming. Many cell preparations for administration to human subjects, e.g., as probiotics, include bacteria that are not capable of forming spores and/or are administered as vegetative cells (e.g., in lyophilized preparations). Common examples of non-spore forming cells that can be administered as vegetative cells include lactobacilli and bifidobacteria. *Lactobacillus* cells are commonly used. Examples of non-spore forming cells that can be administered as vegetative cells include but are not limited to members of the Bacteroidetes family (e.g. *Bacteroides uniformis*, *Bacteroides thetaiotaomicron*, *Bacteroides fragilis*, *Bacteroides ovatus*, *Bacteroides vulgatus*, *Bacteroides caccae*, *Bacteroides eggerthii*, *Bacteroides vulgatus*, and *Parabacteroides distasonis*) and members of the *Lactobacillae* family (e.g., *Lactobacillus paracasei*, *Lactobacillus brevis*, *Lactobacillus casei*, *Lactobacillus delbrueckii*, *Lactobacillus gasseri*, *Lactobacillus rhamnosus*, *Lactobacillus reuteri*, *Lactobacillus plantarum*, and, *Lactobacillus johnsonii*).

[0128] In various embodiments, cells of the present disclosure include commensal bacteria such as *Lactobacilli*, *clostridia*, or *bacteroides*, or other cells recognized as edible in live form, such as *Saccharomyces cerevisiae*.

[0129] In various embodiments a cell is 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*. The present disclosure includes use of bacteria of the genus *Bacillus*, e.g., *B. subtilis*. 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*.

[0130] In certain embodiments, a cell is a bacterial cell that can form a spore. A spore can be a resilient form of a cell. Bacterial spores can be dormant living forms that can exist in a desiccated and/or dehydrated state for significant periods of time and/or indefinitely. Without wishing to be bound by any particular scientific theory, the natural life cycle of various spore forming bacteria can includes germination of the spore in the presence of environmental nutrients and/or factors, proliferation of germinated cells, and re-sporulation in the event of environmental nutrient exhaustion. In various embodiments, a spore can germinate and/or proliferate after administration to a subject, e.g., in the gut of a subject, e.g., in the intestine of a subject. Various spore-forming bacteria are known in the art and can be engineered to include, e.g., a heterologous nucleic acid including a transgene of the present disclosure. In various embodiments, an engineered cell is a cell that is in a spore state. Exemplary engineered spores of the present disclosure include engineered spores of *Bacillus* species, e.g., for administration to a subject.

[0131] A spore can be a useful cell form in the context of the present disclosure at least in part because spores can retain ability to germinate and/or proliferate after a period of storage, e.g., a long period of storage, e.g., at room temperature, and for at least this reason spores are considered to have a long useful shelf life. Shelf-life can be an important factor in safety, satisfaction, and price point for consumer products making it a valuable asset both for the provider and consumer. Additionally, the resilient spore form allows the live organism to traverse the environment of the stomach which is widely known to inactivate enzymes and active

molecules not otherwise formulated to withstand traversal of the stomach. Spore form therefore enables formulation of engineered cells in forms that are pleasing to consumers. For example, while methods exist for formulating pills to enable purified and formulated therapeutic molecules to traverse the stomach, formulations such as liquid formulations (e.g., based at least in part by use of engineered spores) can also be produced and may be preferred by some consumers. Administration of spores obviates the need for certain costly formulation steps. Additionally, administration of spores can provide a staggered release of active cells from a dormant state. Thus, the present disclosure provides, among other things, compositions and methods for extended release of a therapeutic agent in the gut, where the composition includes an engineered cell, e.g., a spore form of an engineered cell. Due to their small size and resilience to heat, spores can also facilitate and/or enable lyophilization at a faster rate than comparable non-spore cells could be lyophilized with a comparable cell survival.

[0132] Those of skill in the art will appreciate that germination from spores can follow exponential decay, e.g., modelled by a constant percentage of spores germinating per minute and defined by the availability of nutrient triggers. Germination can result in a steady release of active live cells over a number of hours, which can further contribute to the period of time over which a dose is effective and/or delivers fiber-synthesizing enzyme to gut. Furthermore, active cells can re-enter a spore state in the absence of sufficient nutrients. Properties of germination and re-sporulation individual and together can allow administered cells to reside in the gut in a resilient spore form until contacted with digesta.

[0133] Spore forming bacteria and spores thereof include, for example, *Bacillus subtilis* PY79 and spores thereof. Other spore forming microbes and spores thereof include but are not limited to *Bacillus subtilis* group members (e.g. *Bacillus subtilis* subsp. *subtilis* str. 168, and *Bacillus subtilis* subsp. *Natto*), *Bacillus* group members (e.g. *Bacillus coagulans*, *Bacillus amyloliquefaciens*, *Bacillus licheniformis*, *Bacillus sp.*, *Bacillus megatarium*, and *Bacillus subtilis* subsp. *Spizizenii*), *Clostridiaceae* family members (e.g., *Clostridium butyricum*, *Clostridium acetobutylicum*, *Clostridium arbusti*, *Clostridium aurantibutyricum*, *Clostridium beijerinckii*, *Clostridium cellulovorans*, *Clostridium cellulolyticum*, *Clostridium thermocellum*, *Clostridium thermobutyricum*, *Clostridium pasteurianum*, *Clostridium khyveri*, *Clostridium novyi*, *Clostridium saccharobutylicum*, *Clostridium thermosuccinogenes*, *Clostridium*

thermopalmarium, *Clostridium saccharolyticum*, *Clostridium saccharoperbutylacetonicum*, *Clostridium tyrobutyricum*, *Clostridium tetanomorphum*, *Clostridium magnum*, *Clostridium ljungdahlii*, *Clostridium autoethanogenum*, *Clostridium butyricum*, *Clostridium puniceum*, *Clostridium diolis*, *Clostridium 5 homopropionicum* and/or *Clostridium roseum*), and various edible yeasts such as *Saccharomyces cerevisiae*, and spores thereof.

[0134] In various embodiments, spore forming microbes and spores thereof include but are not limited to *Bacillus subtilis* group members *Bacillus subtilis* A29, *Bacillus subtilis* Abs3, *Bacillus subtilis* AP254, *Bacillus subtilis* At2, *Bacillus subtilis* B2, *Bacillus subtilis* B7-s, *Bacillus subtilis* BEST7003, *Bacillus subtilis* BEST7613, *Bacillus subtilis* BSn5, *Bacillus subtilis* E1, *Bacillus subtilis* gP20b, *Bacillus subtilis* Hal1, *Bacillus subtilis* HJ5, *Bacillus subtilis* J22, *Bacillus subtilis* J23, *Bacillus subtilis* J24, *Bacillus subtilis* J25, *Bacillus subtilis* J26, *Bacillus subtilis* J27, *Bacillus subtilis* KCTC 1028 = ATCC 6051a, *Bacillus subtilis* LX-8, *Bacillus subtilis* MB73/2, *Bacillus subtilis* Miyagi-4, *Bacillus subtilis* Miyagi-4100, *Bacillus subtilis* PRO 1, *Bacillus subtilis* PS216, *Bacillus subtilis* PTS-394, *Bacillus subtilis* PY79, *Bacillus subtilis* QB928, *Bacillus subtilis* QH-1, *Bacillus subtilis* S1-4, *Bacillus subtilis* SPZ1, *Bacillus subtilis* str. 10, *Bacillus subtilis* strain DY, *Bacillus subtilis* subsp. *amylosacchariticus*, *Bacillus subtilis* subsp. *chungkookjang*, *Bacillus subtilis* subsp. *endophyticus*, *Bacillus subtilis* subsp. *globigii*, *Bacillus subtilis* subsp. *inaquosorum*, *Bacillus subtilis* subsp. *inaquosorum* KCTC 13429, *Bacillus subtilis* subsp. *krietiensis*, *Bacillus subtilis* subsp. *lactipan*, *Bacillus subtilis* subsp. *natto*, *Bacillus subtilis* subsp. *natto* BEST195, *Bacillus subtilis* subsp. *natto* HSF 1410, *Bacillus subtilis* subsp. *niger*, *Bacillus subtilis* subsp. *qingdao*, *Bacillus subtilis* subsp. *sadata*, *Bacillus subtilis* subsp. *spizizenii*, *Bacillus subtilis* subsp. *spizizenii* ATCC 6633, *Bacillus subtilis* subsp. *spizizenii* DV1-B-1, *Bacillus subtilis* subsp. *spizizenii* JCM 2499, *Bacillus subtilis* subsp. *spizizenii* RFWG1A3, *Bacillus subtilis* subsp. *spizizenii* RFWG1A4, *Bacillus subtilis* subsp. *spizizenii* RFWG4C10, *Bacillus subtilis* subsp. *spizizenii* RFWG5B15, *Bacillus subtilis* subsp. *spizizenii* str. W23, *Bacillus subtilis* subsp. *spizizenii* TU-B-10, *Bacillus subtilis* subsp. *stercoris*, *Bacillus subtilis* subsp. *subtilis*, *Bacillus subtilis* subsp. *subtilis* NCIB 3610 = ATCC 6051, *Bacillus subtilis* subsp. *subtilis* 6051-HGW, *Bacillus subtilis* subsp. *subtilis* str. 168, *Bacillus subtilis* subsp. *subtilis* str. L170, *Bacillus subtilis* subsp. *subtilis* str. N170, *Bacillus subtilis* subsp. *subtilis* str. AG1839, *Bacillus subtilis* subsp. *subtilis* str. AUS198, *Bacillus subtilis*

subsp. *subtilis* str. B2, *Bacillus subtilis* subsp. *subtilis* str. BAB-1, *Bacillus subtilis* subsp. *subtilis* str. BSP1, *Bacillus subtilis* subsp. *subtilis* str. JH642, *Bacillus subtilis* subsp. *subtilis* str. JH642 substr. AG174, *Bacillus subtilis* subsp. *subtilis* str. MP11, *Bacillus subtilis* subsp. *subtilis* str. MP9, *Bacillus subtilis* subsp. *subtilis* str. OH 131.1, *Bacillus subtilis* subsp. *subtilis* str. R0179, *Bacillus subtilis* subsp. *subtilis* str. RO-NN-1, *Bacillus subtilis* subsp. *subtilis* str. SC-8, *Bacillus subtilis* subsp. *subtilis* str. SMY, *Bacillus subtilis* TO-A, *Bacillus subtilis* TPK 210909, *Bacillus subtilis* UCMB5014, *Bacillus subtilis* XF-1, and *Bacillus subtilis* YF001.

[0135] In various embodiments, spore forming microbes and spores thereof include but are not limited to *Bacillus subtilis* group members *Bacillus subtilis* Group *Bacillus* amyloliquefaciens, *Bacillus amyloliquefaciens* AB01, *Bacillus amyloliquefaciens* AP143, *Bacillus amyloliquefaciens* AP193, *Bacillus amyloliquefaciens* AP71, *Bacillus amyloliquefaciens* AP79, *Bacillus amyloliquefaciens* CC178, *Bacillus amyloliquefaciens* DC-12, *Bacillus amyloliquefaciens* DSM 7 = ATCC 23350, *Bacillus amyloliquefaciens* EBL11, *Bacillus amyloliquefaciens* EGD-AQ14, *Bacillus amyloliquefaciens* GGI-18, *Bacillus amyloliquefaciens* HB-26, *Bacillus amyloliquefaciens* IT-45, *Bacillus amyloliquefaciens* KHG19, *Bacillus amyloliquefaciens* LFB112, *Bacillus amyloliquefaciens* LL3, *Bacillus amyloliquefaciens* Lx-11, *Bacillus amyloliquefaciens* PGK1, *Bacillus amyloliquefaciens* TA208, *Bacillus amyloliquefaciens* UASWS BA1, *Bacillus amyloliquefaciens* UCMB5007, *Bacillus amyloliquefaciens* UCMB5140, *Bacillus amyloliquefaciens* UMAF6614, *Bacillus amyloliquefaciens* UMAF6639, *Bacillus amyloliquefaciens* XH7, *Bacillus amyloliquefaciens* Y2, *Bacillus siamensis*, *Bacillus siamensis* KCTC 13613, *Bacillus velezensis*, *Bacillus velezensis* A3, *Bacillus velezensis* AS43.3, *Bacillus velezensis* At1, *Bacillus velezensis* CAU B946, *Bacillus velezensis* FZB42, *Bacillus velezensis* M27, *Bacillus velezensis* NAU-B3, *Bacillus velezensis* NJN-6, *Bacillus velezensis* SK19.001, *Bacillus velezensis* SQR9, *Bacillus velezensis* TrigoCor1448, *Bacillus velezensis* UCMB5033, *Bacillus velezensis* UCMB5036, *Bacillus velezensis* UCMB5113, *Bacillus velezensis* variant polyfermenticus, *Bacillus velezensis* YAU B9601-Y2, *Bacillus atrophaeus*, *Bacillus atrophaeus* 1013-1, *Bacillus atrophaeus* 1013-2, *Bacillus atrophaeus* 1942, *Bacillus atrophaeus* BACI051-E, *Bacillus atrophaeus* BACI051-N, *Bacillus atrophaeus* C89, *Bacillus atrophaeus* Detrick-1, *Bacillus atrophaeus* Detrick-2, *Bacillus atrophaeus* Detrick-3, *Bacillus atrophaeus* NBRC 15539, *Bacillus atrophaeus* str. Dugway,

Bacillus atrophaeus subsp. *globigii*, *Bacillus atrophaeus* ATCC 49822, *Bacillus atrophaeus* ATCC 9372, *Bacillus atrophaeus* UCMB-5137, *Bacillus licheniformis*, *Bacillus licheniformis* 10-1-A, *Bacillus licheniformis* 5-2-D, *Bacillus licheniformis* CG-B52, *Bacillus licheniformis* CGMCC 3963, *Bacillus licheniformis* DSM 13 = ATCC 14580, *Bacillus licheniformis* F1-1, *Bacillus licheniformis* F2-1, *Bacillus licheniformis* KRB2009, *Bacillus licheniformis* LMG 17339, *Bacillus licheniformis* LMG 6934, *Bacillus licheniformis* LMG 7559, *Bacillus licheniformis* MKU3, *Bacillus licheniformis* S 16, *Bacillus licheniformis* WX-02, *Bacillus mojavenensis* subgroup, *Bacillus halotolerans*, *Bacillus mojavenensis*, *Bacillus mojavenensis* RO-H-1 = KCTC 3706, *Bacillus mojavenensis* RRC 101, *Bacillus paralicheniformis*, *Bacillus paralicheniformis* ATCC 9945a, *Bacillus paralicheniformis* G-1, *Bacillus sonorensis*, *Bacillus sonorensis* L12, *Bacillus sonorensis* NBRC 101234 = KCTC 13918, *Bacillus subtilis*, *Bacillus subtilis* A29, *Bacillus subtilis* Abs3, *Bacillus subtilis* AP254, *Bacillus subtilis* At2, *Bacillus subtilis* B2, *Bacillus subtilis* B7-s, *Bacillus subtilis* BEST7003, *Bacillus subtilis* BEST7613, *Bacillus subtilis* BSn5, *Bacillus subtilis* E1, *Bacillus subtilis* gtP20b, *Bacillus subtilis* Hal1, *Bacillus subtilis* HJ5, *Bacillus subtilis* J22, *Bacillus subtilis* J23, *Bacillus subtilis* J24, *Bacillus subtilis* J25, *Bacillus subtilis* J26, *Bacillus subtilis* J27, *Bacillus subtilis* KCTC 1028 = ATCC 6051a, *Bacillus subtilis* LX-8, *Bacillus subtilis* MB73/2, *Bacillus subtilis* Miyagi-4, *Bacillus subtilis* Miyagi-4100, *Bacillus subtilis* PRO 1, *Bacillus subtilis* PS216, *Bacillus subtilis* PTS-394, *Bacillus subtilis* PY79, *Bacillus subtilis* QB928, *Bacillus subtilis* QH-1, *Bacillus subtilis* S1-4, *Bacillus subtilis* SPZ1, *Bacillus subtilis* str. 10, *Bacillus subtilis* strain DY, *Bacillus subtilis* subsp. *amylosacchariticus*, *Bacillus subtilis* subsp. *chungkookjang*, *Bacillus subtilis* subsp. *endophyticus*, *Bacillus subtilis* subsp. *globigii*, *Bacillus subtilis* subsp. *inaquosorum*, *Bacillus subtilis* subsp. *inaquosorum* KCTC 13429, *Bacillus subtilis* subsp. *krietiensis*, *Bacillus subtilis* subsp. *lactipan*, *Bacillus subtilis* subsp. *natto*, *Bacillus subtilis* subsp. *natto* BEST195, *Bacillus subtilis* subsp. *natto* HSF 1410, *Bacillus subtilis* subsp. *niger*, *Bacillus subtilis* subsp. *qingdao*, *Bacillus subtilis* subsp. *sadata*, *Bacillus subtilis* subsp. *spizizenii*, *Bacillus subtilis* subsp. *spizizenii* ATCC 6633, *Bacillus subtilis* subsp. *spizizenii* DV1-B-1, *Bacillus subtilis* subsp. *spizizenii* JCM 2499, *Bacillus subtilis* subsp. *spizizenii* RFWG1A3, *Bacillus subtilis* subsp. *spizizenii* RFWG1A4, *Bacillus subtilis* subsp. *spizizenii* RFWG4C10, *Bacillus subtilis* subsp. *spizizenii* RFWG5B15, *Bacillus subtilis* subsp. *spizizenii* str. W23, *Bacillus subtilis* subsp.

spizizenii TU-B-10, *Bacillus subtilis* subsp. *stercoris*, *Bacillus subtilis* subsp. *subtilis*, *Bacillus subtilis* subsp. *subtilis* NCIB 3610 = ATCC 6051, *Bacillus subtilis* subsp. *subtilis* str. 168, *Bacillus subtilis* subsp. *subtilis* str. AG1839, *Bacillus subtilis* subsp. *subtilis* str. AUSI98, *Bacillus subtilis* subsp. *subtilis* str. B2, *Bacillus subtilis* subsp. *subtilis* str. BAB-1, *Bacillus subtilis* subsp. *subtilis* str. BSP1, *Bacillus subtilis* subsp. *subtilis* str. JH642, *Bacillus subtilis* subsp. *subtilis* str. MP11, *Bacillus subtilis* subsp. *subtilis* str. MP9, *Bacillus subtilis* subsp. *subtilis* str. OH 131.1, *Bacillus subtilis* subsp. *subtilis* str. R0179, *Bacillus subtilis* subsp. *subtilis* str. RO-NN-1, *Bacillus subtilis* subsp. *subtilis* str. SC-8, *Bacillus subtilis* subsp. *subtilis* str. SMY, *Bacillus subtilis* TO-A, *Bacillus subtilis* TPK 210909, *Bacillus subtilis* UCMB5014, *Bacillus subtilis* XF-1, *Bacillus subtilis* YF001, *Bacillus tequilensis*, *Bacillus tequilensis* KCTC 13622, *Bacillus vallismortis*, *Bacillus vallismortis* DV1-F-3, *Bacillus vallismortis* NRRL B-14890, *Bacillus* sp. AB01, *Bacillus* sp. AP102, *Bacillus* sp. AP143, *Bacillus* sp. AP183, *Bacillus* sp. AP189, *Bacillus* sp. AP193, *Bacillus* sp. AP215, *Bacillus* sp. AP218, *Bacillus* sp. AP219, *Bacillus* sp. AP295, *Bacillus* sp. AP301, *Bacillus* sp. AP303, *Bacillus* sp. AP305, *Bacillus* sp. AP71, *Bacillus* sp. AP77, *Bacillus* sp. AP79, *Bacillus* sp. D10(2019), *Bacillus* sp. D9(2019), *Bacillus* sp. GeS7V, *Bacillus* sp. LiF4a, *Bacillus* sp. LJF-10, *Bacillus* sp. LJF-11, *Bacillus* sp. LJF-2, *Bacillus* sp. LJF-5, *Bacillus* sp. LJF-8, *Bacillus* sp. LJF-9, *Bacillus* sp. M1(2010), *Bacillus* sp. M100(2010), *Bacillus* sp. M101(2010), *Bacillus* sp. M103(2010), *Bacillus* sp. M110(2010), *Bacillus* sp. M112(2010), *Bacillus* sp. M14(2010), *Bacillus* sp. M2(2010), *Bacillus* sp. M21(2010), *Bacillus* sp. M22(2010), *Bacillus* sp. M25(2010), *Bacillus* sp. M29(2010), *Bacillus* sp. M30(2010), *Bacillus* sp. M35(2010), *Bacillus* sp. M38(2010), *Bacillus* sp. M4(2010b), *Bacillus* sp. M40(2010), *Bacillus* sp. M5(2010), *Bacillus* sp. M63(2010), *Bacillus* sp. M64(2010), *Bacillus* sp. M67(2010), *Bacillus* sp. M68(2010), *Bacillus* sp. M88(2010), *Bacillus* sp. M90(2010), *Bacillus* sp. M94(2010), *Bacillus* sp. M95(2010), *Bacillus* sp. M98(2010), *Bacillus* sp. SE-54, *Bacillus* sp. TT106(2010), *Bacillus* sp. TT45(2010), *Bacillus* sp. TT46(2010), *Bacillus* sp. TT53(2010), *Bacillus* sp. TT80(2010), *Bacillus* sp. ZYJ-1, *Bacillus* sp. ZYJ-10, *Bacillus* sp. ZYJ-11, *Bacillus* sp. ZYJ-12, *Bacillus* sp. ZYJ-13, *Bacillus* sp. ZYJ-15, *Bacillus* sp. ZYJ-16, *Bacillus* sp. ZYJ-17, *Bacillus* sp. ZYJ-18, *Bacillus* sp. ZYJ-19, *Bacillus* sp. ZYJ-2, *Bacillus* sp. ZYJ-20, *Bacillus* sp. ZYJ-21, *Bacillus* sp. ZYJ-22, *Bacillus* sp. ZYJ-23, *Bacillus* sp. ZYJ-24, *Bacillus* sp. ZYJ-26, *Bacillus* sp. ZYJ-27, *Bacillus* sp. ZYJ-28, *Bacillus* sp.

ZYJ-3, *Bacillus* sp. ZYJ-32, *Bacillus* sp. ZYJ-33, *Bacillus* sp. ZYJ-34, *Bacillus* sp. ZYJ-35, *Bacillus* sp. ZYJ-36, *Bacillus* sp. ZYJ-37, *Bacillus* sp. ZYJ-38, *Bacillus* sp. ZYJ-39, *Bacillus* sp. ZYJ-4, *Bacillus* sp. ZYJ-40, *Bacillus* sp. ZYJ-43, *Bacillus* sp. ZYJ-44, *Bacillus* sp. ZYJ-5, *Bacillus* sp. ZYJ-6, *Bacillus* sp. ZYJ-7, *Bacillus* sp. ZYJ-8, *Bacillus* sp. ZYJ-9, *Bacillus subtilis* group sp., and *Nocardia* sp. KY2-1.

[0136] In various embodiments, spore forming microbes and spores thereof include but are not limited to *Bacillus coagulans* bacterium, optionally selected from *Bacillus coagulans* 2-6, *Bacillus coagulans* 36D1, *Bacillus coagulans* CSIL1, *Bacillus coagulans* DSM 1 = ATCC 7050, *Bacillus coagulans* H-1, *Bacillus coagulans* P38, *Bacillus coagulans* XZL4, and *Bacillus coagulans* XZL9. In various embodiments, spore forming microbes and spores thereof include but are not limited to a *Lactobacillus* bacterium, optionally selected from *Lactobacillus paracasei*, *Lactobacillus brevis*, *Lactobacillus casei*, *Lactobacillus delbrueckii*, *Lactobacillus gasseri*, *Lactobacillus rhamnosus*, *Lactobacillus reuteri*, *Lactobacillus plantarum*, and *Lactobacillus johnsonii*. In various embodiments, spore forming microbes and spores thereof include but are not limited to a *Lactobacillus* bacterium, optionally selected from *Lactobacillus paracasei* (*Lactobacillus paracasei* ATCC 334, *Lactobacillus paracasei* COM0101, *Lactobacillus paracasei* N1115, *Lactobacillus paracasei* NRIC 0644, *Lactobacillus paracasei* NRIC 1917, *Lactobacillus paracasei* NRIC 1981, *Lactobacillus paracasei* subsp. *paracasei*, *Lactobacillus paracasei* subsp. *paracasei* 8700:2, *Lactobacillus paracasei* subsp. *paracasei* ATCC 25302, *Lactobacillus paracasei* subsp. *paracasei* BGSJ2-8, *Lactobacillus paracasei* subsp. *paracasei* CNCM I-2877, *Lactobacillus paracasei* subsp. *paracasei* CNCM I-4270, *Lactobacillus paracasei* subsp. *paracasei* CNCM I-4648, *Lactobacillus paracasei* subsp. *paracasei* CNCM I-4649, *Lactobacillus paracasei* subsp. *paracasei* DSM 5622, *Lactobacillus paracasei* subsp. *paracasei* JCM 8130, *Lactobacillus paracasei* subsp. *paracasei* Lpp120, *Lactobacillus paracasei* subsp. *paracasei* Lpp122, *Lactobacillus paracasei* subsp. *paracasei* Lpp123, *Lactobacillus paracasei* subsp. *paracasei* Lpp125, *Lactobacillus paracasei* subsp. *paracasei* Lpp126, *Lactobacillus paracasei* subsp. *paracasei* Lpp14, *Lactobacillus paracasei* subsp. *paracasei* Lpp17, *Lactobacillus paracasei* subsp. *paracasei* Lpp189, *Lactobacillus paracasei* subsp. *paracasei* Lpp219, *Lactobacillus paracasei* subsp. *paracasei* Lpp22, *Lactobacillus paracasei* subsp. *paracasei* Lpp221, *Lactobacillus paracasei* subsp. *paracasei* Lpp223, *Lactobacillus paracasei* subsp.

paracasei Lpp225, *Lactobacillus paracasei* subsp. paracasei Lpp226, *Lactobacillus paracasei* subsp. paracasei Lpp227, *Lactobacillus paracasei* subsp. paracasei Lpp228, *Lactobacillus paracasei* subsp. paracasei Lpp229, *Lactobacillus paracasei* subsp. paracasei Lpp230, *Lactobacillus paracasei* subsp. paracasei Lpp37, *Lactobacillus paracasei* subsp. paracasei Lpp41, *Lactobacillus paracasei* subsp. paracasei Lpp43, *Lactobacillus paracasei* subsp. paracasei Lpp46, *Lactobacillus paracasei* subsp. paracasei Lpp48, *Lactobacillus paracasei* subsp. paracasei Lpp49, *Lactobacillus paracasei* subsp. paracasei Lpp7, *Lactobacillus paracasei* subsp. paracasei Lpp70, *Lactobacillus paracasei* subsp. paracasei Lpp71, *Lactobacillus paracasei* subsp. paracasei Lpp74, *Lactobacillus paracasei* subsp. tolerans, *Lactobacillus paracasei* subsp. tolerans DSM 20258, *Lactobacillus paracasei* subsp. tolerans Lpl14, *Lactobacillus paracasei* subsp. tolerans Lpl7, *Lactobacillus paracasei* TXW), *Lactobacillus brevis* (*Lactobacillus brevis* AG48, *Lactobacillus brevis* ATCC 14869 = DSM 20054, *Lactobacillus brevis* ATCC 367, *Lactobacillus brevis* BSO 464, *Lactobacillus brevis* EW, *Lactobacillus brevis* KB290, *Lactobacillus brevis* subsp. coagulans, *Lactobacillus brevis* subsp. gravesensis, *Lactobacillus brevis* subsp. gravesensis ATCC 27305), *Lactobacillus casei* (*Lactobacillus casei* 12A, *Lactobacillus casei* 21/1, *Lactobacillus casei* 32G, *Lactobacillus casei* 5b, *Lactobacillus casei* A2-362, *Lactobacillus casei* BD-II, *Lactobacillus casei* BL23, *Lactobacillus casei* CRF28, *Lactobacillus casei* DN-114001, *Lactobacillus casei* DSM 20011 = JCM 1134 = ATCC 393, *Lactobacillus casei* Lc-10, *Lactobacillus casei* LC2W, *Lactobacillus casei* LcA, *Lactobacillus casei* LcY, *Lactobacillus casei* LOCK919, *Lactobacillus casei* Lpc-37, *Lactobacillus casei* M36, *Lactobacillus casei* str. Zhang, *Lactobacillus casei* T71499, *Lactobacillus casei* UCD174, *Lactobacillus casei* UW1, *Lactobacillus casei* UW4, *Lactobacillus casei* W56), *Lactobacillus delbrueckii* (*Lactobacillus delbrueckii* subsp. bulgaricus, *Lactobacillus delbrueckii* subsp. bulgaricus 2038, *Lactobacillus delbrueckii* subsp. bulgaricus ATCC 11842 = JCM 1002, *Lactobacillus delbrueckii* subsp. bulgaricus ATCC BAA-365, *Lactobacillus delbrueckii* subsp. bulgaricus CNCM I-1519, *Lactobacillus delbrueckii* subsp. bulgaricus CNCM I-1632, *Lactobacillus delbrueckii* subsp. bulgaricus ND02, *Lactobacillus delbrueckii* subsp. bulgaricus PB2003/044-T3-4, *Lactobacillus delbrueckii* subsp. delbrueckii, *Lactobacillus delbrueckii* subsp. delbrueckii DSM 20074 = JCM 1012, *Lactobacillus delbrueckii* subsp. indicus, *Lactobacillus delbrueckii* subsp. indicus DSM 15996, *Lactobacillus delbrueckii* subsp. jakobsenii, *Lactobacillus delbrueckii* subsp. jakobsenii

ZN7a-9 = DSM 26046, *Lactobacillus delbrueckii* subsp. *lactis*, *Lactobacillus delbrueckii* subsp. *lactis* CRL581, *Lactobacillus delbrueckii* subsp. *lactis* DSM 20072, *Lactobacillus delbrueckii* subsp. *sunkii*), *Lactobacillus gasseri* (*Lactobacillus gasseri* 130918, *Lactobacillus gasseri* 2016, *Lactobacillus gasseri* 202-4, *Lactobacillus gasseri* 224-1, *Lactobacillus gasseri* ADH, *Lactobacillus gasseri* ADL-351, *Lactobacillus gasseri* ATCC 33323 = JCM 1131, *Lactobacillus gasseri* CECT 5714, *Lactobacillus gasseri* DSM 14869, *Lactobacillus gasseri* EX336960VC01, *Lactobacillus gasseri* EX336960VC02, *Lactobacillus gasseri* EX336960VC03, *Lactobacillus gasseri* EX336960VC06, *Lactobacillus gasseri* EX336960VC07, *Lactobacillus gasseri* EX336960VC10, *Lactobacillus gasseri* EX336960VC13, *Lactobacillus gasseri* EX336960VC15, *Lactobacillus gasseri* MV-22, *Lactobacillus gasseri* SJ-9E-US, *Lactobacillus gasseri* SV-16A-US), *Lactobacillus rhamnosus* (*Lactobacillus rhamnosus* 2166, *Lactobacillus rhamnosus* 51B, *Lactobacillus rhamnosus* ATCC 21052, *Lactobacillus rhamnosus* ATCC 8530, *Lactobacillus rhamnosus* BPL15, *Lactobacillus rhamnosus* CASL, *Lactobacillus rhamnosus* CRL1505, *Lactobacillus rhamnosus* DSM 14870, *Lactobacillus rhamnosus* DSM 20021 = JCM 1136 = NBRC 3425, *Lactobacillus rhamnosus* E800, *Lactobacillus rhamnosus* GG, *Lactobacillus rhamnosus* HN001, *Lactobacillus rhamnosus* K32, *Lactobacillus rhamnosus* L31, *Lactobacillus rhamnosus* L33, *Lactobacillus rhamnosus* L34, *Lactobacillus rhamnosus* L35, *Lactobacillus rhamnosus* Lc 705, *Lactobacillus rhamnosus* LMG 25859, *Lactobacillus rhamnosus* LMG 27229, *Lactobacillus rhamnosus* LMS2-1, *Lactobacillus rhamnosus* LOCK900, *Lactobacillus rhamnosus* LOCK908, *Lactobacillus rhamnosus* LR231, *Lactobacillus rhamnosus* LRHMDP2, *Lactobacillus rhamnosus* LRHMDP3, *Lactobacillus rhamnosus* MSUIS1, *Lactobacillus rhamnosus* MTCC 5462, *Lactobacillus rhamnosus* PEL5, *Lactobacillus rhamnosus* PEL6, *Lactobacillus rhamnosus* R0011), *Lactobacillus reuteri* (*Lactobacillus reuteri* 100-23, *Lactobacillus reuteri* 1063, *Lactobacillus reuteri* ATCC 53608, *Lactobacillus reuteri* CF48-3A, *Lactobacillus reuteri* DSM 20016, *Lactobacillus reuteri* F275, *Lactobacillus reuteri* I5007, *Lactobacillus reuteri* JCM 1112, *Lactobacillus reuteri* lpuph, *Lactobacillus reuteri* mlc3, *Lactobacillus reuteri* MM2-2, *Lactobacillus reuteri* MM2-3, *Lactobacillus reuteri* MM4-1A, *Lactobacillus reuteri* SD2112, *Lactobacillus reuteri* TD1), *Lactobacillus plantarum* (*Lactobacillus plantarum* 16, *Lactobacillus plantarum* 19L3, *Lactobacillus plantarum* 2025, *Lactobacillus plantarum* 2165, *Lactobacillus plantarum* 4_3, *Lactobacillus plantarum* 80,

Lactobacillus plantarum AY01, *Lactobacillus plantarum* CMPG5300, *Lactobacillus plantarum* DOMLa, *Lactobacillus plantarum* EGD-AQ4, *Lactobacillus plantarum* IPLA88, *Lactobacillus plantarum* JDM1, *Lactobacillus plantarum* LP91, *Lactobacillus plantarum* ST-III, *Lactobacillus plantarum* subsp. *argenteratensis*, *Lactobacillus plantarum* subsp. *argenteratensis* DSM 16365, *Lactobacillus plantarum* subsp. *plantarum*, *Lactobacillus plantarum* subsp. *plantarum* ATCC 14917 = JCM 1149 = CGMCC 1.2437, *Lactobacillus plantarum* subsp. *plantarum* KCA-1, *Lactobacillus plantarum* subsp. *plantarum* NC8, *Lactobacillus plantarum* subsp. *plantarum* P-8, *Lactobacillus plantarum* subsp. *plantarum* R0403, *Lactobacillus plantarum* subsp. *plantarum* ST-III, *Lactobacillus plantarum* UCMA 3037, *Lactobacillus plantarum* WCFS1, *Lactobacillus plantarum* WHE 92, *Lactobacillus plantarum* WJL, *Lactobacillus plantarum* ZJ316), *Lactobacillus johnsonii* (*Lactobacillus johnsonii* 135-1-CHN, *Lactobacillus johnsonii* 16, *Lactobacillus johnsonii* ATCC 33200, *Lactobacillus johnsonii* DPC 6026, *Lactobacillus johnsonii* FI9785, *Lactobacillus johnsonii* N6.2, *Lactobacillus johnsonii* NCC 533, and *Lactobacillus johnsonii* pf01).

[0137] Cells administered to a subject, e.g., spores and/or non-spore cells administered to a subject and/or delivered to the intestine of a subject, can express and/or deliver a fiber-synthesizing enzyme to the subject. In various embodiments, a cell administered to a subject expresses a fiber-synthesizing enzyme in and/or delivers a fiber-synthesizing enzyme to ascending colon, transverse colon, and/or descending colon. In various embodiments, a cell administered to a subject colonizes the gut. In various embodiments, a cell administered to a subject expresses a fiber-synthesizing enzyme in gut and/or delivers a fiber-synthesizing enzyme to gut before, during, and/or after the cell has colonized the gut. In some embodiments, a cell administered to a subject is non-colonizing and can pass through the entire digestive tract, e.g., while expressing and/or delivering a fiber-synthesizing enzyme of the present disclosure. In various embodiments, an engineered cell is a cell that is able to pass through a subject's gastrointestinal tract without persisting in the microbiome of the subject (orthogonal to the microbiome), optionally wherein the cell is an engineered *B. subtilis* cell.

[0138] In various embodiments, the engineered cell is of a cell type that has been designated by a governmental regulatory authority, such as the United States Food and Drug

Administration or an equivalent body of another country, as safe, or is a cell of a cell type that is otherwise known to be or regarded as safe by those of skill in the art.

[0139] In various embodiments, a cell of the present disclosure is a cell that can be fermented in large batches, e.g., using minimal media, e.g., in a 30-3,000L fermenter. In various embodiments, a cell of the present disclosure is a cell that can be purified as a biomass product including vegetative cells or spores, optionally wherein the cells of the biomass product can be readily purified and/or wherein the biomass product can be stabilized (e.g., by filtering and/or lyophilization, e.g., to facilitate shipping). For example, rapid spore formation upon nutrient starvation enables large fermentation batches to convert to spores at the end of growth. Spores can be readily stored, transported and/or purified from vegetative cells by heating the growth medium and removing the supernatant.

[0140] Various methods are known for the preparation and/or purification of spores. To provide one example, spores (e.g., *Bacillus subtilis* spores) can be prepared by a technique that includes growing cells to high density in a nutrient limited medium to stimulate sporulation, followed by a step of purifying spores. In one example, cells can be grown in Difco Sporulation Medium (0.8% Nutrient Broth, 0.1% KCl, 0.012% $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0.5 mM NaOH, 1 mM $\text{Ca}(\text{NO}_3)_2$, 0.01 mM MnCl_2 , 1 μM FeSO_4) at 37°C for 24-48 hours, with 250 rpm shaking to provide oxygen. The resulting mixture of cell debris and spores can be heated to 80°C to kill and lyse vegetative cells followed by pelleting and rinsing three times with distilled water to remove cell debris and purify the spore layer. Spores prepared in this manner can be stable in water for more than a year.

[0141] In certain embodiments, a transgene of the present disclosure is introduced into a cell that includes a sequence having significant homology to the transgene, e.g., where the transgene has at least 80% identity (e.g., at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity) with a sequence of the cell (e.g., a sequence of at least 50 nucleotides, e.g., at least 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 1,000, 2,000, 3,000, 4,000, or 5,000 nucleotides). In various such instances, homology-based interaction of the cell nucleic acid sequence and transgene can compete for and/or inhibit integration of the transgene at a different target locus. Moreover, integration of the transgene (e.g., at a target locus not homologous to the transgene) can in various instances result in genome instability and/or loss of transgene function.

For at least these reasons, in various embodiments, a homologous endogenous nucleic acid sequence of a cell is modified or deleted from the cell genome (e.g., by knockout) prior to introduction of the transgene nucleic acid sequence.

Isolated Fiber-Synthesizing Enzymes

[0142] In various embodiments, the present disclosure includes isolated fiber-synthesizing enzymes and compositions including the same. For the avoidance of doubt, while the present disclosure provides various engineered cells useful for administration to subjects such that the cells deliver to the subject a fiber-synthesizing enzyme, isolated fiber-synthesizing enzymes of the present disclosure need not be derived, isolated from such cells. To the contrary, the present disclosure includes fiber-synthesizing enzymes isolated from any source. In various embodiments, a fiber-synthesizing enzyme of the present disclosure is produced by and/or isolated from a plant or bacterium in which the fiber-synthesizing enzyme is naturally expressed. In various embodiments, a fiber-synthesizing enzyme of the present disclosure is produced by and/or isolated from a plant, bacterium, or animal (e.g., a rat, mouse, cow, pig, or horse) engineered to express a fiber-synthesizing enzyme of the present disclosure, e.g., using standard techniques of molecular biology known to those of skill in the art.

[0143] Fiber-synthesizing enzymes of the present disclosure can be isolated by any means known to those of skill in the art. Those of skill in the art will appreciate that isolation of proteins is a well-established practice with a wide variety of known techniques at the disposal of those of skill in the art for that purpose. Methods of isolating fiber-synthesizing enzymes can include, e.g., expressing fiber-synthesizing enzymes in cell culture (e.g., bacterial cell culture or mammalian cell culture, e.g., *E. coli* cell culture), tagging fiber-synthesizing enzymes with an affinity tag to facilitate isolation (e.g., by engineering and expressing a nucleic acid sequence encoding a fiber-synthesizing enzyme and an affinity tag), and/or isolating fiber-synthesizing enzyme by methods such as chromatography (e.g., high performance liquid chromatography (HPLC) or reversed-phase chromatography) or enzyme-linked immunosorbent assay (ELISA) or other immunoaffinity methods. In various embodiments, isolated fiber-synthesizing enzyme can be processed, e.g., by lyophilization or ultrafiltration.

[0144] The present disclosure includes that methods and compositions including isolated fiber-synthesizing enzymes disclosed herein can be delivered to subjects and can provide various health benefits disclosed herein.

Formulations and Applications

[0145] The present disclosure includes fiber-synthesizing enzyme formulations, e.g., where the fiber-synthesizing enzyme formulation can be or include one or more engineered cells of the present disclosure (e.g., in a spore form) and/or isolated fiber-synthesizing enzyme of the present disclosure. Enzymes (e.g., isolated fiber-synthesizing enzymes) and cells of the present disclosure can be delivered to a subject in any of a variety of formulations. The present disclosure includes, for example, formulation of isolated fiber-synthesizing enzymes and/or engineered cells of the present disclosure in any manner known for use in administration of probiotics. The present disclosure includes administration of a fiber-synthesizing enzyme formulation to a subject. In various embodiments, the present disclosure includes pharmaceutically acceptable fiber-synthesizing enzyme formulations that include a pharmaceutically acceptable carrier or excipient.

[0146] In some embodiments, a fiber-synthesizing enzyme formulation can be formulated as a dietary supplement. In some embodiments, a fiber-synthesizing enzyme formulation can be formulated in a freeze-dried powder form. In some embodiments, a fiber-synthesizing enzyme formulation can be formulated as a powder in stick packaging or sachets.

[0147] In some embodiments, a fiber-synthesizing enzyme formulation can be formulated in an oral dosage form such as a capsule or tablet (e.g., chewable tablets). In some embodiments, a fiber-synthesizing enzyme formulation can be formulated as a gummy (e.g., including ingredients such as one or more of sugar and gelatin). The present disclosure includes administration of a fiber-synthesizing enzyme formulation formulated in a delayed release oral dosage form, e.g., an enterically coated oral dosage form.

[0148] Many foods and/or ingredients can be, or be utilized as vehicles for delivery of, a fiber-synthesizing enzyme formulation, including without limitation chocolate, crackers, cereal, yogurt, nattō, kombucha, or flour. Accordingly, the present disclosure includes a fiber-synthesizing enzyme formulation formulated in or as a food product and/or food ingredient such

as chocolate, crackers, cereal, yogurt, nattō, kombucha, or flour. In various embodiments, a fiber-synthesizing enzyme formulation is formulated in or as a food product or food ingredient that is a liquid such as a juice or drink. In various embodiments, spores can be formulated in a suspension. In various embodiments, a unit dose or dosage form of a food product or food ingredient of the present disclosure has volume of 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. In various embodiments, a unit dose or dosage form of a food product or food ingredient of the present disclosure has volume that has a lower bound of, e.g., 0.05 mL, 0.5 mL, 1 mL, 10 mL, 15 mL, 20 mL, 25 mL, 50 mL, or 100 mL and an upper bound of, e.g., 50 mL, 100 mL, 150 mL, 200 mL, 250 mL, 300 mL, 350 mL, 400 mL, 450 mL, or 500 mL.

[0149] The present disclosure further includes a fiber-synthesizing enzyme formulation that is a sachet including one or more engineered cells of the present disclosure (e.g., in a spore form) and/or isolated fiber-synthesizing enzyme of the present disclosure. In various embodiments, a sachet is formulated for use of the one or more engineered cells of the present disclosure (e.g., in a spore form) and/or isolated fiber-synthesizing enzyme of the present disclosure with a food product for consumption, e.g., by combination of the food product with the fiber-synthesizing enzyme formulation, e.g., by mixing with the food product and/or distribution onto or into the food product (e.g., by sprinkling).

[0150] In various embodiments, a fiber-synthesizing enzyme formulation is included in or added to a food product (e.g., a food ingredient) prior to administration to and/or consumption of the food product by a subject. In various embodiments a food product is a food product for human consumption. In various embodiments the food product is a food for animals, such as animal feed or biscuits. In various embodiments, a fiber-synthesizing enzyme formulation is included for activity after administration and/or consumption of the food product and/or food ingredient, e.g., in the gut of a subject. monosaccharides and/or disaccharides and increase concentration or amount of fiber in the food product and/or food ingredient prior to consumption by a subject.

[0151] Various formulations can be stored at ambient conditions for extend periods of time, e.g., at least 1 week, 2 weeks, 3 weeks, 1 month, 2 months, 3 months, 6 months, or 1 year, e.g., up to 1 month, 2 months, 3 months, 6 months, or 1 year.

[0152] In various embodiments, a unit dose or dosage form of a formulation that includes an engineered cell of the present disclosure includes a number of colony-forming units (CFU) that is between $1\text{E}+04$ and $1\text{E}+14$ (e.g., about $1\text{E}+04$, $1\text{E}+05$, $1\text{E}+06$, $1\text{E}+07$, $1\text{E}+08$, $1\text{E}+09$, $1\text{E}+10$, $1\text{E}+11$, $1\text{E}+12$, $1\text{E}+13$, or $1\text{E}+14$ CFU, or within a range having a lower bound of about $1\text{E}+04$, $1\text{E}+05$, $1\text{E}+06$, $1\text{E}+07$, $1\text{E}+08$, $1\text{E}+09$, or $1\text{E}+10$ CFU and an upper bound of about $1\text{E}+08$, $1\text{E}+09$, $1\text{E}+10$, $1\text{E}+11$, $1\text{E}+12$, $1\text{E}+13$, or $1\text{E}+14$ CFU). In certain embodiments, a unit dose or dosage form includes $1\text{E}+11$ CFU of engineered cells (e.g., spores). Various examples of live cell formulation administration are known in the art. For example, in one study, $2.25\text{E}+11$ CFU was found to be a tolerable dose with no observable adverse effects. Nattō, a food made through fermentation with *bacillus subtilis* can contain as much as $1\text{E}+09$ CFU per gram and a typical consumed portion of nattō (~90-100g) can include $1\text{E}+11$ CFU. Some recommendations for consumption of *Bacillus coagulans* recommend a daily dosage of $1\text{E}+11$ - $2\text{E}+11$ CFU. Animal studies support safety and efficacy of CFU dosage levels provided herein, e.g., demonstrating that $7\text{E}+12$ and $9\text{E}+12$ as a safe dosages, e.g., for a 70kg human.

[0153] In various embodiments, a fiber-synthesizing enzyme formulation of the present disclosure is administered to a subject in need thereof. In various embodiments, administration of a fiber-synthesizing enzyme formulation of the present disclosure causes production (e.g., in a 1, 3, 6, 12 or 24 hour period) of at least 1g soluble fiber, e.g., at least 1g, 2g, 3g, 4g, 5g, 10g, 15g, 20g, 25g, 30g, 35g, or 40g. In various embodiments, administration of a fiber-synthesizing enzyme formulation of the present disclosure causes production (e.g., in a 1, 3, 6, 12 or 24 hour period) of at least 5g soluble fiber. In various embodiments, administration of a fiber-synthesizing enzyme formulation of the present disclosure causes production (e.g., in a 1, 3, 6, 12 or 24 hour period) of at least 10g soluble fiber. In various embodiments, administration of a fiber-synthesizing enzyme formulation of the present disclosure causes production (e.g., in a 1, 3, 6, 12 or 24 hour period) of an amount of soluble fiber that is in a range having a lower bound of 1g, 2g, 3g, 4g, or 5g and an upper bound of 5g, 10g, 15g, 20g, 25g, 30g, 35g, or 40g. In various embodiments, production of fiber is caused by administration of a single dose or dosage form of a fiber-synthesizing enzyme formulation of the present disclosure. Those of skill in the art will appreciate that, in various embodiments, a recommended total amount of daily fiber (e.g., by consumption and/or synthesis by isolated fiber-synthesizing enzymes and/or engineered cells of

the present disclosure) is 25 to 30g per day, including 6g to 8g per day soluble fiber (about ¼ total fiber).

[0154] In various embodiments, administration of a fiber-synthesizing enzyme formulation of the present disclosure causes incorporation into fiber (e.g., in a 1, 3, 6, 12 or 24 hour period) of at least 1g carbohydrate, e.g., at least 1g, 2g, 3g, 4g, 5g, 10g, 15g, 20g, 25g, 30g, 35g, or 40g. In various embodiments, administration of a fiber-synthesizing enzyme formulation of the present disclosure causes incorporation into fiber (e.g., in a 1, 3, 6, 12 or 24 hour period) of at least 5g carbohydrate. In various embodiments, administration of a fiber-synthesizing enzyme formulation of the present disclosure causes incorporation into fiber (e.g., in a 1, 3, 6, 12 or 24 hour period) of at least 10g carbohydrate. In various embodiments, administration of a fiber-synthesizing enzyme formulation of the present disclosure causes incorporation into fiber (e.g., in a 1, 3, 6, 12 or 24 hour period) of an amount of carbohydrate that is in a range having a lower bound of 1g, 2g, 3g, 4g, or 5g and an upper bound of 5g, 10g, 15g, 20g, 25g, 30g, 35g, or 40g. In various embodiments, production of fiber is caused by administration of a single dose or dosage form of a fiber-synthesizing enzyme formulation of the present disclosure. Those of skill in the art will appreciate that, in various embodiments, typical daily consumption of carbohydrates can include, e.g., about 45g-77g sucrose.

[0155] In various embodiments, administration of a fiber-synthesizing enzyme formulation of the present disclosure causes cleavage of (e.g., in a 1, 3, 6, 12 or 24 hour period) of at least 1g substrate carbohydrate, e.g., at least 1g, 2g, 3g, 4g, 5g, 10g, 15g, 20g, 25g, 30g, 35g, or 40g. In various embodiments, administration of a fiber-synthesizing enzyme formulation of the present disclosure causes cleavage of (e.g., in a 1, 3, 6, 12 or 24 hour period) of at least 5g substrate carbohydrate. In various embodiments, administration of a fiber-synthesizing enzyme formulation of the present disclosure causes cleavage of (e.g., in a 1, 3, 6, 12 or 24 hour period) of at least 10g substrate carbohydrate. In various embodiments, administration of a fiber-synthesizing enzyme formulation of the present disclosure causes cleavage of (e.g., in a 1, 3, 6, 12 or 24 hour period) of an amount of substrate carbohydrate that is in a range having a lower bound of 1g, 2g, 3g, 4g, or 5g and an upper bound of 5g, 10g, 15g, 20g, 25g, 30g, 35g, or 40g.

[0156] Accordingly, administration of a fiber-synthesizing enzyme formulation of the present disclosure that includes an engineered cell of the present disclosure can dramatically impact

ratio of substrate carbohydrate to a synthesized fiber type in gut. As an example, if 60g of sucrose and 4 g of soluble fiber are consumed by a subject, the weight ratio of soluble fiber:sucrose is about 6.7%. If sucrose is utilized as a substrate for fiber synthesis by a fiber-synthesizing enzyme expressed by an engineered cell, such that 5g of fiber (e.g., levan) is produced, the total simple sugar content is decreased to 55g and the total soluble fiber is increased to 9g, fiber:sucrose ratio becomes 16%. It is likely that many consumers have an intake of soluble fiber significantly lower than 4 g a day which would make the impact of administration of a fiber-synthesizing enzyme formulation of the present disclosure even more dramatic.

[0157] Various means of measuring carbohydrate (e.g., consumed and/or gut carbohydrate, e.g., monosaccharides and/or disaccharides) and/or fiber (e.g., consumed and/or gut fiber) are known in the art. However, those of skill in the art will appreciate that direct sampling of the gut (e.g., human intestinal tract) is not always possible. Accordingly, those of skill in the art will appreciate that activity and/or efficacy of a fiber-synthesizing enzyme formulation can be demonstrated by a variety of *in vivo* or *ex vivo* alternative assays that do not require direct sampling of the gut. For example, synthesis of fiber from fructose generated by cleavage of sucrose can be measured by cognate release of glucose. In certain embodiments, engineered cells are cultured in growth media together with sucrose and change in the amount of glucose in media can be quantified, e.g., by a variety of enzymatic assays and/or by direct high pressure liquid chromatography measurement. Other methods include detection of fiber-synthesizing enzyme in stool, e.g., where the fiber-synthesizing enzyme is secreted in gut by engineered cells. Samples can be derived from excrement or through invasive or other means. Another method of measuring soluble fiber production includes chromatography and/or NMR of stool. This method can be particularly effective for rarely consumed fibers such as levan. Microbiome composition itself can also provide an indication of fiber-synthesizing enzyme activity. Microbiome composition can be routinely measured by metagenomics. A decrease in number or ratio of strains associated with simple sugar dysbiosis and/or an increase in strains associated with healthy gut function from a high fiber diet demonstrate therapeutic efficacy.

[0158] In various embodiments, a fiber-synthesizing enzyme formulation is administered in a therapeutically effective amount and/or delivers a therapeutically effective amount of fiber-

synthesizing enzyme. In various embodiments, isolated fiber-synthesizing enzymes are administered in a therapeutically effective amount (e.g., a therapeutically effective number of cells) and/or deliver a therapeutically effective amount of fiber-synthesizing enzyme. In various embodiments, engineered cells are administered in a therapeutically effective amount (e.g., a therapeutically effective number of cells) and/or deliver a therapeutically effective amount of fiber-synthesizing enzyme. In various embodiments, the engineered cells are non-colonizing cells. In various embodiments the engineered cells are spore-forming cells. In various embodiments, the engineered cells germinate in the gut of a subject, e.g., the gut of a human subject, following administration to the subject. In various embodiments, engineered cells administered to a subject germinate transiently in the gut of the subject. In various embodiments, engineered cells administered to a subject secrete fiber-synthesizing enzyme in the gut of the subject. In various embodiments, engineered cells of a single administered dose of a fiber-synthesizing enzyme formulation, fiber-synthesizing enzyme produced by engineered cells of a single administered dose of a fiber-synthesizing enzyme formulation, and/or isolated fiber-synthesizing enzyme delivered by a single administered dose of a fiber-synthesizing enzyme formulation can be detected in the gut of the subject, e.g., for a period of at least 6 hours (e.g., at least 6, 12, 18, 24, 48, or 72 hours, e.g., for a period of time having a lower bound of 6, 12, 18, or 24 hours and an upper bound of 12, 24, 48, 72, 96, or 120 hours).

[0159] In various embodiments, a fiber-synthesizing enzyme formulation the present disclosure is administered to a subject, before (e.g., up to 1, 2, 3, 4, 5, 6, 12, 24, or 48 hours prior to), during, or after (e.g., up to ., up to 1, 2, 3, 4, 5, 6, 12, 24, or 48 hours after) consumption of food by the subject, optionally wherein the food includes carbohydrates (e.g., monosaccharides and/or disaccharides, e.g., sucrose). In various embodiments, a fiber-synthesizing enzyme formulation the present disclosure is administered in a dosage regimen, e.g., where a subject is administered a single daily dose, multiple daily doses (e.g., two or three daily doses), or one dose for each period of a certain number of days, e.g., 1, 2, 3, 4, 5, 6, or 7 days. To enable consumers to successfully use this technology it is highly desirable to limit the number of administration events required by a user of a supplement. The present inventors have determined one tablet a day is preferred by consumers and, among other things, is particularly preferable to taking a tablet or other form of administration with each meal. This is partly

driven by enabling privacy, wherein meals can be enjoyed under any circumstances without drawing attention to the use of the supplement. In one embodiment, efficacy in a single dose is achieved by delivering the active enzyme in *Bacillus subtilis*, which traverses the intestinal tract in a manner different to that of formulated supplements.

[0160] The present disclosure includes the recognition that, in various embodiments, it can be beneficial for the expected time passage of a composition of the present disclosure through a subject's gastrointestinal tract to be longer than the expected time for passage of typical foods through the gastrointestinal tract of the subject, which can be referred to herein as slow passage. One benefit of slow passage is that a single administration of a composition provided herein can express an enzyme of the present disclosure in the subject for a period of time that includes multiple meals and/or a period of time that is at least 3 hours (e.g., at least 3, 6, 9, 12, 18, or 24 hours, e.g., a period of time that is at least 3 to 6 hours, 3 to 9 hours, 3 to 12 hours, 3 to 18 hours, 3 to 24 hours, 6 to 9 hours, 6 to 12 hours, 6 to 18 hours, 6 to 24 hours, 12 to 18 hours, or 12 to 24 hours). In various embodiments, slow passage can be characteristic of compositions that transiently adhere to the lining of the intestine. Those of skill in the art will appreciate that, by contrast, purified enzymes delivered (e.g., using various slow release formulations) pass through the intestine at the same rate as digesta and therefore generally must be administered with or at about the same time as ingested food with respect to which therapeutic effect is desired.

[0161] The present disclosure includes the recognition that, in various embodiments, it can be beneficial for a composition of the present disclosure to pass through the stomach of a subject without significant loss of cells (e.g., survival of cells) or cell viability (e.g., capacity to germinate and/or proliferate). In various embodiments, loss of cells or cell viability is less than 50%, e.g., less than 50%, 40%, 30%, 20%, 10%, 5%, 4%, 3%, 2%, or 1%. The present disclosure includes the recognition that, in various embodiments, it can be beneficial for a composition of the present disclosure to pass through the stomach of a subject without significant loss of isolated fiber-synthesizing enzyme activity. In various embodiments, loss of fiber-synthesizing enzyme activity is less than 50%, e.g., less than 50%, 40%, 30%, 20%, 10%, 5%, 4%, 3%, 2%, or 1%.

[0162] The present disclosure includes the recognition that, in various embodiments, it can be beneficial for a composition of the present disclosure to pass through the gut within a limited timeframe (gut residency time). Agents that include protective layers can be subject to passage times that vary with the activity of the intestinal tract, which, for example, can increase or decrease depending on eating behaviors of the subject and content consumed. Agents that colonize the gut can provide long-term delivery of an expressed transgene product, but the extended residency of colonizing bacteria can allow time for mutations in the bacteria, e.g., mutations that reduce transgene expression, and proliferation or population size can be difficult to monitor, stabilize, and/or control. An inflammatory and/or immune response against a colonizing bacteria could also arise over time. Colonization can also impact the composition of the microbiome with respect to other microbes, e.g., by causing loss of a similar strain present in the microbiome prior to administration of an agent of the present disclosure. Extended residency associated with colonization can also reveal unexpected characteristics, such as the discovered carcinogenic effects of colibactin produced by the colonizing probiotic *E. coli Nissile*, a strain that had historically been developed for human consumption. In various embodiments, gut residency time of a composition of the present disclosure is less than 10 days, e.g., less than 10, 9, 8, 7, 6, 5, 4, 3, 2, or 1 days. In various embodiments, residency time is measured from administration of a dose of a fiber-synthesizing enzyme formulation of the present disclosure to a time at which the subject includes a number of cells administered with or derived from the cell composition that is no more than 5%, 4%, 3%, 2%, 1%, 0.5%, 0.1%, 0.05%, 0.01%, 0.005%, or 0.001% the total number of cells administered in the dose. In various embodiments, clearance is measured as the inability to detect DNA characteristic of administered cells in DNA samples of the subject, e.g., isolated from stool of the subject. Those of skill in the art will appreciate from the present disclosure will appreciate that the present disclosure includes that a cell composition administered in repeated doses can therefore be advantageous as compared to a colonizing cell composition, e.g., where individual doses have a limited residency time and a plurality of doses over time can be used to extend duration of treatment. In various embodiments, cells of the present disclosure are cells that lack one or more functions required for survival in and/or colonization of the gut, including but not limited to, reliance on oxygen as an electron acceptor, susceptibility to defenses of host microbes, inability to participate in the cross-feeding ecosystem

of the established microbiome, and/or lack of offensive self-preservation against host microbiome strains.

[0163] Those of skill in the art will appreciate that administration to a subject of a fiber-synthesizing enzyme formulation before, during, or after consumption of carbohydrate (e.g., monosaccharides and/or disaccharides, e.g., sucrose) can decrease concentration or amount of monosaccharides and/or disaccharides in the gut of the subject and/or decrease the rate or amount of fructose absorbed by the intestine of the subject and/or decrease the rate or amount of monosaccharides and/or disaccharides processed by microbiome cells of the subject, and also increase the production of fiber. In some embodiments, for every weight or mole units of fructose utilized as a substrate by a fiber-synthesizing enzyme an equal weight or mole units of fiber can be produced, leading to commensurate beneficial health effects from the increase in the concentration or amount of gut soluble fiber.

[0164] In various embodiments, administration of a fiber-synthesizing enzyme formulation in need thereof can treat a condition associated with or caused by consumption of carbohydrate (e.g., monosaccharides and/or disaccharides), e.g., high blood pressure, heart disease, obesity, diabetes, high blood glucose and other health problems. In various embodiments, administration of a fiber-synthesizing enzyme formulation of the present disclosure to a subject in need thereof can treat a condition associated with or caused by low soluble fiber consumption, e.g., high blood cholesterol, heart disease, obesity, diabetes, high blood glucose, and other health problems. In various embodiments, administration of a fiber-synthesizing enzyme formulation of the present disclosure to a subject in need thereof can provide health benefits associated with consumption of soluble fiber, including without limitation reduced risk of high blood cholesterol, heart disease, obesity, diabetes, high blood glucose, and other health problems. In various embodiments, administration of a fiber-synthesizing enzyme formulation of the present disclosure to a subject in need thereof to treat dysbiosis.

[0165] In various embodiments, a fiber-synthesizing enzyme formulation is administered to a subject in need thereof, wherein the subject has, has been diagnosed as having, or is at risk of developing one or more of a condition associated with or caused by consumption of carbohydrate (e.g., monosaccharides and/or disaccharides), caused by low soluble fiber

consumption, or preventable by soluble fiber consumption, e.g., where the condition is selected from high blood pressure, heart disease, obesity, diabetes, high blood glucose and other health problems.

[0166] In various embodiments, a fiber-synthesizing enzyme formulation is administered to a subject for weight management, e.g., for dieting and/or for treatment of obesity. In various embodiments, the subject is overweight or obese. In various embodiments, obesity refers to a condition of having a body mass index equal to or greater than 30. In various conditions, overweight refers to a condition of having a body mass index equal to or greater than 25. Those of skill in the art will appreciate that a fiber-synthesizing enzyme formulation of the present disclosure can also be used by an individual who is neither overweight nor obese (body mass index below 25) for purposes of weight management or for any other purposes provided herein, e.g., for gut health.

[0167] In various embodiments, a fiber-synthesizing enzyme formulation is administered to a subject having consumed in the preceding 24 hours an amount of carbohydrate (e.g., an amount of total carbohydrate or monosaccharides and/or disaccharides) that is equal to or greater than about 45g (e.g., at least about 45g, 50g, 55g, 60g, 65g, 70g, 80g, 85g, 90g, 95g, 100g, 105g, 110g, 115g, 120g, 125g, 130g, 135g, 140g, 145g, 150g, 155g, 160g, 165g, 170g, 175g, 180g, 185g, 190g, 195g, 200g, or more). In various embodiments, a fiber-synthesizing enzyme formulation is administered to a subject having consumed in the preceding 24 hours an amount of fiber that is equal to or less than about 50g (e.g., less than about 50g, 45g, 40g, 35g, 30g, 25g, 20g, 15g, 10g, 5g, or 1g). In various embodiments, a fiber-synthesizing enzyme formulation is administered to a subject having consumed in the preceding 24 hours an amount of soluble fiber that is equal to or less than about 20g (e.g., less than about 20g, 15g, 10g, 9g, 8g, 7g, 6g, 5g, 4g, 3g, 2g, or 1g).

[0168] In various embodiments, a fiber-synthesizing enzyme formulation is administered to a subject having consumed during at least about 10% (e.g., at least about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100%) of the preceding 5 or more days (e.g., 5, 7, 10, 14, 28, 42, 50, 100, 150, 200, 250, 300, 350, or 365 days) an amount of carbohydrate (e.g., an amount of total carbohydrate or monosaccharides and/or disaccharides) that is equal to or greater than about 45g (e.g., at least about 45g, 50g, 55g, 60g, 65g, 70g, 80g, 85g, 90g, 95g, 100g, 105g,

110g, 115g, 120g, 125g, 130g, 135g, 140g, 145g, 150g, 155g, 160g, 165g, 170g, 175g, 180g, 185g, 190g, 195g, 200g, or more). In various embodiments, a fiber-synthesizing enzyme formulation is administered to a subject having consumed during at least about 10% (e.g., at least about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100%) of the preceding 5 or more days (e.g., 5, 7, 10, 14, 28, 42, 50, 100, 150, 200, 250, 300, 350, or 365 days) an amount of fiber that is equal to or less than about 50g (e.g., less than about 50g, 45g, 40g, 35g, 30g, 25g, 20g, 15g, 10g, 5g, or 1g). In various embodiments, a fiber-synthesizing enzyme formulation is administered to a subject having consumed during at least about 10% (e.g., at least about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100%) of the preceding 5 or more days (e.g., 5, 7, 10, 14, 28, 42, 50, 100, 150, 200, 250, 300, 350, or 365 days) an amount of soluble fiber that is equal to or less than about 20g (e.g., less than about 20g, 15g, 10g, 9g, 8g, 7g, 6g, 5g, 4g, 3g, 2g, or 1g).

[0169] In various embodiments, a fiber-synthesizing enzyme formulation is administered to a subject that is a human. In various embodiments, a fiber-synthesizing enzyme formulation is administered to a subject that is an animal, e.g., a domestic animal (e.g., a cat, dog, or other domestic animal), a livestock animal (e.g., a horse, cow, sheep, goat, pig, or other livestock animal), or a zoo animal (e.g., a non-human primate, elephant, hippopotamus, rhinoceros, bear, or other zoo animal).

[0170] In various embodiments, treatment achieved by administration of a fiber-synthesizing enzyme formulation is advantageous as compared to direct consumption of fiber. For example, direct consumption of fiber does not prevent an overabundance of purified sugars in the colon. Additionally, a single daily dose of fiber only mediates one bolus passing through the digestive tract and may not affect other food portions containing sugar consumed throughout the day. Large doses of fiber also have physicochemical and biochemical properties that are abnormal for food, leading to changes in osmotic pressure and blooms in bacterial growth with associated negative experiences. Ingesting a large amount of fiber quickly can promote intestinal gas, abdominal bloating and cramping. No attempts have been made, to the knowledge of the present inventors, to decrease sugar after consumption prior to absorption by the intestine or microbiome. Nor, to the knowledge of the present inventors, is there a suitable technology for

sustained delivery of therapeutically effective amounts of soluble fiber over, e.g., a period of 12 or 24 hours or longer.

[0171] The present disclosure includes the discovery that solutions to various challenges disclosed herein are solved through the use of a living engineered cells as a delivery vector. In various embodiments, engineered cells temporarily adhere to and/or reside within the mucus layer lining the wall of the intestine. The proximity of the mucus layer to passing digesta allows for rapid interaction of engineered cells and/or fiber-synthesizing enzyme with passing digesta. Engineered bacterial spores and other live cells are known to transit the intestine over a period of greater than 9 hours, allowing for interaction of engineered cells and/or fiber-synthesizing enzyme produced thereby with multiple meals (e.g., two meals consumed within a period of 9 hours). Bacterial spores are extremely robust, with demonstrated ability to survive intact through the environment of the stomach. We were further able to identify, in particular embodiments, a microbial spore former that is active in the intestinal environment but does not colonize the intestine long term. This microbe, *Bacillus subtilis*, has a long history of being consumed in food as an active reagent in fermentation of soy beans, producing the food nattō, and is considered food safe. For at least these reasons, administration of an engineered cell of the present disclosure can provide longer lasting activity than ingestion of purified enzyme. Engineered cells (e.g. spores) of the present disclosure can pass safely pass through the stomach and subsequently express fiber-synthesizing enzyme in intestine. We identified that *B. subtilis* is a preferred candidate engineered cell at least in part because, when administered to a subject, it becomes caught in the mucosa (sticky ball model). *B. subtilis* is therefore exemplary of the limited set of delivery vectors that are food safe, non-colonizing (orthogonal to the microbiome), and spore forming.

REFERENCES

- Davis, L. M. G., I. Martínez, J. Walter, and R. Hutkins. 2010. “A Dose Dependent Impact of Prebiotic Galactooligosaccharides on the Intestinal Microbiota of Healthy Adults.” *International Journal of Food Microbiology* 144 (2): 285–92.
- Desai, Mahesh S., Anna M. Seekatz, Nicole M. Koropatkin, Nobuhiko Kamada, Christina A. Hickey, Mathis Wolter, Nicholas A. Pudlo, et al. 2016. “A Dietary Fiber-Deprived Gut

- Microbiota Degrades the Colonic Mucus Barrier and Enhances Pathogen Susceptibility.” *Cell* 167 (5): 1339–53.e21.
- Dong, Yutong, Li Chen, Bernard Gutin, and Haidong Zhu. 2019. “Total, Insoluble, and Soluble Dietary Fiber Intake and Insulin Resistance and Blood Pressure in Adolescents.” *European Journal of Clinical Nutrition* 73 (8): 1172–78.
- Earle, Kristen A., Gabriel Billings, Michael Sigal, Joshua S. Lichtman, Gunnar C. Hansson, Joshua E. Elias, Manuel R. Amieva, Kerwyn Casey Huang, and Justin L. Sonnenburg. 2015. “Quantitative Imaging of Gut Microbiota Spatial Organization.” *Cell Host & Microbe* 18 (4): 478–88.
- Kim, H. J., M. Camilleri, S. McKinzie, M. B. Lempke, D. D. Burton, G. M. Thomforde, and A. R. Zinsmeister. 2003. “A Randomized Controlled Trial of a Probiotic, VSL#3, on Gut Transit and Symptoms in Diarrhoea-Predominant Irritable Bowel Syndrome.” *Alimentary Pharmacology & Therapeutics* 17 (7): 895–904.
- Marlett, J. A., and T. F. Cheung. 1997. “Database and Quick Methods of Assessing Typical Dietary Fiber Intakes Using Data for 228 Commonly Consumed Foods.” *Journal of the American Dietetic Association* 97 (10): 1139–48, 1151; quiz 1149–50.
- Parnell, Winsome, Noela Wilson, Donnell Alexander, Mark Wohlers, Micalla Williden, Joel Mann, and Andrew Gray. 2008. “Exploring the Relationship between Sugars and Obesity.” *Public Health Nutrition* 11 (8): 860–66.
- Schmidt, Kristin, Philip J. Cowen, Catherine J. Harmer, George Tzortzis, Steven Errington, and Philip W. J. Burnet. 2015. “Prebiotic Intake Reduces the Waking Cortisol Response and Alters Emotional Bias in Healthy Volunteers.” *Psychopharmacology* 232 (10): 1793–1801.
- Townsend, Guy E., 2nd, Weiwei Han, Nathan D. Schwalm 3rd, Varsha Raghavan, Natasha A. Barry, Andrew L. Goodman, and Eduardo A. Groisman. 2019. “Dietary Sugar Silences a Colonization Factor in a Mammalian Gut Symbiont.” *Proceedings of the National Academy of Sciences of the United States of America* 116 (1): 233–38.
- Wei, Q., C. Wolf-Hall, and K. C. Chang. 2001. “Natto Characteristics as Affected by Steaming Time, Bacillus Strain, and Fermentation Time.” *Journal of Food Science* 66 (1): 167–73.

EXAMPLES

[0172] The present Examples demonstrate the construction of transgenes that encode a fiber-synthesizing enzyme. The present Examples further demonstrate production of engineered cells that include and express a transgene encoding fiber-synthesizing enzyme. As will be appreciated by those of skill in the art, engineered cells of the present disclosure are suitable for administration to subjects, including human subjects, and express fiber-synthesizing enzyme. Accordingly, the present Examples illustrate engineered cells representative of compositions disclosed herein and suitable for various applications provided herein, including without limitation applications that include administration to a human subject.

Example 1: Construction and Experimental Validation of Engineered Cells that express an Enzyme for Synthesis of Soluble Fiber from Carbohydrate Substrate

[0173] A transgene was constructed for expression of a fiber-synthesizing enzyme. The transgene was produced using standard techniques of molecular biology. The transgene included nucleic acid sequences encoding a fiber-synthesizing enzyme that included a secretion polypeptide. The present example utilizes SacB levansucrase enzyme. The nucleic acid encoding the fiber-synthesizing enzyme was derived from B.s. Natto *sacB* and has the following sequence, with the secretion polypeptide encoding sequence bolded and underlined:

atgaacatcaaaaagtgtgcaaaacaagcaacagtattaacctttactaccgcactgctggcaggaggcgcaactcaag
cgtttgcgaaagaaacgaacaaaagccatataaggaaacatacggcatttcccatattacacgcatgatgctgcaaatccctgaacag
 caaaaaatgaaaaatatcaagttcctgaattcgattcgccacataaaatatcttctgcaaaaggcctggacgtttgggacagctggc
 cattacaaaacgctgacggcactgtcgcaactatcacggctaccacatcgctttgcattagccggagatcctaaaaatgcggatgacacat
 cgatttacatgttctatcaaaaagtcggcgaaacttctattgacagctggaaaaacgctggccgctcttaagacagcgacaaattcgatg
 caaatgattctatcctaaaagaccaaacacaagaatggtcaggttcagccacattacatctgacggaaaaatccgtttattctactgatttct
 ccggtaaacattacggcaaacaaacactgacaactgcacaagttaacgtatcagcatcagacagctcttgaacatcaacgggtgtagaggat
 tataaatcaatctttgacggtgacggaaaaacgtatcaaatgtacagcagttcatcgatgaaggcaactacagctcaggcgacaaccatac
 gctgagagatcctcactacgtagaagataaaggccacaatacttagtatttgaagcaaacactggaactgaagatggctaccaaggcgaa
 gaatctttatttaacaaagcatactatggcaaaagcacatcattctccgtcaagaaagtcaaaaacttctgcaaaagcgataaaaaacgcacg
 gctgagttagcaaacggcgctctcggtatgattgagctaaacgatgattacactgaaaaaagtgatgaaccgctgattgcatctaacaca

gtaacagatgaaattgaacgcgcgaacgtctttaaataacggcaaatggtacctgttactgactcccgcggatcaaaaatgacgattga
 cggcattacgtctaacgatattacatgcttggtatgtttctaattcttaactggccatacaagccgctgaacaaaactggccttggttaaaa
 atggatcttgatcctaacgatgaacctttacttactcacacttcgctgtacctcaagcgaaaggaaacaatgtcgtgattacaagctatatgac
 aaacagaggattctacgcagacaaacaatcaacgtttgcgccaagcttctgctgaacatcaaaggcaagaaaacatctgttgtaagaca
 gcctccttgaacaaggacaattaacagttaacaaataa (SEQ ID NO: 20).

[0174] The protein product SacB has the following amino acid sequence, with the secretion polypeptide sequence bolded and underlined:

MNIKKFAKQATVLTFTTALLAGGATQAFAKETNQKPYKETYGISHITRHDML
 QIPEQQKNEKYQVPEFDSSTIKNISSAKGLDVWDSWPLQNADGTVANYHGYHIVFALAG
 DPKNADDTSIYMFYQKVGETSIDSWKNAGR VFKDSDKFDANDSILKDQTQEWSGSATF
 TSDGKIRLFYTD FSGKH YGKQTLTTAQVNV SASDSSLNINGVEDYKSIFDGDGKTYQNV
 QQFIDEGNYSSGDNHTLRDPHYVEDKGHKYL VFEANTGTEDGYQGEESLFNKAYYGKS
 TSFFRQESQKLLQSDKKRTAELANGALGMIELNDDYTLKKVMKPLIASNTVTDEIERAN
 VFKMNGKWYLF TDSRGSKMTIDGITSNDIYMLGYVSNLTGPYKPLNKTGLVLKMDLD
 PNDVTFTYSHFAVPQAKGNNVVITSYMTNRGFYADKQSTFAPSFLNKGKKT SVVKDS
 ILEQGQLTVNK* (SEQ ID NO: 21).

[0175] The nucleic acid sequence encoding fiber-synthesizing enzyme derived from *B.s. Natto* was operably linked with a nucleic acid sequence encoding a secretion polypeptide, such that the encoded fiber-synthesizing enzyme was a fusion polypeptide including a secretion polypeptide. The nucleic acid sequence encoding the fusion polypeptide was analyzed for secondary structure, at least in part because significant secondary structure can cause translation termination and poor expression. Base pairs that significantly contributed to secondary structure in the nucleic acid sequence were modified by silent modifications of nucleic acid sequence (e.g., non-coding and/or synonymous modifications of nucleic acid sequence) that reduced the contribution to secondary structure.

[0176] The nucleic acid sequence encoding the fiber-synthesizing enzyme was transformed into *B. subtilis* PY79 cells and operably linked to *hag* promoter by homologous recombination at the endogenous *hag* locus of *B. subtilis* PY79 genome with the addition of a modifying mutation in the promoter sequence by homologous integration as described in US 16/048,147 (published as US 2019/0076489) and PCT/US18/50957 (published as WO

2019/055707) herein incorporated by reference in their entirety and with respect to regulatory sequences and systems disclosed therein.

[0177] Fiber-synthesizing enzymatic activity was measured by the following assay. Because levansucrase cleavage of sucrose releases glucose, accumulation of glucose provides a measure of fiber-synthesizing enzyme activity. Accumulation of glucose is directly proportional to the cleavage of sucrose and formation of fructooligosaccharides. Various glucose meters are known in the art and can be standardized for measurement of glucose concentrations in bacterial media. Engineered bacteria of the present Example were cultured in media together with a defined concentration of sucrose. After culture for a defined period, bacteria were pelleted by centrifugation (21xg for 2 minutes) and remaining supernatant was contacted with a glucose meter testing strip. Levan production (mg/dL over time) is shown in Fig. 1.

[0178] Controls demonstrated no significant consumption of glucose occurred in the assay, consistent with expectations that glucose consumption due to cell growth would be <0.1% of available glucose. Further controls demonstrated that addition of sucrase to collected supernatant for complete cleavage of remaining sucrose resulted in a final yield of glucose equal to the starting amount of sucrose.

[0179] Fructose polymer formation was confirmed by addition of xylose isomerase to collected supernatant. Xylose isomerase can convert glucose to fructose, and addition of xylose isomerase equilibrates free, non-polymerized fructose with free glucose. If fructooligosaccharide fibers have formed, then the level of free, non-polymerized fructose would be significantly lower than glucose, such that xylose isomerase would catalyze conversion of glucose into fructose causing the glucose level of the supernatant to drop. If fructooligosaccharides have not formed, then the level of free, non-polymerized fructose would be comparable or higher than that of glucose, in which case xylose isomerase would not catalyze conversion of glucose to fructose and the level of glucose would rise or remain unchanged. Results demonstrated that addition of xylose isomerase catalyzed conversion of glucose into fructose causing the glucose level of the supernatant to drop, indicative of fructooligosaccharide fiber formation. Those of skill in the art will appreciate that a variety of alternative methods are readily available for quantification (e.g., of the amount and/or concentration) of glucose, fructose, fructooligosaccharide, and the like, including without limitation HPLC and HPAEC. To confirm that Levan is produced from ZB423

the strain was cultured in productive media (Per Liter in water: 2.5 g yeast extract, 1.5 g NH_4SO_4 , 7.2 g K_2HPO_4 , 0.2 g $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, and 200 g sucrose) for 72 hours. The cell culture was then pelleted to remove cells and cold ethanol (1:1.5) was added to the supernatant. A white precipitate was observed, collected by centrifugation, and further purified by twice repeating dissolution in water and precipitation in cold ethanol followed by centrifugation to collect the pellet. The pellet was dried under vacuum for 24 hours before analysis, becoming clear, hard, and brittle after drying. A strain of *Bacillus subtilis* PY79 that was not engineered to express SacB did not produce a precipitate or pellet from the supernatant and could not be analyzed. The pellet was analyzed by High Performance Anion Exchange Chromatography (HPAEC) [Figure 3.] to confirm the degree of polymerization, and by ^1H -NMR [Figure 4], ^{13}C -NMR [Figure 5], and HSQC-NMR [Figure 6] to confirm the structure of the polymer as Levan. A comparison of ^{13}C -NMR chemical shifts to those previously measured in literature shows that they closely correlate with known values for Levan [Figure 7].

[0180] Secretion of fiber-synthesizing enzyme by engineered cells was assayed in LB media. Engineered cells were cultured for 18 hours, after which period sucrose was added to media and glucose accumulation was measured periodically as described above. Additionally, a gel of the supernatant shows protein of the correct size in supernatant and not in cell fractions. Results are shown in Fig. 2.

[0181] The present Example demonstrated that transgene encoding a fiber-synthesizing enzyme derived from F1-sacB of *B. subtilis* sp. *natto* expressed a highly active fiber-synthesizing enzyme.

[0182] The present disclosure further includes the recognition that levansucrase (SacB) derived from B.s. Natto has certain exemplary qualities that confirm the utility of SacB expression in gut. The optimum pH range for levansucrase activity of sacB is 6-7 which matches that of the small intestine. SacB enzyme is naturally secreted and therefore stable outside of the engineered cell, amenable to secretion through a gram positive cell envelope.

Example 2: Engineered Cells for Increased Efficiency of Transgene Integration

[0183] The present Example demonstrated engineering of target cells to increase the efficiency of transformation of the target cells with a transgene including a nucleic acid sequence

encoding a fiber-synthesizing enzyme, where the nucleic acid sequence that encodes the fiber-synthesizing enzyme has at least 80% identity to an endogenous sequence of a target cell type. The present Example provides in particular a transgene as set forth in Example 1, encoding F1-SacB, and target cells that are *B. subtilis* PY79 cells. The transgene was flanked by homology regions (often referred to as “homology arms”) of about 800 bp designed to target the transgene for integration at the *B. subtilis* PY79 *hag* locus by homologous recombination. The transgene included a sequence of about 1400 bp having significant (greater than 80%) identity with *B. subtilis* PY79 *sacB* gene. Because of the 1400 bp region of significant identity between the transgene and the endogenous *sacB* gene, homologous recombination between the transgene and the endogenous *sacB* gene can complete with homologous recombination between the homology arms and the endogenous *hag* locus. Due to its large size, the 1400 bp region may be more likely to recombine with the endogenous genome than the desired homology arms. To reduce and/or obviate recombination between the transgene and the endogenous *SacB* gene, the corresponding sequence of the endogenous *sacB* gene was knocked out by deletion using standard molecular biology techniques.

Example 3: Engineered Cells for Constitutive Expression of Fiber-Synthesizing Enzyme

[0184] While those of skill in the art will appreciate that any of a variety of promoters are known to cause constitutive expression of an operably linked coding sequence, the present Example provides an exemplary demonstration of cells engineered for constitutive expression of a transgene-encoded fiber-synthesizing enzyme. The present Example includes a transgene that includes a flagellin gene promoter operably linked to a nucleic acid sequence encoding a fiber-synthesizing enzyme. Many bacteria have a flagellin gene homolog. Flagellin genes go 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. Various embodiments including a *hag* promoter operably linked to a nucleic acid sequence encoding a fiber-synthesizing enzyme are disclosed, and at least one representative embodiment thereof is exemplified. The disclosures of US 16/048,147 (published as US 2019/0076489) and PCT/US18/50957 (published as WO 2019/055707) are herein incorporated by reference in their entirety and with respect to regulatory

sequences and systems disclosed therein. Those of skill in the art will appreciate that the flagellin gene expression system disclosed herein is merely one of many expression systems available for expression (e.g., constitutive expression) of coding sequences in bacteria, and that the flagellin system disclosed herein was used as a matter of experimental convenience based on available reagents.

[0185] In the present Example, flagellar regulatory machinery of *B. subtilis* was adapted to accomplish constitutive and robust expression of a fiber-synthesizing enzyme, levansucrase. Various bacteria, including *B. subtilis*, regulate motility, at least in part, by a sophisticated system involving several positive and negative regulators. The constitutive expression strategy of this Example includes removing negative regulators of the highly expressed *B. subtilis* flagellin gene, called *hag*.

[0186] The gene encoding the flagellar subunit of *B. subtilis* flagellin is *hag*, and the protein it encodes is expressed 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. Transcription is mediated by a sigma factor, SigD, which is repressed by the FlgM protein. Inactivation (e.g., by deletion or mutation) of *flgM* expression 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. Translation of transcripts generated from the *hag* gene is enabled by a highly robust ribosome binding site that is bound and repressed post-transcriptionally by a protein called CsrA. However, a single point mutation in the CsrA binding site abrogates its binding and results in increased constitutive translation of the Hag protein.

[0187] Using the combination of inactivation the *flgM* gene and making the single point mutation in the CsrA-binding site can achieve extremely high levels of Hag protein expression 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, replacement of the *hag* gene with a gene encoding a fiber-synthesizing enzyme, inactivation of expression of the *flgM* gene, and introduction of a single point mutation at the CsrA binding site of the *hag* promoter can result in robust and constitutive expression of fiber-synthesizing enzyme in *B. subtilis*.

[0188] Accordingly, in certain embodiments provided herein, bacteria include a point mutation in the binding site of CsrA in combination with a mutation that inactivates *flgM* expression. By making these two mutations, constitutive expression from any SigD-based or flagellin promoter system is increased. Without wishing to be limited by any particular scientific theory, it is believed that mutation of a CsrA binding site is distinct from simple deletion of CsrA because CsrA is a pluripotent regulator in many bacterial species and its deletion could have many other potentially undesirable phenotypic effects on a cell. By making a point mutation only in the binding site, the mutation reduces or precludes CsrA repression of the *hag* promoter specifically and/or uniquely, rather than removing CsrA repression from any other targets it may have.

[0189] Certain nucleic acids disclosed herein thus include a transgene comprising a flagellin gene transcription regulatory sequence (e.g., a flagellin gene promoter, e.g., a *hag* promoter) operatively linked with a heterologous nucleic acid sequence encoding a fiber-synthesizing enzyme. 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 reduced or repressed compared to a reference with wild-type or canonical CsrA binding sites.

[0190] In certain embodiments, 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 cell 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.

[0191] As used herein, the term “Hag” (or “*hag*” or “hag”) can refer 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. *B. subtilis* Hag is encoded by the following sequence:

```
atgagaattaaccacaatattgcagcgcttaacacactgaaccgtttgtcttcaacaacagtgcgagccaaaagaacatggaga
aactttctcaggtcttcgatcaaccgtgcgggagatgacgcagcaggtcttcgatctctgaaaaatgagaggacaaatcagaggtctt
gaaatggcttctaaaaactctcaagacggaatctcttattccaaacagctgagggtgcattaactgaaactcatgcgatccttcaacgtgttc
gtgagctagtgttcaagctggaacactggaactcaggacaaagcaactgattgcaatctattcaagatgaaatttcagctttaacagatga
```

aatcgatgggtatttcaaatcgtagaattcaatggtaagaaattgctcgatggcacttacaagttgacacagctactcctgcaaatcaaaag
 aacttggtattccaaatcgagcaaatgctacacagcaaatctctgtaaatattgaggatatgggtgctgacgctcttggaaattaaagaagctg
 atggttcaattgcagctcttcattcagtaaatgatcttgacgtaacaaaattcgagataatgcagcagatactgctgatatcggttcgatgctca
 attgaaagttgtgatgaagcgatcaaccaagtttcttctcaacgtgctaagcttggtgcggtacaaaatcgctagagcacacaattaacaact
 taagcgcttctggtgaaaacttgacagctgctgagtctcgtatccgtgacgttgacatggctaaagagatgagcgaattcacaaagaacaac
 attcttctcaggcttctcaagctatgcttgcctcaagcaaccaacagccgcaaaacgtacttcaattattacgttaa (SEQ ID NO: 22).

[0192] 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 (e.g., at least 80%, e.g., 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% 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, having the nucleic acid sequence:

ggaattgacgccccaaagcatattgatattcacaggaaagaaatttacttgaccattcaggaagaaaataaccgtgcagcagcggtatccag
 cgatgtgatctccgattatcctcacaaaaaagtgaggattttttattttgtatttacaaaatcagagacaatccgatattaatgatgtag
ccgggaggaggcgaaaagactcagccagttacaaaataagggcacaaggacgtgccttaacaacatattcaggagggaacaa
aaca(ATG) (SEQ ID NO: 23) (where “ATG” represents the start codon of *hag*). The sequence in bold beginning with TTAA (underlined) through the start codon ATG is sufficient to promote expression of an operably linked coding sequence.

[0193] In particular, the *hag* promoter includes 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: 24) (high-affinity binding site 1, or “BS1”) (underlined) and “attcaggaggaa” (SEQ ID NO: 25) (low-affinity binding site 2, or “BS2”) (underlined). The *hag* promoter also includes a Shine-Dalgarno sequence: agggagga (SEQ ID NO: 26).

[0194] 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

nucleic acid 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 encoded in *B. subtilis* by the sequence:

atgctagttttatcgcgaaataaacgaagcgattcaaataggtgctgatattgaagtaaaagtattgcggtgaaggggatcaagtgaag
cttggaattgacgccccaaagcatattgatattcacaggaaagaaatttacttgaccattcaggaagaaaataaccgtgcagcagcggtatcc
agcgatgtgatctccgattatcctcacaaaaaagtga (SEQ ID NO: 27).

[0195] In certain embodiments, transgenes of the present disclosure include genetic modifications in a flagellin gene promoter that, upon transcription from the promoter into a transcript, such as mRNA, reduce and/or repress CsrA inhibition of mRNA translation. The present disclosure contemplates several genetic modifications to a flagellin gene promoter, and, in particular, to a *hag* promoter, to reduce and/or repress CsrA inhibition of mRNA translation. 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. A genetic modification can include one or more point mutations to CsrA BS1 (binding site 1P. 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 include one or more mutations in the 12-base-pair BS1 binding site or in the surrounding bases on either side of the nucleotides that form the stem of the stem-loop secondary structure of BS1. This includes, for example, modification of one or a plurality of nucleotides (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, or at least 12 nucleotides) of the BS1 binding site, gcacaaggacgt (SEQ ID NO: 24). Alternatively, a genetic modification 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 **taagggcacaaaggacgtgcctta** (SEQ ID NO: 28) 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: 29). A genetic

modification can also or alternatively include one or more point mutations to CsrA BS2 (binding site 2). This includes, for example, modification of one or a plurality of nucleotides (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, or at least 13 nucleotides) of 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. Alternatively, a genetic modification can disrupt 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: 30). In certain embodiments, the modification does not include an alteration of nucleotides in the Shine-Dalgarno sequence agggagga (SEQ ID NO: 26).

[0196] 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.

[0197] 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 (e.g. inactivation) of expression and/or activity of Sigma factor repressors such as FlgM.

[0198] As used herein, the term “FlgM” can refer 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*, FliA 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:

atgaaaatcaatcaatttgaacacaatccgtaatccatatcaaaaaattatgataagcaagcgggtgcaaaaaactgtgcacaacctcaag
ataaaattgaaatttcacacaggctaagaaatgcaacatgcacgcagtcactggttcacgacaggaaaaattgcgcagcttaa
gcgcaaattgaaaacgggtcatacaagtagacgcaaatcatattgcgaaaaatgattaattttataaaaagcaataa (SEQ ID
NO: 31).

[0199] As used herein, the term “SigD” can refer 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 FlhA in *E. coli*, or the broader denotation of sigma-28 in several species. SigD is defined in *B. subtilis* by the sequence:

```
atgcaatccttgaattatgaagatcaggtgcttggacgcgctggaaagagtggaaagatcctaaagccggtgacgacttaatgc
gccgttacatgccgcttgcacatatcatgtaggcagaattctgtcggactgccgaaatcagtgcataaagacgatcttatgagccttggtatg
cttggtttatatgatgcccttgaaaaattgacccagccgggacttaaaattgatacctacgcctcgtttagaattcgcgcgcaatcataga
cgggcttcgtaaagaagattggctgccagaaacctgcgcgaaaaacaaaaagggtgaagcagcaattgaaaagcttgacagcggta
tcttcggaatgtatgcccgcggaattgcagaggaactcggaatgacgggtacaggatgtcgtgtcaacaatgaatgaagggtttttgcaa
tctgtgtcaattgatgaaaagctccatgatcaagatgacgggaaaacattcaagtcagatgatcagagatgacaaaaatgttcgcctgaaga
aaagattatgaaggatgaactgattgcacagcttcgggaaaaattcacgaactctctgaaaaagaacagctggtgtcagttgttctacaaa
gaggagtgacactgacagaaatcggaagaagtattaaatcttctacgtccgcataatcagatccattcaaaggcattattaaattaaagaa
tctgtggaaaaagtatacaataa (SEQ ID NO: 32).
```

[0200] FlgM can bind 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.

[0201] In various embodiments, a cell includes a genetic modification that reduces the ability FlgM to inhibit SigD activity. For example, one such genetic modification can be 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.

[0202] 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.

[0203] A heterologous nucleic acid sequence such as a transgene encoding a fiber-synthesizing enzyme can be operably linked with a flagellin promoter by any method known in the art. For example, a heterologous nucleic acid sequence can be integrated into the bacterial chromosome. Alternatively, a heterologous nucleic acid sequence can be attached to a flagellin promoter in a plasmid that is introduced into the microorganism. A heterologous nucleic acid 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 an attached nucleic acid sequence at a target locus. Various transposon systems are known in the art.

[0204] The present Example provides engineered cells that express a fiber-synthesizing enzyme, where the cells are PY79 strain cells that include an inactivation deletion of *flgM* (Δ flgM), a deletion of endogenous *sacB* (Δ sacB), and a modified *hag* promoter (“Pso3”) engineered to reduce binding of CsrA to transcripts (Δ hag, the engineering including a mutation referred to alternatively as “sow3” or “so3”) operably linked to a nucleic acid sequence encoding a fiber-synthesizing enzyme, where the nucleic acid sequence encoding the fiber-synthesizing enzyme has the following structure: Pso3::nucleic acid sequence encoding secretion polypeptide::nucleic acid sequence encoding sacB. These engineered cells were assayed for enzymatic activity *in vitro*. A unit dose of 1E+09 engineered spore form cells was found to produce 10.4 mg of glucose through cleavage of sucrose for incorporation into fiber per hour. Scaled to an exemplary therapeutic unit dose of 1E11 CFU of spore form cells, a single unit dose of 1E11 engineered cells yields at least 1.04g glucose from cleavage of sucrose for incorporation into fiber per hour, or at least about 25 g of glucose from cleavage of sucrose for incorporation into fiber per day. These engineered cells produce a stoichiometric amount of fiber in the reaction, and a single unit dose of 1E11 engineered cells therefore synthesizes 25g of fiber per day.

Example 4: Engineered Cells for Conversion of Lactose to Human Milk Oligosaccharide Prebiotic Fibers

[0205] Genetically modified *Bacillus subtilis* strains were engineered to express transgenes encoding Human Milk Oligosaccharide-synthesizing enzymes. Transgenes and bacterial strains were produced using standard techniques of molecular biology. Each transgene included a nucleic acid sequence encoding a lactose-degrading/oligosaccharide-synthesizing enzymes. In particular, four engineered strains, each including a distinct transgene, were produced. The strains respectively included transgenes for expression of LacZ (*E. coli*), YesZ (*Bacillus subtilis*), GanA (*Bacillus subtilis*), and LacZ (*Bacillus coagulans*) beta-galactosidase enzymes in accordance with the following sequence information:

[0206] Nucleic acid sequence encoding LacZ (*E. coli*):

atgacatgattacggattcactggccgtcgtttacaacgtcgtgactgggaaaaccctggcgttacccaactaatcgcttgca
gcacatcccccttccgagctggcgtaatagcgaaggcccgaccgatcgccctccaacagttgcgcagcctgaatggcgaatgg
cgctttgcctggttccggcaccagaagcgggtgccgaaagctggctggagtgcgatcttctgaggccgatactgtcgtcgtccctcaaa
ctggcagatgcacggttacgatgcgccatctacaccaacgtgacctatccattacggtaacccgcttgggtccacggagaatccga
cgggtgttactcgtcacttaattgtgatgaaagctggctacaggaaggccagacgcgaattattttgatggcgtaactcggcggttcat
ctgtggtgcaacggcgctgggtcggttacggccaggacagtcgttgcgtctgaatttgacctgagcgcattttacgcgccggagaaaa
ccgctcgcggtgatggtgctgcgtggagtacggcagttatctggaagatcaggatatgtggcgatgagcggcattttccgtgacgtc
cgttgctgcataaaccgactacacaaatcagcgatttccatgttgcactcgtttaatgatgatttcagccgcgtgtactggaggctgaagtt
cagatgtgcggcgagttgcgtgactacctacgggtaacagtttctttatggcagggtgaaacgcaggtcgccagcggcaccgcgcttccg
gcggtgaaattatcgtatgagcgtggtgttatgccgatcgctcacactacgtctgaacgtcgaaaacccgaaactgtggagcgcggaat
cccgaatctctatcgtgcggtggtgaactgcacaccgacggcagcgctgattgaagcagaagcctgcgatgtcggttccgcgaggtg
cggattgaaaatggtcgtcgtcgtgaacggcaagccgttgcgtgattcgaggcgtaaccgtcacgagcatcctctcgtatggtcaggtc
atggatgagcagacgatggtgcaggatatcctgctgatgaagcagaacaactttaacgccgtgcgtgttcgattatccgaaccatccgct
gtggtacacgctgtgcgaccgtacggcctgtatgtggtgatgaagccaatattgaaaccacggcatggtgccaatgaatcgtctgacc
gatgatccgcgtggctaccggcgatgagcgaacgcgtaacggaatggtgcagcgcgatcgaatcaccgagtgatgatcctgtgctg
ctggggaatgaatcaggccacggcgtaatcacgacgcgtgtatcgtggatcaaatctgtcgtatcctccgcccgggtgcagtatgaag
gcggcgagcggacaccacggccaccgatattttgcccgatgtacgcgcgtggatgaagaccagccctcccggtgtgcccgaat
gggtccatcaaaaaatggcttccgtacctggagagacgcgcccgtgatcctttgcaatacggcacgcgatgggtaacagtcttggcggt
ttcgctaatactggcaggcgttctcgtcagatccccgttacaggcggtctcgtggactgggtgatcagtcgtgattaaatgatg

aaaacggcaaccggtggtcggttacggcggtgattttggcgatacgcgaacgatcgccagttctgtatgaacggctggtctttgccgac
cgcacgccgatccagcgctgacggaagcaaacaccagcagcagttttccagttccgtttatccgggcaaaccatcgaagtgaccagc
gaatacctgttccgtcatagcgataacgagctcctgcactggatgggtggcgctggatggtaagccgctggcaagcggtgaagtgcctctgg
atgtcgctccacaaggtaaacagttgattgaactgcctgaactaccgcagccggagagcgccgggcaactctggctcacagtacgcgtagt
gcaaccgaacgcgaccgatggtcagaagccgggcacatcagcgctggcagcagtggtgctggtgggaaaacctcagtgtagcgtc
cccgcgcgtcccacgccatcccgcattctgaccaccagcgaaatggattttgcatcgagctgggtaataagcgttggaatttaaccgcca
gtcaggctttctttcacagatgtggattggcgataaaaaacaactgctgacgcccgtgcgcgatcagttcacccgtgcaccgctggataacg
acattggcgtaagtgaagcgacccgcattgaccctaaccgctgggtcgaacgctggaaggcgggcgccattaccaggccgaagcagc
gttgttcagtgacggcgagatacactgctgatgcggtgctgattacgaccgctcacgcgtggcagcatcaggggaaaacctatttatcag
ccggaacacctaccgattgatgtagtggtcaaatggcgattaccgttgatgtgaagtggcgagcgataaccgcatccggcgcggtgatt
ggcctgaactgccagctggcgaggttagcagagcggttaactggctcggttagggccgcaagaaaactatcccaccgccttactgc
cgctggttttaccgctgggactgctccattgtcagacatgtataccccgtactgttcccagcgaaaacgggtctgcgctgcccggacgcgcg
aattgaattatggccacaccagtggtgcggcgacttcagttcaacatcagccgctacagtaacagcaactgatggaaccagccatcg
ccatctgctgcacggaagaaggcacatggctgaatatcgacggtttccatatggggattggtggcgacgactcctggagcccgctagta
tcggcggaattccagctgagcgccggtcgctaccattaccagttggtctggtgtcaaaaataa. (SEQ ID NO: 71).

[0207] LacZ (*E. coli*) amino acid sequence (encoded by SEQ ID NO: 71):

MTMITDSLAVVLQRRDWENPGVTQLNRLAAHPPFASWRNSEEARTDRPSQQLR
SLNGEWRFAWFAPEAVPESWLECDLPEADTVVPSNWQMHGYDAPITYTNVTYPITVN
PPFVPTENPTGCYSLTFNVDESWLQEGQTRIIFDGVNSAFHLWCNGRWVGYGQDSRLPS
EFDLSAFLRAGENRLAVMVLRWSDGSYLEDQDMWRMSGIFRDVSLHKKPTTQISDFHV
ATRFNDDFSRAVLEAEVQMCGELRDYLRVTVSLWQGETQVASGTAPFGGEIIDERGGY
ADRVTLRLNVENPKLWSAEIPNL YRAVVELHTADGTLIEAEACDVGFREVRIENGLLLL
NGKPLLIRGVNRHEHHPLHGQVMDEQTMVQDILLMKQNNFNNAVRC SHYPNHPLWYTL
CDRYGLYVVDEANIETHGMVPMNRLTDDPRWLPAMSERVTRMVQRDRNHPSVIIWSL
GNESGHGANHDALYRWIKSVDPSRPVQYEGGGADTTATDIICPMYARVDEDQPFPAVP
KWSIKKWL SLPGETRPLILCEYAHAMGNSLGGFAKYWQAFRQYPRLQGGFVWDWVDQ
SLIKYDENGNPWSAYGGDFGDTPNDRQFCMNGLVFADRTPHPALTEAKHQQQFFQFRL
SGQTIEVTSEYLF RHSDNELLHWMVALDGKPLASGEVPLDVAPQGKQLIELPELPQPESA
GQLWLTVRVVQPNATAWSEAGHISAWQQWRLAENLSVTLPAAASHAIPHLTTSEMDFCI
ELGNKRWQFNRSQFSLQMWIGDKKQLLTPLRDQFTRAPLDNDIGVSEATRDPNAWV

ERWKAAGHYQAEAALLQCTADTLADAVLITTAHAWQHKGKTLFISRKTYRIDGSGQM
 AITVDVEVASDTPHPARIGLNCQLAQVAERNWLGLGPQENYPDRLTAACFDRWDLPL
 SDMYTPYVFPSENGLRCGTRELNYGPHQWRGDFQFNISRYSSQQQLMETSHRLLHAAEE
 GTWLNIDGFHMGIGGDDSWSPSVSAEFQLSAGRYHYQLVWCQK* (SEQ ID NO: #72),

[0208] Nucleic acid sequence encoding YesZ (*Bacillus subtilis*):

atgagaaaactgtatcatggcgcttgctattatccggaattatgggatgaagagacgattcagcaggacattgacatcatgcgtga
 agttggcgtaaatgtgtgctggatcggcgaatttgcctggcagtcagtggaacctgaagaaggaaaaattgacgtcggtttttcaagaaatc
 atgccccggctgtatgatacgggatcgaacgattatgtgcacgccagcgcctaccccgcgatttggttctcacatggccggcccgaaac
 gcatgcatgccaatgaaaaagagagatcatggggcatggctcccgctcagcatgcctgtacgaacaacccgtatttccgaaaaaagccg
 ccatcatcaccacagccatcgccaaggagcttgccggctccggggctgacggatggcagctagacaatgagtttaaatgcatgttgc
 agaatgcatgtgtgagacatgcttgcgcctatggcatgactggctcaaaatcgctacgggtaattgagcgttgatgaagcttggggaa
 ccgatgtgtggagcgagacatcatgacgtttgagcaagtcggcgacccgggacggccccgttctgcatcatgcctctctacgcactatg
 tatcagctgtttcgtatggagatgatcgttctgttgcggatgaacaggccaaaatcatccgctgctattcagatgcgcccgatcacgcataacg
 gatcagtcatttgcagctggacaatgagcgaatgttcagaatctcgattttgcctctatgacacgtacgcttcgaggaaaacgcctctgc
 cttttattgaactgtgatttatggagaaatctgaaacaagggcgcccgttttgattttgaaacgagtcgctgatccgcctcgcttgaaa
 gctccgcttaccgcacgcagacgggtatttgcaggccgaagccgtatcgtctacgccttagggagccaggggtttgctactggctatgg
 cgacagcagcgttcaggcagcgagatttccacgggttcgttctcagtcgctggggcgaaccaccatcggtatcaaatgtgctggcgg
 ttgagcgggcaagaaggaaatcagcctattattctatcgactgaacccgttcaagccgagggcgcgatgactactctgacagagcaaa
 agcatttataaaactgagcctaccggggactccggcatcgttcgttgacgcattttatgaacgtatttctaacacggggattaccgtg
 acctattccggaaggcgctccactggacggctatcgttctgtttacgccatttgcctgatttgccttctgaattatcaaaaaagcttcggc
 attcgtgaagcggggcgcatctggatcaccgggcccgtgacaggaggacgcacatgcgagcataccattcataccgattgcggacttgg
 cgaacttgagaaaacgtcagggatcaaaacactttttaccttccgatgaatgagaacgtgaatacaggaaaagcgtttggcatcacggcgc
 cgctcggactgtggagcgcggtgtttgacacagagagcggaaacacccctggcacggtgaagcaggaccggggcgggccatgctttt
 ctgacggaaacgaattacggcgaggggaaaattgtcatgctgggctcgttccatccgggaaagaaggggatgcgatgctggaagcgct
 cgtcaggcattatgcggaggaagctgttattccagccggtcggatgtgacacccggcacgatcgttccccgcgtataggcgaaaacgg
 ccttgtgtgatcgttgaatatggatggaaaaggcgggagcgtgacattgccggaatcgggaacggatttgtgacgcaccgcttgaa
 aaggcggggagactggcggtcggaccgatgaataccgtgtgattcaatttgacaatcacagctga (SEQ ID NO: 73)

[0209] YesZ (*Bacillus subtilis*) amino acid sequence (encoded by SEQ ID NO: 73):

MRKLYHGACYYPELWDEETIQDIDIMREVGVNVRIGEFAWSVMEPEEGKIDV
 GFFKEIARLYDSGIETIMCTPTPTPIWFSHGRPERMHANEKREIMGHGSRQHACTNNPY

FRKKAIIITTAIAKELGRLPGLIGWQLDNEFKCHVAECMCETCLRLWHDWLKNRYGVIE
 RLNEAWGTDVWSETYQTFEQVPQGPAPFLHHASLRTMYQLFSMEMIASFADEQAKIIR
 CYSDAPITHNGSVMFSDNERMFQNLDFASYDTYASQENASAFLLNCDLWRNLKQGRP
 FWILETSPSYAASLESSAYPHADGYLQAEAVSSYALGSQGFCYWLWRQQRSGSEISHGS
 VLSAWGEPTIGYQNVLAVERARKEIEPIILSTEPVQAEAAMTYSDRAKAFIKTEPHRGLR
 HRSLVTHFYERILNTGIHRDLIPEGAPLDGYRLLFTPFVPYLSSEFIKKASAF AEAGGIWIT
 GPLTGGRTCEHTIHTDCGLGELEKTSGIKTLFTFPMNENVNTGKAFGITAPLGLWSAVFD
 TESGNTLGTVEAGPGAGHAFLTERNYGEGKIVMLGSLPSGKEGDAMLEALVRHYAEEA
 VISSRSDVTPGTIVAPRIGENGLVWIVVNMDGKGGSVTLPESGTDLLTHRLEKAGRLAV
 GPHEYRVIQFDNHS* (SEQ ID NO: 74

[0210] Nucleic acid sequence encoding GanA (*Bacillus subtilis*):

atgtcaaaagcttgaacacgcacgtaacaaaagcaaaatttatgctccatgggggagactacaaccccgatcagtggtggt
 cggcccgatatttagctgacgatatcaactgatgaagctttctacatgaatacgtttctgctggcattttgcatggagcgcacttgagccg
 gaggagggcgatatcaattgaatggctggatgatattttgagcggattcacagtataggcggccgggtcatattagcaacgcccagcgg
 agcccgctccggcctggctgtcgaaacctatccggaagtttgcgcgtcaatgcctcccgcgtcaaacagctgcacggcggaaggcaca
 ccactgcctcacatcaaaagtctaccgagaaaaaacacggcacatcaacgcttattagcagaacgatacggacatcacccggcgctgta
 atgtggcacatttcaaacgaatacgggggagattgccactgtgattatgccagcatgctttccgggagtggtgaaatcgaaatgacaac
 agcctaagacattgaacctgcgtggtggacccttttggagccatacgttcaatgactggtcacaattgaaagcccttcgcccagtcggt
 gaaatggcttgcacgctgaatttagattggcggcgttcgtcaccgatcaaacgatttcgtttatgaaatgaaatcattccgctgaaaga
 attgacgctgatatccctatcacaaatgttgctgacacacgggatttgatcccgtatcaggcgctgactacagcaatttgcaaag
 catgtcgatgccatcagctgggacgcttatcctgtctggcacaatgactgggaaagcacagctgattggcgatgaaggtcggctttatcaat
 gatctgtaccgaagcttgaagcagcagcccttcttataatggagtgtacgccaagcgcggtcaattggcataacgtcaacaaggcaagc
 gcccgggcatgaatctgctgtcatccatgcaaatgattgccacggctcggacagcgttctctatttccaataccgcaaatcacgggggtcat
 cagaaaaattacaggagcgggttggtgatcatgacaatagcccgaagaaccgcgtcttcaagaagtggccaaggtaggcgagacattgg
 aacggctgtccgaagttgtcggaacgaagaggccggtcacaacggcgattttatgactgggaaatcattgggcgctcgaggatgctca
 ggggtttcggaaggcgacaaaacgttatccgcaaacgcttcagcagcattaccgcacattctgggaacacgatccctgtcgacgtcatca
 cgaaagaacaagactttcaccatataaactgctgatcgtccgatgctgtatttaacagcgaggacaccgtttcccggttaaaagcgtttacg
 gctgacggcgccaccttagtcacgtatatcagcggggttgatgagcatgacttaacatacacaggcggtatggcatccggatcttca
 agctatatttgagttgagcctcttgaacggacaccctgtatccgaaggatcgaaacgctgtcagctaccgcagccaaatatatgaaatgaa
 ggattatgaacgctgattgatgtaagacagcttcagtgaagcgggtgatcaagaagattttatgcgcgcacgcccagcgggtcacaagcc

atgagatcagcagggcaaggcgattttatcggcgcgcgttggaggatcaatttcagcgtgatttctatagggctgatcacagacctgtc
tctctctccagttttccgggttcggcacggaaaaggcgtctccgtacaagcaggcaggatcaggacaatgattatattttgtcatgaatttca
cggaagaaaaacagctggtcacgttggatcagagtggaaggacataatgacaggagacatattgtcaggcgacctgacgatggaaaagt
atgaagtgagaattgtcgtaaacaacacattag (SEQ ID NO: 75)

[0211] GanA (*Bacillus subtilis*) amino acid sequence (encoded by SEQ ID NO: 75):

MSKLEKTHVTKAKFMLHGGDYNPDQWLDPRDILADDIKLMKLSHTNTFSVGIFA
WSALEPEEGVYQFEWLDDIFERIHSIGGRVILATPSGARPAWLSQTYPEVLRVNASRVKQ
LHGGRHNHCLTSKVYREKTRHINRLLAERYGHHPALLMWHISNEYGGDCHCDLCQHAF
REWLKSKYDNLKTLNHAWWTPFWSHTFNDWSQIESPSPIGENGLHGLNLDWRRFVTD
QTISFYENEIPLKELTPDIPITNFMADTPDLIPYQGLDYSKFAKHVDAISWDAYPVWHN
DWESTADLAMKVGFINLDRSLKQQPFLLMECTPSAVNWHNVNKAKRPGMNLLSSMQ
MIAHGSDSVLYFYQYRKSRSSEKLHGA VVDHDNSPKNRVFQEVAKVGETLERLSEVVG
TKRPAQTAILYDWENHWALEDAQGF AKATKRYPQTLQQHYRTFWEHDIPVDVITKEQD
FSPYKLLIVPMLYLISEDTVSRLLAFTADGGTLVMTYISGVVNEHDLTYTGGWHPDLQAI
FGVEPLETDTLYPKDRNAVSYRSQIYEMKDYATVIDVKTASVEAVYQEDFYARTPAVTS
HEYQQGKAYFIGARLEDQFQRDFY EGLITDLSPVFPVRHGKGVSVQARQDQDNDYIF
VMNFTEEKQLVTFDQSVKDIMTGDILSGDLTMEKYEVRIVVNTH* (SEQ ID NO: 76)

[0212] Nucleic acid sequence encoding LacZ (*Bacillus coagulans*):

atgctcaaaaagcacgaaaagttctactatggcgggtgattataatcctgaacaatgggacgaaaagcgtctggaaagaggatagc
gcttgatgaagaagcaggtgttaactatgtaccataaacattttctctgggcacgtctccaacctgatgaagaacatatgattttctacgct
tgataaaataatggatagctggctgaaaacggaattggtgctgacctggctaccgccacggctgctccgccggcctggctgtcacgtaagt
atcctgattctttccgggtcgacaagatggctcccggttctgcccggatctcgccaacactactgtccgaactctaaagactatgtagact
cgcagctaaattggtgagaaagatcgtgagcgtataaaagtcacccagcattagttatgtggcatgtaaacgaatacggctgccaca
tatctgaatgctactgcgataattgtaaaaagggtttcaaacgtggctcaaggagaaatatggaacgatcgagaactggaataagagctgga
gtaccgatttctggtcacagcgtactatgagtggaagaaattgcctccctggaaaaacacctacctttgcgaatccaatgcagcagctcg
attataaggcctttatggatgatagcctgttagcactgtataaatggagcgtgacatactgaaaacttataccagacgtcccagtcagac
gaatttaattggggcttcataaacagtgagcggctttcactgggctaaggagatggatttggtfacctgggacgcgtatcctgatcctttcag
gacatcccgtacgtcagttcatggcgacgatctgacacgcagcttgaaagaaacaccttttctctcatggaacaggccgccccggcgccg
taaattggcgcgacagaacgtgttaaggcgccaggggttatgctttatggtcatacgaagcagcgcgcatggtgctgacggtataat
gtttttcaatggcgggcaagtaaggaggcgcggaataattcatagcgggatggtaccgattcaggagatgaggagtctcggaattttc

gggagggtcgtagcttaggaaatgaacttaagaatttgaaaaagtaacgggaagtgcgtacgcgtccgacgtagcaatagttttgattgg
 aaaaactgggtggcggttggaactggacagtaagccgagctctctgggtcacttatataaaacaactcctccggttctatcgggtttgacacgc
 agaacataggtgtcgactttatccatccagatgaagctatggatcgctacaaggtggttttcgctccggcgagctaccgggtgacaaagacg
 ttgcagataaggtaaggcatagtagagaacggaggatatttcgcgacaaacttctcagcgggtagctgatgagaatgaacgtgtgta
 ccttgagggttaccaggcgcttaccgtgacattttgggtatatatgtggaagagtttgcggcgatgaaaaaggagcggtacatcagatccg
 gactggatagcgagatgctgcgatacgctgtgggaagagaaaattcattgaaaggcgccgaggcactcgctggtttaaggatggttat
 ctggccggctcaccggcggtgaccgcacatcactgtggcaaggcaagcatactatatggcacacagccagatgagcaatactatcct
 cactgctgaaggaaattctcaaggagggtgacgttcgcccggccctcgatgctccgctggagtagaagtcgcggttcgcaaaacggtc
 atgaaaaatttcttcttactgaaccatacagatcaggtgcaattcgtatgcccggcggtacttatccagaactgattacggtcgcaccgaa
 gccgaaccgtgagactctcaccacgcgacgtgaaaatcctcaggtcatagagaaataa (SEQ ID NO: 77)

[0213] LacZ (*Bacillus coagulans*) amino acid sequence (encoded by SEQ ID NO: 77):

MLKKHEKFYYGGDYNPEQWDESVWKEDMRLMKKAGVNYVSINIFSWARLQPD
 EETYDFSTLDKIMDMAENGIGADLATATAAPPAWLSRKYPDSLVPDKDGSRFLPGSRQ
 HYCPNSKDYARLA AKLV RKIAERYKSH PALVMWHVNNEYGCHISECYCDNCKKGFQT
 WLKEYGTIENLNKSWSTDFWSQRYYEWEICLPKGTPTFANPMQQLDYKAFMDDSL
 ALYKMERDILKTYTPDVPVMTNLMGLHKPVDGFHWAKEMDLVTWDAYPDPFEDIPYA
 QFMAHDLTRSLKKQPFLMEQAAGAVNWRAQNAV KAPGVMRLWSYEAAAHGADGI
 MFFQWRASQGGAEKFHSGMVPHSGDEESRNFREV VQLGNELKNLEKVTGSAYASDVAI
 VFDWKNWWALELDSKPSSLVTYIKQLLPFYRVLHTQNIGVDFIHPDEAMDRYKVVFA
 ASYRVTKTFADKVKAYVENGGYFATNFFSGIADENERVYLGGYPGAYRDILGIYVEEFA
 PMKKGAVHQIRTGYGDAAIRVWEEKIHLKGAEALAWFKDGYLAGSPAVTAHHCGK GK
 AYYIGTQPDEQYLSSLLKEILKEADVRLDAPRGVEVAVRKN GHEKFLFLNHTDQVQ
 FVDAGGTYP ELIYGRTEAETVRLSPRDVKILQVIEK* (SEQ ID NO: 78).

[0214] In the case of expressing yesZ or ganA constructs in *Bacillus subtilis*, endogenous sequences encoding corresponding sequences were deleted using standard molecular biology techniques, to prevent undesirable recombination. The present Example further includes the insight that, while not required, fiber synthesis can be increased by genetic modification of cells to increase transport of Lactose across the cell membrane. Accordingly, strains were additionally, optionally engineered to express one or more lactose transporters. In the case of cells engineered to express LacZ (*E. coli*), cells were further modified to express a heterologous

nucleic acid sequence encoding polar linked lacY transporter, the heterologous nucleic acid sequence including the following sequence:

atgtactatttaaaaaacacaaacttttgatgttcggtttattcttttctttactttttatcatgggagcctacttcccgttttcccatt
 tggctacatgacatcaaccatatcagcaaaagtgtacgggtatttttgcgcgtatttctgttctcgctatttccaaccgctgttggctctg
 ctttctgacaaactcgggctgcgcaaacctgtgtggattattaccggcatgttagtgatgttgcgcgttcttttttcttctgggccact
 gttacaatacaacatttttagtaggatcgattgttgggtggtatttatctaggcttttgtttaacgcgggtgcgccagcagtagaggcatttattgag
 aaagtcagccgtcgcagtaatttcgaatttggtcgcgcgcggatgttggctgtgtggctgggcgctgtgtgcctcgattgtcggcatcatgtt
 caccatcaataatcagtttcttggctgggctctggctgtgcacatcctcgcggtttactcttttcgcaaaacggatgcgcctcttctg
 ccacgggtgccaatgcggtagggtgccaacctcggcatttagccttaagctggcactggaactgttcagacagccaaaactgtggttttgc
 actgtatgttattggcgtttctgcacctacgatgttttgaccaacagtttgaatttcttactcgttcttctgctaccgggtgaacagggtacgcg
 ggtatttggctacgtaacgacaatgggcgaattacttaacgcctcgattatgttcttgcgccactgatcattaatcgcatcggtgggaaaaacg
 ccctgctgctggctggcactattatgtctgtacgtattattggctcatcgttccacctcagcgctggaagtgggtattctgaaaacgctgcata
 tgtttgaagtaccgttctgctggtgggctgcttaaatatattaccagccagtttgaagtgcgttttcagcgacgatttatctggtctgttctgct
 tctttaagcaactggcgatgttttatgtctgtactggcgggcaatatgtatgaaagcatcggttccagggcgcttactgtgctgggtctgg
 tggcgtggtggtcaccttaattccgtgttcacgcttagcggcccgccgcttccctgctgcgtcgcaggtgaatgaagtcgcttaa
 (SEQ ID NO: 79) which encodes the following amino acid sequence:

MYYLKNTNFWMFGLFFFFYFFIMGAYFPFFPIWLHDINHISKSDTGIIFAAISLFSL
 LFQPLFGLLSDKLGLRKYLLWIITGMLVMFAPFFIFIFGPLLQYNILVGSIVGGIYLGFCFN
 AGAPAVEAFIEKVSRRSNFEFGRARMFGCVGWALCASIVGIMFTINNQFVFWLGSGCAL
 ILAVLLFFAKTDAPSSATVANAVGANHSASFSLKLALELFRQPKLWFLSLYVIGVSCTYDV
 FDQQFANFFTSFFATGEQGTRVFGYVTTMGELLNASIMFFAPLIINRIGGKNALLLAGTIM
 SVRIIGSSFATSALVVILKTLHMFEPFLLVGCIFYITSQFEVRFSATIYLVCFCKQLA
 MIFMSVLAGNMYESIGFQGAYLVLGLVALGFTLISVFTLSGPGPLSLLRRQVNEVA*
 (SEQ ID NO: 80)

[0215] In the case of cells engineered to express LacZ (*B. coagulans*), cells were further modified to express a heterologous nucleic acid sequence lacY (*B. megatarium*), the heterologous nucleic acid sequence including the following sequence:

atgaaaagtagtaagtcactctactggaagctttctgcgtatttcttcttttctttacctggagctctagtactctctgtttccattt
 gggtgggacaggagataaagctgaatggctcagccacggggctcatatttagtgtaacgccatattcgctctttgtatgaaccattatacgg
 atatatctccgacagaatcggcctcaagaagcatatttttttataaagtgccctctgtattgttgggccattctacatatttgtatgggccg

ttattgcagtataatgtgctcataggtgccattattggtggcctgtacttgggcgtggcattttggcaggaataggcgcgatagaaacgtatatt
 gagaaggtatctcgcaagtacaagttcgagtatggaaagtctcgatgtgggggagcttgggtgggcccgcgcgacgtttttgcgggcca
 acttttcaatatcaacccgcacatcaattttgggtggccagcgtatccgctgttatacttatggctataatcttctcagtaaaagtgaatgagct
 cttatgaaatggagaaggcagaatcagtcgctccgctgatgtaggttaactgttctctaaaggaattctgggttttcatgatctatgctgtagg
 tgtaacatgtgtctatggggtgtacgaccaacagttcccaatatactatgcgtctttattccaaccgagtcacatcggaatcaagtgttcgggta
 tctcaatagtttccaagtctttctcgaggcaggatgatgttcgccgcgccatttattgtaacaaaataggcgcgagaattccttaactctggc
 tggtttctcatgggctttagaattattggttccgggtggtgtgggtcctataggaatcagttctatgaagcttatacacgcgctgaactccta
 taatgctcatagccatttttaagtacctcgccgcaattttgatacaagattatcatctattttgtacctggttggcttccaatttgcagtcagattg
 gcgcctctgtcctctccctatcgccggtggctgtatgactcagtcggatttagtcgcacttatctgatcatgggtgggatggtactgtttttaa
 tgttatttcaatgttcacattgttgaatgcaaaaagcataaatttatccggaaggacgttcaagaaaagactcagataatt (SEQ ID NO:
 81) which encodes the following amino acid sequence:

MKSSKSLYWKL SAYFFFFFFWSSSYSLFSIWL GQEIKLNGSATGLIFSVNAIFALC
 MQPLYGYISDRIGLKKHILFFISCLLVFVGPFYIFVYGPLLQYNVLIGAIIGGLYLGVAF LA
 GIGAIETYIEKVS RKYKFEY GKS RMWGS LGWAAATFFAGQLFNINPHINFWVASVSAVIL
 MAIIFSVK VEMSSYEMEK AESVRLRDVGNL FLLKEFWFFMIYVVG VTCVYGVYDQQFPI
 YYASLFPTESIGNQVF GYLNSFQVFLEAGMMFAAPFIVNKIGAKNSLILAGFLMGFRIIGS
 GLVVGPIGISSMKLIHALELPIMLIAIFKYLAANFDTRLSSILYLVGFQFASQIGASVLSPIA
 GGLYDSVGF SRTYLIMGGMVL VFNVISMFTLLNSKKHKFIRKDVQEKTQII (SEQ ID NO:
 82)Fiber synthesis by various strains of the present Example was measured by release of glucose
 as has been established in literature, where release of glucose corresponds to degradation of
 lactose. In brief, a suspension of active cells was mixed with Lactose and incubated. Periodically
 an aliquot was taken, the cells were removed from the aliquot by centrifugation, and the glucose
 content of the remaining sample of aliquot was assayed by glucose meter. Glucose is released
 upon degradation of Lactose and galactose is concurrently incorporated into transgalactosylation
 reactions depending on the reaction conditions and enzyme present. A culture of *B. subtilis*
 engineered to express LacZ (ZB420) was capable of degrading Lactose at a rate of 163 uM/min,
 13-fold faster than probiotic *Lactobacillus acidophilus* (Fig. 8).

Example 5: Construction and Experimental Validation of Engineered Cells that Express an Enzyme for Synthesis of Soluble Trehalulose Fiber from Carbohydrate Substrate

[0216] A transgene was constructed for expression of a trehalulose disaccharide fiber-synthesizing enzyme. The transgene was produced using standard techniques of molecular biology. The transgene included nucleic acid sequences encoding a fiber-synthesizing enzyme that included a secretion polypeptide. The present example utilizes mutB trehalulose synthase enzyme. The nucleic acid encoding the fiber-synthesizing enzyme was derived from *Pseudomonas mesoacidophila* and has the following sequence, with the secretion polypeptide encoding sequence bolded and underlined:

atgaacattaagaagttcgccaagcaagcaacgggtttaacgtttacaacagcactgctggcaggcggagcgacacag
gcttttgcattgatgaaaagattgttcgcgccctcattaatgtctcgcttctcctctgtgtcgtccgttcgcgcggaagaggccgtgaagccg
 ggagcaccatggtgaaatctgctgtgtttatcaggtgtacccgcggtctttcaaggatacgaacggggatgggaattggagactttaaggtg
 ctcacggaaaagctggattaccttaaggggctgggcattgacgccatatggataaacctcattacgctagtccgaacaccgataatggta
 cgatatctcagactatcgggaagtaataaagagtatggcactatggaagactttgaccggcttatggcagaattgaagaagagaggcatg
 aggctgatggtggacgtggaatcaaccactcaagcgatcagcacgaatggttaagtcttctagggcatccaaagacaaccctaccgtga
 ctactactctgcgcgacggtaagatggccatgaaccgaacaattaccctgctattcttcggcgatcggttgggagaaggaccggta
 actggccagtattatcttactacttcggtaggcagcaaccggattgaaactgggacacacctaagttgctgaggagtatacgcgatgctc
 agattctggctggacaaaggtgtgtctggcatgcgtttgacaccgtagcaacctacttaagacaccgggattccggatcttactctgaa
 cagatgaagaattttgcggaagcatacactcagggtcctaacctgcaccgatacttacaagaaatgcacgaaaaggtctttgatcactacga
 cgctgtcacggcgggagagatcttcggtgcccctcacaacaaagtaccgcttttcacgagtcgcccgaaggaattagatatggccttca
 ctttcgacttaataagatacgcgtgtcttgacagatggcacaccattccgaggacattagctgattccgtcaaacaatcgataaagttgac
 gcgattgcaggcggagtaggggtggaacaccttctttaggaacctagacaacctagagcagtgacatttcggcgacgacgcccac
 aatggagagaggcaagtgcgaaggcgtggctaccgtgactttaacacagcggggaacaccgttcattctccaggagacgagcttggga
 atgaccaattaccatttaagacactgcaagactttgatgacatcgaggttaagggtctttcaagactacgtcgagactggttaaggccaca
 gccgaggaattactgacaaacgtggccttgactagtcgtgacaatgcgagaacgccttccaatgggatgactcagctaacgtggattcac
 aaccggcaagccttggctcaaggtcaatcctaactacacagaaataaacgctgcgcgagattggggatcccaagtcagtctactcttct
 atcgcaacctgatctcgatccgcatgagactctgccccttcgaccggatcgatatagatatagacccagtaatgcagatgtatacgct
 atacgcgtcccaagacggagaaacctacttggtggtggtcaatttcaaagccgagcctaggagttcaccttaccagacgggtatgcatatc
 gccgagaccctattgagcttctagtccagctgcgctgtgccgggtgcagcgagcttagagttacaaccgtggcaatcgggcatctacaa
 ggtgaagtag (SEQ ID NO: 83).

[0217] The protein product MutB has the following amino acid sequence, with the secretion polypeptide sequence bolded and underlined:

MNIKKFAKQATVLTFTTALLAGGATQAFALMKRLFAASLMLAFSSVSSVRAE
EAVKPGAPWWKSAVFYQVYPRSFKDTNGDGIGDFKGLTEKLDYLGKLGIDAIWINPHY
ASPNTDNGYDISDYREVMKEYGTMEDFDRLMAELKKRGMRLMVDVVINHSSDQHEWF
KSSRASKDNPYRDYYFWRDGGKDGHEPNNYPSSFFGGSWEKDPVTGQYYLHYFGRQQP
DLNWDTPKLREELYAMLRFWLDKGVSGMRFDTVATYSKTPGFDDLTPQMKNF AEAY
TQGPNLHRYLQEMHEKVFDHYDAVTAGEIFGAPLNQVPLFIDSRRELDMAF TFDLIRY
DRALDRWHTIPRTLADFRQTIDKVD A IAGEYGWNTFFLG NHDNPRAVSHFGDDR PQWR
EASAKALATVTLTQRGTPFIFQGDELGMTNYPFKTLQDFDDIEVKGFFQDYVETGKATA
EELLTNVALTSRDNARTPFQWDDSANAGFTTGK PWLKVNP NYTEINAAREIGDPKSVYS
FYRNLSIRHETPALSTG SYRDIDPSNADVAYTRSQDGETYL VVVNFKAEP RSFTLPDG
MHIAETLIESSPAAPAAGAASLELQPWQSGIYKVK* (SEQ ID NO: 84).

[0218] The nucleic acid sequence encoding fiber-synthesizing enzyme derived from *P. mesoacidophila* was operably linked with a nucleic acid sequence encoding a secretion polypeptide, such that the encoded fiber-synthesizing enzyme was a fusion polypeptide including a secretion polypeptide. The nucleic acid sequence encoding the fusion polypeptide was analyzed for secondary structure, at least in part because significant secondary structure can cause translation termination and poor expression. Base pairs that significantly contributed to secondary structure in the nucleic acid sequence were modified by silent modifications of nucleic acid sequence (e.g., non-coding and/or synonymous modifications of nucleic acid sequence) that reduced the contribution to secondary structure.

[0219] The nucleic acid sequence encoding the fiber-synthesizing enzyme was transformed into *B. subtilis* PY79 cells and operably linked to *hag* promoter by homologous recombination at the endogenous *hag* locus of *B. subtilis* PY79 genome with the addition of a modifying mutation in the promoter sequence by homologous integration as described in US 16/048,147 (published as US 2019/0076489) and PCT/US18/50957 (published as WO 2019/055707) herein incorporated by reference in their entirety and with respect to regulatory sequences and systems disclosed therein.

[0220] Fiber-synthesizing enzymatic activity was measured by the following assay. Because trehalulose synthase rearranges sucrose into trehalulose it renders the sucrose inaccessible for sucrase. The accumulation of glucose after treatment with sucrase reveals remaining sucrose only, as trehalulose is not cleaved and does not release glucose. The glucose released by sucrase is directly proportional to the amount of sucrose remaining, which when subtracted from the amount of sucrose at the beginning of the reaction, yields the amount of trehalulose synthesized. Various glucose meters are known in the art and can be standardized for measurement of glucose concentrations in bacterial media. Engineered bacteria of the present Example were cultured in media to a defined density of 1E9 cells per ml (Optical Density of 1). The cells were then removed by centrifugation (21xg for 2 minutes) and the supernatant, containing secreted enzyme, was assayed by the addition of a defined amount of sucrose. At the beginning of the assay and at the conclusion of 18 hours the level of sucrose in the reaction was measured by cleavage to glucose using sucrase and the difference in sucrose concentration relative to a control strain lacking trehalulose synthase is directly proportional to trehalulose produced. Trehalulose production (g/L) using supernatant containing secreted enzyme or using whole cell broth is shown in Fig. 9 and Fig. 10. Those of skill in the art will appreciate that a small fraction of isomaltulose can also be produced, but that such production would not reduce or negate the advantages or utility of disclosed embodiments.

[0221] Those of skill in the art will appreciate that a variety of alternative methods are readily available for quantification (e.g., of the amount and/or concentration) of glucose, disaccharide trehalulose, and the like, including without limitation HPLC and HPAEC.

[0222] The present Example demonstrated that a transgene encoding a fiber-synthesizing enzyme derived from F1-mutB of *P. mesoacidophila* expressed a highly active fiber-synthesizing enzyme.

[0223] The present disclosure further includes the recognition that trehalulose synthase (mutB) derived from *P. mesoacidophila* has certain exemplary qualities that confirm the utility of MutB expression in gut. The optimum pH range for trehalulose synthase activity of MutB is 5.5-6.5 which matches that of the small intestine. MutB enzyme is naturally secreted and therefore stable outside of the engineered cell, amenable to secretion through a gram positive cell envelope.

OTHER EMBODIMENTS

[0224] While we have described a number of embodiments, it is apparent that our basic disclosure and examples may provide other embodiments that utilize or are encompassed by the compositions and methods described herein. Therefore, it will be appreciated that the scope of is to be defined by that which may be understood from the disclosure and the appended claims rather than by the specific embodiments that have been represented by way of example.

SEQUENCE LISTING

SEQ ID NO: 1

MNTDQQPYQGQTDYTGQPGNGQSSEQDYDQYGQPLYPSQADGYYDPNVAAGTEADMYGQQPPNESYD
 QDYTNGEYYGQPPNMAAQDGENFSDFSSYGPPGTPGYDSYGGQYTASQMSYGEPNSSGTSTPIYGNYDPN
 AIAMALPNEPYPAWTADSQSPVSIEQIEDIFIDL TNRLGQFQRDSMRNMFDHFMVLLDSRSSRMSPDQALLSL
 HADYIGGDTANYKKWYFAAQLDMDDDEIGFRNMSLGKLSRKARKAKKKNKKAMEEANPEDTEETLNKIE
 GDNSLEAADFRWKAKMNQLSPLERVRHIALYLLCWGEANQVRFTAECCLFIYKCALDYLDSPLCQQRQEP
 MPEGDFLNRVITPIYHFIRNQVYEIVDGRFVKRERDHNKIVGYDDLNLQFLWYPEGIAKIVLEDGTKLIELPLE
 ERYLRLGDVVWDDVFFKTYKETRTWLHLVTNFNRIWVMHISIFWMYFAYNSPTFYTHNYQQLVDNQPLA
 AYKWASCALGGTVASLIQIVATLCEWSFVPRKWAGAHL SRRFWFLCIIFGINLGPPIFVFA YDKD TVYSTA
 AHVVAAMVFFVAVATIIFFSIMPLGGLFTSYMKKSTRRYVASQTFTAFAPLHGLDRWMSYLVWVTVFAA
 KYSESYFVLVSLRDPILSTTAMRCTGEYWWGAVLCKVQPKIVLGLVIATDFILFDTYLVYIIVNTIFS
 VGKSFYLGISILTPWRNIFTRLPKRIYSKILATTDMEIKYKPKVLISQVWNAIIISMYREHLLAIDHVQKLLYH
 QVPSEIEGKRTLRAPTFFVSQDDNNFETEFFPRDSEAERRISFFAQSLSSTPIPEPLPVDNMTFTVLT PHYAERI
 LLSLREIIREDDQFSRVTLLEYLQLHPVEWECFVKDTKILAEETAA YEGNENEAEKEDALKSQIDDLPFYCI
 GFKSAAPEYTLRTRI WASLRSQTL YRTISGMNYSRAIKLLYRVENPEIVQMFGGNAEGLERELEKMARRK
 FKFLVSMQRLAKFKPHELENAEFLLRAYPDLQIAYLDEEPLTEGEEPRISALIDGHCEILDNGRRRPKFRV
 QLSGNPILGDGKSDNQNHALIFYRGEYIQLIDANQDNYLEECLKIRSVLAEFEELNVEQVNPYAPGLRYEEQ
 TTNHPVAIVGAREYIFSSENSGVLDVAAGKEQTFGTLFARTLSQIGGKLHYGHPDFINATFMTTRGGVSKA
 QKGLHLNEDIYAGMNAMLRGGRKHCEYYQCGKGRDLGFGTILNFTTKIGAGMGEQMLSREYYLGTQL
 PVDRFLTFFYYAHPGFHLNNLFIQLSLQMFMLTLVNLSSLAHESIMCIYDRNPKPTDVLVPIGCVNFQPAVD
 WVRRYTLSIFIVFWIAFVPIVVQELIERGLWKATQRFFCHLLSLSPMFEVFAAGQIYSSALLSDLAIGGARYIST
 GRGFATSRIPFSILYSRFAGSAIYMGARSMLMLLFGTVAHWQAPLLWFWASLSSLIFAPFVFNPHQFAWEDF
 FLDYRDYIRWLSRGNNQYHRNSWIGYVRMSRARITGFKRKL VGDESEKAAGDASRAHRTNLIMAEIIPCAI
 YAAGCFIAFTFINAQTGVKTTDDDRVNSVLRHICTLAPIAVNLGVLF FCMGMSCCSGPLFGMCKKTGSVM
 AGIAHGVAIVIVHIAFFIVMWVLESFNFVRMLIGVVTICQQRILFHCMTALMLTREFKNDHANTAFWTGK
 WYGKGMGYMAWTQPSRELTAKVIELSEFAADFLGHVILICQLPLIIPKIDKFHSIMLFWLKPSRQIRPPIYS
 LKQTRLRKRMMVKKYCSLYFLVLAIFAGCIIPAVASAKIHKHIGDSL DGVVHNLFPINTNNDTGSQMSTY
 QSHYYTHPSLKTWSTIK

SEQ ID NO: 2

ATGAATACGGACCAACAACCATATCAAGGCCAAACAGACTACACGCAAGGTCCTGGGAACGGGCAAT
 CCCAGGAACAAGACTACGATCAGTATGGCCAGCCATTGTATCCGAGCCAGGCCGATGGGTATTATGAT
 CCGAATGTGGCAGCGGGAACGGAGGCAGATATGTACGGACAACAACCCCGAATGAGTCGTATGATC
 AAGACTATACTAATGGCGAATACTATGGCCAACCGCCTAATATGGCGGCGCAGGATGGCGAAAAATTTT
 TCGGACTTTTCAAGTTACGGTCCGCCGGGCACACCGGGGTACGATTCATATGGCGGCCAATATACCGC
 AAGCCAAATGTCTTATGGCGAGCCGAACAGCTCAGGCACATCAACCCCGATTACGGAAATTACGACC
 CTAACGCAATTGCCATGGCTCTGCCGAATGAACCATATCCCGCATGGACGGCTGATAGCCAGAGTCCG
 GTCAGTATTGAACAAATTGAAGATATTTTATTGATTTGACGAACCGGTTAGGATTTACGCGGGATAG
 CATGCGGAATATGTTTGATCATTTTATGGTTCTGCTTGACTCACGGTCCAGCAGAATGTCACCTGATCA
 AGCCTTATTATCATTACACGCCGACTACATTGGAGGAGATACAGCAAACCTACAAAAAGTGGTATTTTCG
 CGGCGCAGTTAGATATGGATGATGAGATCGGGTTTCGTAATATGTCTCTCGGAAAACCTGAGCAGGAAA
 GCAAGGAAGGCAAAAAAAGAATAAAAAAGGCGATGGAGGAGGCGAATCCGGAHRTNLIMAEIIPCAI
 GACACTGAATAAGATCGAAGGCGACAATTCTTTAGAAGCCGCGAGATTTCCGTTGGAAGGCAAAAAATG
 AACCAGTTGAGTCCACTGGAACGAGTTCGACATATCGCGCTGTATCTTTTATGCTGGGGTGAGGCTAA
 CCAAGTCCGGTTCACCGCCGAATGTTTATGTTTTATTTACAAATGTGCCTTAGACTACCTTGACTCTCCT
 CTGTGCCAACAACGCCAGGAACCGATGCCCCGAAGGCGACTTCTTAAACCGCGTGATTACGCCGATTTA
 TCATTTTCATTGCAATCAAGTGTATGAAATCGTTGATGGACGTTTTGTTAAACGCGAACGCGATCATAA
 TAAAAATCGTTGGCTACGATGATCTCAACCAACTTTTTTGGTACCCGGAAGGTATTGCTAAAAATGTATT
 AGAAGACGGCACGAAACTCATTGAGTTACCGTTGGAGGAAAGATACTTACGCCCTGGCGACGTGGTCT
 GGGACGATGTTTTTTTTTAAACCTACAAAGAAACACGTACGTGGTTACATCTTGTAAACGAATTTTAAACA
 GAATTTGGGTAAATGCATATCTCCATTTTTTGGATGTACTTTGCCTACAATAGCCCTACCTTTTATACACA
 TAACTACCAGCAGTTAGTCGACAATCAACCGCTGGCCGCATACAAATGGGCGTCTGTGCTTTAGGGG
 GCACAGTTGCGAGCTTAATACAGATCGTAGCAACACTGTGCGAATGGAGCTTCGTTCCGAGAAAAATGG
 CGGGGCGCTCAGCATTTGTACGCGCTTCTGGTTTCTTTGTATCATCTTCGGTATCAACCTGGGTCCG

ATTATATTTGTTTTTGTCTACGACAAAGATACCGTCTATTCTACTGCAGCCCATGTAGTTGCAGCAGTA
ATGTTCTTTGTGGCGGTAGCGACTATTATTTTTTTTCAATCATGCCTCTGGGCGGCCTGTTACCTCGT
ATATGAAGAAATCGACACGACGCTATGTAGCATCGCAAAACATTTACAGCCGCGTTTGCCCCGCTGCAT
GGGCTTGACCGTTGGATGTCATATCTGGTCTGGGTGACCGTATTCGAGCCAAATATTCCGAAAGTTAT
TACTTTCTTGCTTATCTTTGCGAGATCCGATTCTGATCTTAAGTACGACAGCAATGAGATGTACGGGG
GAGTATTGGTGGGGCGCCGTTCTTTGTAAAGTTCAGCCGAAAATCGTCTTGGGACTGGTGATTGCGAC
AGACTTTATTTTATTTTTTCTTGATACATATTTGTGGTATATTATCGTGAATACTATTTTTTCTGTTGGAA
AATCATTTTATCTGGGAATCTCGATTCTGACGCCTTGGCGCAACATCTTTACACGCCTTCCTAAAAGAA
TCTATAGTAAAAATTTTGGCCACAACCGATATGGAAATCAAATATAAACCGAAGGTGCTTATTAGTCAG
GTGTGGAATGCTATTATATATCGATGTATCGCGAACATCTTTAGCAATCGACCATGTTCAAAAGTTG
CTGTATCACCAGGTTCTTCAGAGATCGAAGGAAAGAGGACGTTGAGGGCGCCACCTTTTTCTGTGAG
TCAAGATGATAACAACCTTTGAAACAGAATTTTTTCCAAAGGACTCCGAGGCCGAAAAGACGGATTTCTT
TTTTTGACAATCTTTATCTACACCAATCCCGGAACCTCTCCAGTCGACAATATGCCGACGTTTACAG
TGCTCACACCCCACTATGCCGAAAGAATCCTCTTAAGCCTGAGGGAAATAATCCGAGAGGATGATCAG
TTTTCTCGTGTTACGCTGCTGGAGTATTTAAACAATTACATCCTGTTGAGTGGGAGTGCTTCGTGAAG
GATACGAAAATTTTGGCAGAAGAGACGGCTGCTTATGAAGGCAATGAAAACGAAGCTGAAAAAGAAG
ACGCGCTCAAGTCACAGATAGATGATCTCCCTTTTTATTGCATAGGCTTTAAATCTGCGGCCCGGAAT
ATACGCTTCGCACAAGAATATGGGCATCTTTAAGATCCCAAACGTTATACCGGACCATTAGTGGATTT
ATGAACTATTCTCGGGCAATTAACCTTTTTGTATAGAGTGGAACCCCGAAATTGTACAAATGTTTCGG
CGGAAATGCTGAAGGCCTGGAGAGGGAACCTCGAAAAAATGGCTCGCAGGAAATTTAAATTTCTGGTTT
CCATGCAACGCCTTGCGAAATTCAAACCCCATGAATTAGAGAATGCCGAATTTCTGTTGAGGGCATAT
CCGGAATTGCAAATAGCATACCTTGATGAAGAACCGCCACTGACGGAAGGTGAAGAGCCGAGAATAT
ATTCTGCACCTATTGACGGACACTGCGAGATCCTTGACAATGGCCGACGTAGGCTAAATTTAGAGTC
CAACTTTCTGGAACCCGATTCTTGGTGATGGGAAAGTCTGATAATCAAACCATGCTGTGATCTTCTAT
CGGGGAGAATATATTCAACTTATAGATGCAATCAGGATAACTATCTTGAGGAGTGCTGAAAAATTCG
GTCTGTTCTGGCTGAGTTTGAGGAACTTAACGTTGAACAGGTGAATCCCTACGCACCGGGACTCAGAT
ATGAGGAACAGACAACCAACCCAGTCGCTATCGTAGGCGCACGCAATACATATTCTCAGAGAA
TTCCGGTGTATTAGGCGACGTGCGCCCGCGTAAAGAACAACGTTTCGGCACCTTGTTTGCGCGGACGC
TCTCTCAAATTTGGGGGAAAACCTGCATTACGGACATCCTGATTTTATTAATGCAACGTTTATGACAACGC
GAGGAGGAGTGAGTAAAGCTCAGAAAGGCCTTCATTTAAATGAAGACATTTATGCCGGCATGAATGC
GATGTTGAGGGGCGGCAGAATAAAACATTGTGAATATTATCAGTGTGGAAGGTCGGGATTAGGA
TTCCGGAACAATCCTTAATTTTACGACCAAAAATCGGCGCTGGTATGGGCGAGCAAATGTTATCAGAGA
GTATTATTATCTTGGAACGCAATTACCGGTTGATCGATTCTTGACCTTTTACTATGCTCATCCGGTTTT
CATCTTAACAATCTGTTTATTCAACTGTCCCTGCAATGTTTCATGCTGACGCTTGTAATTTGTATCTC
TGGCCCATGAATCGATCATGTGTATCTACGATCGCAATAAACCAAGACAGATGTGCTGGTACCGATC
GGCTGCTATAATTTCAACCCGGCTGTGGACTGGGTAAGACGATATACACTTTCCATATTTATTGTCTTC
TGGATCGCTTTTGTACCCATTGTCTGTTTCAGGAGCTTATTGAACGCGGCTTGTTGGAAGCGACTCAAAG
ATTCTTTTGCCATCTTTTATCCCTCTCTCCAATGTTTGAGGTCTTCGCGGGTCAAATTTATTCTTCAGCG
CTGCTTAGCGACTTGGCTATTGGCGGCGGAGGTATATCTCAACAGGTTCGCGGGTTTGCTACGTCCCGT
ATACCTTTTTTCGATCCTCTATTCAAGATTTGCAGGCAGCGCTATTTACATGGGCGCACGATCTATGTTA
ATGTTGTTATTTGGAACAGTTGCTCATTGGCAGGCGCCACTTCTTTGGTTCTGGGCATCCTTGTCATCAT
TAATCTTCGCGCCGTTGTTTTTAAATCCGCATCAATTTGCCCTGGGAAGATTTTTTTTAGATTATCGGGA
TTACATTGCTGGCTGAGCCGAGGAAATAAACAGTATCATCGTAATTCATGGATTGGTTACGTACGAA
TGTCTCGTGCCCGTATTACAGGCTTTAAAGAAAGCTCGTCGGCGACGAATCCGAAAAAGCAGCGGGA
GATGCAAGCCGAGCCCATCGTACTAACCTGATCATGGCTGAAATTATCCCGTGCCTATCTATGCGGC
AGGGTGTTCATAGCGTTCACGTTTATAAACGCTCAGACAGGCGTAAAGACAACCGACGATGACCGCG
TCAATTCGGTTCTGAGGATCATCATATGTACCCTTGACCGATTGCCGTGAATTTGGGAGTGCTTTTCT
TTTGTATGGGTATGTCATGCTGCAGCGGTCTTTATTCGGTATGTGTTGCAAGAAGACAGGCTCCGTTA
TGGCTGGTATTGCACACGGTGTGCTGTGATTGTTTCATATTGCGTTCTTTATCGTGATGTGGTACTGG
AGTCTTTCAATTTTGTTCGCATGCTTATTGGTGTGGTGACATGTATTCAAGTGTGAGCGCTTGATCTTTCA
CTGCATGACGGCACTGATGCTGACTCGAGAATTTAAAAACGACCACGCGAATACAGCCTTTTGGACGG
GCAAATGGTATGGCAAAGGAATGGGATATATGGCCTGGACCCAGCCGAGTAGAGAATTGACGGCAA
AGTCATTGAACTTAGTGAATTTGCAGCAGATTTTCGTATTGGGTCACGTTATTCTGATCTGTCAGTTACC
ACTGATCATCATCCCCAAAATCGACAAGTTTCACTCAATTATGCTCTTCTGGCTGAAACCATCACGTCA
GATCCGTCGCCCTATCTATAGTCTCAAACAAACGAGATTGAGAAAAAGAATGGTTAAAAAATATTGCA

GCCTTTACTTTTTGGTCTTGGCCATTTTTGCGGGGTGTATCATCGGCCCCGCTGTAGCATCAGCCAAGA
TCCATAAACACATCGGTGATTCTTAGATGGAGTCGTACATAATCTTTTCCAACCGATAAATACCACAA
ACAATGACACGGGCAGTCAGATGTCAACTTACCAGAGTCACTACTATACTCATACGCCGCTCTTAAAA
ACCTGGTCCACAATTAAGTAA

SEQ ID NO: 3

MAQRREPDPPPPQRRILRTQTVGSLGEAMLDSEVVPSSLVEIAPILRVANEVEASNPRVAYLCRFYAFEKAH
RLDPTSSSGRGVRQFKTALLQRLERENETTLAGRQKSDAREMQSFYQHYYKKYIQALLNAADKADRAQLT
KAYQTAAVLFEVLKAVNQTEDVEVADEILETHNKVEEKTQIYVPYNILPLDPDSQNQAIMRLPEIQAAVAA
LRNTRGLPWTAGHKKLDEDILDWLQSMFGFQKDNVLNQREHLILLANVHIRQFPKPDQOPKLDDRALT
IVMKKLFRNYKKWCKYLGRKSSLWLPTIQQEVQQRKLLYMGLYLLIWGEAANLRFMPECLCYIYHHMAF
ELYGMLAGSVSPMTGEHVKPAYGGEDEAFLOKVVTPITYQTISKEAKRSRGGKSKHSVWRNYDDLNEYFW
SIRCFRLGWPMRADADFFCQTAELRLERSEIKSNSGDRWMGKVNFEIRSFWHIFRSFDRLSFYILCLQA
MIVIAWNGSGELSAIFQGDVFLKVLVSFITAAILKLAQAVLDIALSWKARHSMSLYVKLRYVMKVGAAAV
WVVVMAVTYAYSWKNASGFSQTIKNWFGGSHNSPSLFIYAILIYLSNMLSALLFLFPFIRYLERSDYKI
MMLMMWWSQPRLYIGRGMHESALSLFKYTMFVIVLLISKLAFSYYAEIKPLVGPTKDIMRIHISVYSWHEF
FPHAKNNLGVVIALWSPVILVYFMDTQIWIYAVSTLVGGLNGAFRRLGEIRTLGMLRSRFSIPGAFNDCLV
PQDNSDDTKKKRFRATFSRKFDQLPSSKDKEAARFAQMWNKIISFREEDLISDREMELLLVPYWSDPDL
LIRWPPFLASKIPIALDMAKDSNGKDRELKRLAVDSYMTCAVRECYASFKNLINYLVVGEREGQVINDIF
SKIDEHIEKETLITELNLSALPDLYGQFVRLIEYLLNREEDKDQIVIVLLNMLELVTDRIMEEVPSSLET
NGSYVKYDVMTPLHQQRKYFSQLRFPVYSQTEAWKEKIKRLHLLTVKESAMDVPSNLEARRRLTFFSNS
LFMDMPAPKIRNMLSFSVLTPYFSEDVLFISIFGLEQQNEDGVSYLFLQKIFPDEWTFNLERVKCGNEEELR
AREDLEELRLWASYRGQTLTKTVRGMYYRKALELQAFDMAKDEELLKGYKALETSEEASKSGGSL
WAQCQALADMKFTFVVSQCQYSIHKRSQDQRAKDILRLMTTYPYSIRVAYIDEVEQTHKESYKTEEKIYYS
ALVKAAPQTKPMDSSSVQTLQDIYRIKLPAILGEGKPNQNHAIIFTRGEGQLTIDMNQDNYMEEAFK
MRNLLQEFLEKHGGVRCPTILGLREHIFTGSVSSLAWFMSNQENSFVTIGQVRVLASPLKVRFHYPHDIFDR
LFHLTRGGICKASKVINLSEDFAGFNSTLREGNVTHHEYIQVGKGRDVGLNQISMFEAKIANGNGEQLSR
DLYRLGHRFDFRMLSCYFTTIGFYFSTMLTVLTVYVFLYGRLYLVLGLEEGSSQRAFRNNKPLEAALAS
QSFVQIGFLMALPMMMEIGLERGFHNLIEFVLMQLQLASVFFTFQLGKTHYYGRTLFHGGAEYRGTGR
GFVVFHAKFAENYRFYSRSHFVKGIELMILLVYQIFGQSYRGVVITYILITVSIWFMVVTWLFAPFLFNPSGF
EWQKIVDDWTDWKNWYINRGGIGVPPEKSWESWEKEHLRHSGVRGITLEIFLALRFFIFQYGLVYHL
STFKGKNQSFVYVYGASWFVILFILLIVKGLGVGRRRFFSTNFQLLFRIKGLVFLTFVAILITFLALPLITIKDLFI
CMLAFMPTGWGMLLIAQACKPLIQQLGIWSSVRTLARGYEIVMGLLLFTPVAFLAWFPFVSEFQTRMLFNQ
AFSRGLQISRILGGQRKDRSSKNKE

SEQ ID NO: 4

ATGGCACAAACGACGGAACCGGACCCGCCACCCCGCAGCGTAGAATATTGCGGACGCAAACGGTTG
GCTCTCTGGGAGAAGCCATGCTTGATTCTGAGGTCGTTCCCTTCATCACTTGTGCAAAATAGCGCCAATTT
TACGCGTCGCAAACGAAGTAGAAGCTTCAAATCCGCGTGTGCTTACTTATGTAGATTCTACGCATTTG
AAAAAGCTCACCGACTTGATCCAACCTCTTCAGGCCGCGGAGTTAGACAATTTAAGACGGCATTGCTT
CAACGGTTAGAGCGTGAGAACGAAACCAATTGGCGGGAAGACAAAAAGCGACGCGCGAGAGATG
CAAAGTTTTTACCAGCATTACTACAAAAATACATTCAAGCGTTACTGAATGCAGCGGACAAGGCAGA
TCGGGCGCAACTTACAAAAGCGTACCAGACAGCTCGAGTCCTTTTTGAAGTGCTTAAAGTGCTCAATC
AGACTGAAGACGTCGAAGTTGCCGATGAAATCTGGAGACACACAATAAAGTCGAGGAGAAAACGCA
GATTTATGTGCTTATAACATTCTTCCCTTAGATCCTGACAGCCAAAATCAGGCCATTATGCGCTTGCC
AGAAATTCAGGCGGCTGTGCGCCGACTGCGCAACACTCGTGGTTTACCTTGGACAGCAGGACACAAAA
AGAAATTGATGAGGATATTCTTGATTGGCTGCAATCGATGTTTCGGTTTTCAAAAAGATAATGTGCTC
AATCAAAGAGAACATCTGATTTTGTCTTGGCTAACGTTTCATATCCGCCAATTTCCCGAAGCCTGATCAG
CAGCCAAAACCTTGACGACCGGGCTCTTACGATTGTAATGAAAAAGTTATTTCCGAATTACAAGAAGTG
GTGCAAGTATCTTGGTCGCAAGTCATCCCTTTGGTCCCTACAATTCAACAGGAAGTGCAACAGCGTA
AATTACTTTATATGGGGCTCTACTTGTTAATATGGGGAGAACCCGCGAATCTTCGCTTTATGCCGGAAT
GTCTGTGCTATATCTATACCATATGGCATTGAGTTGTATGGGATGTTAGCAGGCAGCGTGTCTCCGA
TGACTGGCGAACACGTTAAGCCTGCATATGGAGGGGAGGACGAAGCCTTTTACAGAAGGTCTGACG
CCAATTTATCAAATATTTCTAAAGAAGCAAAAAGATCAAGAGGAGGAAAAAGCAAACATAGCGTGT
GGCGGAATATGATGATCTCAATGAGTATTTTGGAGTATTCGGTGTTTTCGGCTTGCTGGCCCATGC

GCGCGGATGCTGATTTTTTTTTGTCAAACAGCTGAAGAATTGAGGTTAGAGAGAAGCGAGATTAAATCG
AATAGCGGGGATCGTTGGATGGGGAAAGTAAACTTTGTAGAGATTAGATCATTTTGGCACATTTTATG
ATCTTTTGATAGATTGTGGTCTTTTACATACTGTGCTTGCAAGCGATGATTGTAATCGCTTGGAAATGG
CTCGGGTGAATTGTCGGCGATTTTTCAAGGAGATGTATTTTGAAGGTGCTCTCTGTCTTTATTACCGC
GGCGATCCTTGGCGCAAGCCGTTTCGATATTGCCCTGTCTGGAAGGCGCGTCAATTCGATGA
GCCGTGATGTTAAACTTCGTTACGTATGAAAAGTGGGTGCTGCGGCAGTTTGGGTGCTGCTTATGGCGG
TAACATACGCATATTCATGAAAAACGCGTCTGGCTTCTCCCAGACCATTAAAGAACTGGTTCGGCGGA
CATTCACATAATTCCCCGTCACCTCTTTATTGTGGCTATCCTGATTTATCTGAGCCCTAACATGTTGTCAG
CACTTCTTTTTCTCTTCCCGTTTATCCGTCGGTATTTGGAACGATCTGATTACAAAATCATGATGCTTAT
GATGTGGTGGTCCAGCCACGACTGTATATCGGAAGGGGGATGCATGAATCAGCTCTGTCTCTGTTTA
AATACACTATGTTTGGATTGTTCTGCTGATTTCGAACTTGCCTTTTCGTATTACGCGGAAATTAAC
CCCTCGTAGGCCGACAAAAGACATTATGCGAATTCATATTAGTGTTTATTCGTGGCATGAGTTTTC
CACATGCAAAAAACAATCTGGGTGTAGTCAATTGCACTTTGGTCACCCGTCATCCTGGTATATTTTCATGG
ACACACAAATTTGGTACGCTATCGTCTCCACCCTGGTGGGAGGCTTAAACGGTGCTTTTAGACGTTTAG
GGGAGATCAGAACATTAGGTATGTTACGTTGCGGCTTCCAGAGTATCCAGGTGCATTTAATGATTGTC
TTGTCCCGCAGGACAACTCCGACGACACTAAAAAAAAGCGCTTTAGAGCAACTTTTAGTCGGAAATTT
GATCAGCTTCCATCATCAAAGGACAAAGAGGCGGCAAGATTTCGCACAAATGTGGAATAAAATCATTTTC
AAGTTTTCGTGAAGAAGACCTGATTTCAGACCGGGAATGGAACCTTTGCTTGTACCTTACTGGAGTG
ATCCTGATTTGGACCTGATCAGGTGGCCGCCGTTTTTATTAGCATCCAAAATTCCTATCGCGCTGGACA
TGGCTAAAGACTCTAACGGTAAGGACCGTGAACCTCAAAAAGAGACTCGCCGTTGATTCTATATGACC
TGTGCAGTCCGTGAATGCTACGCGTCTTTCAAAAATTTAATTAATTATTTAGTTGTTGGAGAACGCGAA
GGGCAAGTCATTAATGATATCTTTTCAAAGATCGATGAACATATAGAAAAAGAGACCTTAATTACAGA
ACTTAATTTGAGCGCGCTGCCGATTTATACGGACAATTCGTGAGACTTATTGAATATCTGCTGGAAAA
TCCGTAAGAGGATAAAGATCAGATTGTTATAGTCTTATTAATATGCTGGAATTGGTAAGGACGCGGACA
TTATGGAGGAAGAAGTTCGCTCTTTGTTAGAAACGGCTCATAATGGATCTTACGTTAAGTATGATGTG
ATGACACCACTCCATCAGCAGCGTAAATATTTTAGTCAACTGCGGTTTCCGGTTTACAGCCAAACGGA
GGCGTGGAAAGAGAAAAATCAAACGACTGCATTTGCTGTTGACGGTCAAAGAATCGGCAATGGACGTA
CCGTCAAACCTTGGAAAGCGCGAAGAAGATTAACCTTCTTTTCTAATTCACGTTCATGGATATGCCTCCT
GCACCGAAAAATTCGTAATATGTTATCATTTTCAGTCTTAACTCCGTATTTCTCTGAAGATGTCCTTTTTA
GCATCTTCGGCCTTGAACAGCAGAATGAAGATGGAGTGTCATTCTTTTCTATCTTCAAAAAATTTTTTC
CGGATGAATGGACCAATTTTTTAGAGCGGGTCAAAATGCGGCAATGAGGAAGAACTGCGGGCCGTTGA
AGATCTTGAAGAAGAATTGCGACTTTGGGCCTCATATAGAGGTCAAACACTGACAAAAACAGTACGTG
GGATGATGTATTATAGAAAAGCTCTGGAAGTGCAGGCATTTTTAGACATGGCTAAAGATGAAGAATTA
TTAAAAGGTTACAAAGCTCTGGAGCTTACATCCGAGGAAGCGAGTAAGAGCGGAGGTTCTTTGTGGGC
TCAATGTCAAGCGTTGGCTGACATGAAGTTCACCTTCGTTGTTTCTTGCCAACAATATAGTATTCATAA
GCGTAGCGGTGATCAAAGAGCGAAGGATATCCTTCGTTGATGACAACGTATCCGAGCATCCGAGTTG
CATATATAGACGAGGTAGAGCAAAACGCACAAAGAGTCCATATAAAGGCACGGAAGAAAAAGATATATTA
CTCTGCTCTTGTGAAAGCCGCTCCACAGACAAAAGCCGATGGATTCTTCAGAAAGCGTACAAACATTGG
ATCAGTTGATTTACCGTATCAAACCTTCCGGGGCCAGCAATCCTGGGAGAAGGCAAAACCGGAAAAATCAG
AATCACGCAATCATTTTCACAAGAGGCGAAGGCCTTCAAACAATCGATATGAATCAGGATAATTATAT
GGAAGAAGCTTTCAAATGCGCAATCTGTTACAGGAATTCCTTGAAAAACATGGAGGCGTTAGATGCC
CTACAATCCTGGGCCCTTCGCGAACACATTTTTACTGGCAGTGTCAGCTCTTTAGCGTGGTTTATGTCCA
ACCAAGAAAACTCATTTGTCACTATAGGCCAGAGAGTCTTAGCGAGCCCTCTGAAAGTACGCTTTTAC
TATGGTCAATCCGGATATTTTTGATAGATTGTTTACCTTACCAGGGGTGGGATTTGTAAAGCCTCTAAG
GTCATCAACCTCAGCGAAGACATCTTTGCTGGCTTTAACAGCACACTTCGCGAAGGCAATGTCACCCA
CCATGAATATATTCAAGTTGGTAAGGGACGTGATGTGGGGTTGAATCAAATCTCGATGTTTGAAGCGA
AAATTGCCAATGGCAATGGAGAACAAACCTTGTCCCGGGATCTTTACCGGTTGGGTATCGTTTCGATT
TCTTTTCGTATGCTTTCTTGCTATTTTACCACGATTGGGTTTTATTTTTCTACCATGTTGACCGTCTGACT
GTATACGTCTTCTGTATGGGCGGCTGTATCTTGCTGAGCGGTCTGGAAGAAGGACTTAGCTCCCAA
CGGGCCTTTTCGCAACAACAAGCCTTTAGAAGCCGCACTTGCAATCACAAGTTTTGTTTCAGATCGGATTT
TTAATGGCATTGCCTATGATGATGGAAATTGGGCTCGAAAGGGGTTTTTACAATGCTTTAATAGAATTT
GTGCTTATGCAATTGCAGCTTGCATCCGATTTCTTTACTTTTCAATTAGGAACCAAGACTCACTACTAT
GGAAGGACATTATTTACGGCGGAGCAGAGTATAGAGGCACAGGCCGTGGATTTGTCTGTTTCATGC
GAAATTTGCCGAAAAATTATCGGTTCTATTCTAGGTCTCATTTTGTCAAAGGCATCGAGCTGATGATCCT
TCTTTTGGTCTATCAAATCTTCGGTCAATCATATAGGGGTGTAGTGACATACATCCTTATCACTGTAAG

CATATGGTTTATGGTTGTGACGTGGCTGTTTGC GCCTTTTCTGTTTAACCCAGTGGATTGAATGGCA
GAAGATCGTGGATGACTGGACCGATTGGAACAAATGGATCTATAATCGCGGCGGCATTGGGGTTCCAC
CGGAGAAATCTTGGGAGTCATGGTGGGAAAAAGAACTGGAACACCTCCGCCATAGCGGAGTCCGGGG
AATTACATTGGAATTTTCCTTGCCTTCGCTTTTTCATCTTTCAGTACGGGCTTGTCTACCATCTCAGC
ACATTTAAAGGCAAAAACCAAGTCATTTTGGGTCTACGGGGCCTCATGGTTTGTCTATTGTTTATTTTG
TTAATTGTAAAAGGTTTGGGCGTTGGACGGCGCCGCTTTTCTACTAATTTCCAGCTGCTTTTCCGCATC
ATTAAAGGTCTCGTTTTCTTAACGTTCTGTGCAATTCTGATTACTTTTCTGGCACTGCCGCTGATAACCA
TAAAGGATTTATTTATCTGTATGCTTGCATTTATGCCTACGGGGTGGGGTATGCTTCTGATAGCTCAGG
CGTGCAAACCGCTGATCCAGCAACTGGGAATCTGGTCTTCCGTCAGAACCTTAGCACGGGGGTACGAA
ATCGTTATGGGCTGTTACTGTTTACCCAGTGGCATTCTTGCCTGGTTTCCGTTTGTAAAGCGAATTCC
AAACGCGCATGCTGTTTAAATCAGGCGTTTCTCGCGGGTTGCAAAATAGCCGCATTTTAGGCGGCCAAC
GGAAGGATAGGAGCAGCAAAAATAAGGAG

SEQ ID NO: 5

MAPAVAGGGGRRNNEG VN GNAAAPACVCGFPVCACAGAAVASAASSADMDIVAAGQIGAVNDESWV
AVDLSDDSDAPAGDVQ GALDDRPVFRTEKIKGVLLHPYRVLIFVRLIAFTLFVIWRIEHKNPDAMWLWVT
SIAGEFWGFSWLLDQLPKLPINRVPDLAVLRRRFDHADGTSSLPGLDIFVTTADPIKEPILSTANSILSILAA
DYPVDRNTCYLSDDSGMLLTYEAMAEAAKFATLWVPFCRKHAIEPRGPESYFELKSHPYMGRAQEEFVND
RRRVRKEYDDFKARINGLEHDIKQRSDSYNAAAGVKDGEPRATWMADGSQWEGTWIEQSENHRKGDHA
GIVLVLLNHPSHARQLGPPASADNPLDFSGVDVRLPMLVYVAREKRP GCNHQKKAGAMNALTRASAVLS
NSPFILNLD CDHYINNSQALRAGICFMLGRDSDTVAFVQFPQRFEGVDPTDLYANHNRIFFDGLRALDGLQ
GPIYVGTGCLFRITLYGFEP PRINVG GPCFPRLLGGMFAKNRYQKPGFEMTKPGAKPVAPPPAATVAKGKH
GFLPMPKKAYGKSDAFADTIPRASHPSPYAAEAAVADEAAIAEAVMVTAAAYEKKTGWGS DIGWVYGT
VTEDVVTGYRMHIK WRSRYCSYPHAFIGTAPINLTERLFQVLRWSTGSLEIFFSRNNPLFGSTFLHPLQRV
AYINITTYPTALFLIFYTTVPALSFTVGHFIVQRPTTMFYVYLAIVLGTLLILAVLEVKWAGVTVFEWFRNG
QFWMTASCSAYLA AVLQVVTKVVFRRDISFKLTSKLPAGDEKDPYADLYVVRWTWLMITPIIHLVNIIGS
AVAFAKVLDGEWTHWLKVAGGVFFNFVWLFHLYPFAKGILGKHGKTPVVVLVWWAFTFVITAVLYINIP
HIHGPRHGAASPSHGHS AHGTKKYDFTYAWP

SEQ ID NO: 6

ATGGCGCCGGCTGTGCGGGGAGGCGGTGGCCGTCGAAATAATGAAGGAGTCAACGGCAACGCCGAG
CTCCAGCCTGCGTATGCGGCTTTCCGGTTTGC GCATGCGCCGGGGCTGCCGCTGTTGCTAGTGCGGCCT
CTTCCGCTGATATGGACATTGTGGCGGCCCGCCAGATCGGCGCGGTCAACGACGAATCATGGGTGCC
GTCGATCTGAGCGATAGTGACGACGCCCGCGCGGGTGATGTACAAGCGCTCTCGATGACCGTCC
GGTGTTTAGAACAGAGAAAATTAAGGCGTTCTCCTTACCCGTATAGAGTGCTTATTTTGTACGGCT
GATCGGTTTACGCTTTTCTGCATCTGGAGGATAGAGCATAAGAACCCGGATGCGATGTGGCTGTGGG
TTACATCTATAGCAGGTGAATTTTGGTTTGGTTTTCTTGGCTTCTTGATCAGTCCCCGAACTTAATCC
TATAAACAGAGTTCCAGACTTGGCTGTATTGCGGAGACGATTCGATCATGCGGATGGTACTTCGTCATT
ACCCGGA CTGGATATCTTCGTTACAACAGCTGATCCAATTAAGAGCCTATTTTGAGCACTGCGAACA
GTATCCTGTCAATTTTAGCTGCCGATTATCCGGTTGATCGGAACACATGTTATCTTAGCGATGATTGAG
GAATGCTGCTCACATATGAAGCAATGGCTGAGGCAGCTAAATTTGCCACATTATGGGTGCCGTTTTGT
CGTAAACATGCTATTGAGCCGAGGGGACCGGAAAGCTATTTGAATTAAAATCTCATCCTTATATGGG
TAGGGCACAGGAAGAATTTGTTAACGATCGTCAAGGGTAAGGAAAGAATATGATGATTTTAAAGCA
CGGATTAACGGCTTAGAACACGATATCAACAACGCTCAGATTATACAAACGCGGCAGCAGGTGTTAA
AGACGGGGAACACGCTGCTACATGGATGGCAGATGGCTCACAGTGGGAAGGGACATGGATCGAACA
TCGGAATATCATGAAAAGGCGATCACGCAGGTATAGTGCTTGTCTGTTAAATCATCCTTCACATGC
ACGTC AATTAGGACCACCGGCAAGCGCAGATAATCCTTTGGACTTTTCCGGAGTTGACGTTAGGTTAC
CTATGCTTGTATACGTAGCCCGTGAAAAACGGCCCGGATGCAACCACCAAAAAAAGCGGGAGCAAT
GAACGCGTTGACGCGCGCTTCCGCGGTCCTCAGCAATTCTCCGTTTATACTTAATCTGGACTGCGATCA
CTACATTAACAACTCACAAGCGCTTCGTGCAGGGATATGCTTTATGTTGGGACGTGATAGTGACACAG
TGGCTTTCTGTGCAATTTCCGCAAGATTTGAAGGCGTAGATCCGACCGACCTGTACGCAAAACCACAAT
AGAATTTTGTGATGGTACCCTGAGAGCTTTAGATGGGTTACAGGGGCCGATTATGTTGGTACGGGC
TGCTTGTTTAGACGCATAACCTTATATGGTTTCAACCGCCGAGAATCAATGTTGGTGGCCCCCTGCTTT
CCTCGCTCGGTGGGATGTTTGCCAAGAATCGGTATCAGAAGCCTGGCTTCGAGATGACTAAGCCGGG
AGCGAAACCGGTTGCACCTCCACCGGCTGCTACAGTAGCCAAGGGGAAGCATGGCTTTTACCGATGC

CGAAAAAGCGTATGGGAAAAAGCGACGCATTTGCGGATACAATTCCTAGAGCGTCACACCCTAGTCC
GTATGCCGCCGAAGCAGCTGTTGCCGCCGACGAGGCTGCAATTGCCGAAGCGGTGATGGTCACAGCTG
CCGCTTATGAGAAAAAGACGGGTTGGGGGTCAGACATCGGCTGGGTTTACGGAACAGTTACAGAGGA
TGTCGTTACTGGATATAGAATGCACATCAAAGGCTGGCGGTACGCTACTGTAGCATCTACCCGCATG
CATTCATCGGTACCGCTCCGATTAATTTGACAGAACGACTGTTTCAAGTCTTACGCTGGAGCACAGGAT
CTTTGGAAATCTTCTCTCTAGAAATAATCCGTTATTCGGTAGCACATTCTCCATCCATTGCAGAGGG
TTGCGTATATTAATATCACACCTATCCATTTACTGCACTGTTTCTTATTTTTTATAACAACAGTCCCGGC
GCTGTCTTTTGTACGGGTCACTTCATTGTACAACGCCCCGACAACGATGTTTTACGTGTATCTCGCAAT
AGTACTGGGCACGCTGCTGATCTTAGCCGCTCTAGAAAGTTAAGTGGGCTGGGGTAACGGTATTCTGAAT
GGTTCCGCAACGGACAATTTTGGATGACGGCATCTTGTTCAGCTTATTTAGCTGCTGTCTTCAAGTCG
TTACGAAGGTTGTGTTTCGCCGGGACATTTTCGTTTAACTGACATCGAAATTGCCCGCTGGGGATGAA
AAGAAAGATCCATATGCGGATCTTTACGTCTGTGCGGTGGACCTGGTTAATGATTACGCCGATCATTAT
CATCTCTGTTAACATCATCGGCTCTGCAGTCGCGTTTCGCTAAAGTGCTTGATGGAGAATGGACCCATTG
GTTAAAAGTGGCGGGCGGCGTCTTTTTTAACTTCTGGGTCTTGTTCATTTATACCCTTTTCGCAAAGGG
AATTTTAGGCAAGCATGGGAAAACGCCTGTTGTGGTCTGTTTGGTGGGCTTTCACCTTTGTTATTAC
AGCGGTTCTTTACATCAACATCCCTCATATCCATGGGCCGGGAAGACATGGCGCCGCAAGCCCGTCTC
ATGGACATCATTCTGCACATGGCACAAAAAATATGATTTACGTACGCATGGCCATAA

SEQ ID NO: 7

MRLQRNSIICALVFLVSFVLGDVNIVSPSSKATFSPSGGTVSVPVEWMDNGAYPSLSKISTFTFSLCTGPNNN
IDCVAVLASKITPSELQDDKVYSYAEFASTLTGNGQYYIQVFAQVDGQGYTIHYTPRFQLTSMGGVTAY
TYSATTEPTPQTSIQTTTTNNAQATTIDRSFTVPYTKQTGTSRFAPMQMQPNTKVTATTWTRKFATSAVTY
YSTFGSLPEQATTTTPGWSYTISSGVNYATPASMPSDNGGWYKPSKRLSLSARKINMRKV

SEQ ID NO: 8

ATGCGGCTTCAACGTAACAGTATCATTGTGCTTAGTTTTTTAGTATCTTTTGTCTTGGGTGATGTCA
ACATTGTGTCGCCGTCAGCAAAGCTACATTTTCTCCAAGTGGCGGCACAGTGTCCGTTCTCTGTCGAAT
GGATGGACAACGGAGCTTATCCGAGCCTTTCAAAAATTTCTACATTTACATTTAGTCTGTGCACGGGTC
CAAATAATAATATTGATTGTGTTGCTGTCTGGCAAGTAAAATAACGCCTAGCGAATTAACCTCAGGAT
GACAAAGTTTATTCTTATACGGCTGAATTTGCTTCAACGTTGACCGGGAATGGTCAATACTATATCCAA
GTATTTGCCCAAGTAGATGGCCAAGGCTACACAATCCATTACACGCCGCGGTTCCAGCTCACGTCAAT
GGGGGGCGTGACAGCGTACACATATAGTGCAACGACCGAACCTACACCCAGACTAGTATTCAAACA
ACAACGACAAATAATGCGCAGGCAACTACCATAGATTCCCGGTCCTTACAGTCCCTTATACTAAACA
AACGGGCACCTCACGCTTTGCGCCTATGCAGATGCAGCCGAATACTAAAGTGACCGCGACAACGTGGA
CGAGGAAATTCGCCACCTCAGCAGTTACCTATTACTCTACTTTTGGATCGTTGCCTGAACAAGCAACAA
CAATTACACCCGGCTGGTCGTATACTATATCAAGCGGGGTAACTATGCGACACCTGCTAGTATGCCT
AGCGATAACGGAGGTTGGTATAAGCCAAGCAAACGATTGAGCCTTTCAGCGCGGAAATTAACATGC
GAAAAGTTTGA

SEQ ID NO: 9

MLMKRLFAASLMLAFSSVSSVRAEEAVKPGAPWWKSAVFYQVYPRSFKDTNGDGIGDFKGLTEKLDYLK
GLGIDAIWINPHYASPNTDNGYDISDYREVMEKEYGTMEDFDRLMAELKKRGMRLMVDVVINHSSDQHEW
FKSSRASKDNPYRDYYFWRDGDGHEPNYPSFFGGSAAWEKDPVTGQYYLHYFGRQQPDNLNWDTPKLE
ELYAMLRFWLDKGVSGMRFDTVATYSKTPGFPLDTPQMKNF AEAYTQGPNLHRYLQEMHEKVFDHYD
AVTAGEIFGAPLNQVPLFIDSRRKELDMAFTFDLIRYDRALDRWHTIPRTLADFRQTIDKVDIAIGEYGWNT
FFLGNHDNPRAVSHFGDDRPQWREASAKALATVTLTQRGTPFIFQGDELGMTNYPFKTLQDFDDIEVKGFF
QDYVETGKATAEELLTNVALTSRDNARTPFQWDDSANAGFTTGKPLKVNPNYTEINAAREIGDPKSVYS
FYRNLISIRHETPALSTGSYRIDPSNADVAYTRSQDGETYL VVVNFKAEP RSFTLPDGMHIAETLIESSPA
APAAGAASLELQPWQSGIYKVK

SEQ ID NO: 10

ATGCTGATGAAGAGATTATTCGCGGCCAGCTTAATGCTGGCCTTCAGTTCAGTATCAAGTGTTAGAGC
AGAGGAAGCAGTAAACCTGGAGCGCCGTGGTGGAAGTCAGCAGTTTTTTATCAAGTTTATCCTAGGA
GTTTTAAGGATACAAATGGTGATGGGATTGGCGACTTTAAGGGTCTTACGGAGAAACTTGACTATCTT
AAAGGCCTTGGAAATCGATGCGATCTGGATAAATCCTCATTATGCATCCCCTAACACGGATAATGGCTA

TGATATTAGCGACTATCGTGAAGTAATGAAAGAGTATGGGACGATGGAGGATTTTCGACCGGCTGATGG
 CGGAGCTGAAAAACCGCGCATGCGCCTTATGGTAGATGTAGTGATCAACCATTTCATCTGATCAGCAC
 GAATGGTTCAAATCCTCTCGTGCCAGTAAAGATAATCCGTATCGTGATTACTATTTTTGGCGTGATGGA
 AAAGACGGTCATGAACCTAATAATTATCCTTCTTTCTTCGGCGGCTCAGCATGGGAAAAGGATCCGGT
 TACCGGGCAATACTATCTCCATTACTTCGGAAGACAGCAGCCAGACTTGAAGTGGGATACACCGAAAC
 TCGGGAGGAGCTGTATGCGATGCTTCGTTTCTGGCTGGATAAAGGCGTCTCTGGAATGCGGTTTGAT
 ACGGTGGCCACATATTCTAAAACGCCGGGATTTCCGGACCTGACGCCTGAGCAGATGAAAACTTTGC
 CGAAGCTTATACCCAGGGTCCGAACCTCCACCGGTATCTGCAAGAAATGCATGAAAAAGTGTGATC
 ATTATGACGCGGTACCCGCTGGGGAGATTTTGGCGCGCCGCTTAACCAAGTTCGCTGTTTATTGATT
 CAAGAAGAAAAGAGCTTGATATGGCTTTCACCTTCGACTTGATTAGATACGACCGTGCTTTAGACCGG
 TGGCATAACAATTCCGCGCACGTTAGCTGATTTCGTCAGACGATCGACAAGGTGGATGCCATAGCCGG
 CGAATACGGCTGGAATACATTTTCTTGGGGAACCATGACAACCCTCGCGCAGTGAGCCATTTTGGTG
 ACGATAGACCTCAGTGGAGAGAAGCGAGCGCTAAAGCACTCGCCACAGTCACGCTGACCCAGCGCGG
 CACGCCTTTTATTTTTCAAGGGGATGAATTAGGGATGACGAATTACCCTTTTAAAACGCTTCAGGATTT
 TGATGACATCGAAGTCAAAGGATTCTTCCAAGACTATGTTGAAACAGGGAAAAGCTACTGCTGAAGAGC
 TTTTGACAAACGTTGCGCTTACGTCAAAGGGATAATGCCCGAACGCCTTTTCAATGGGACGATAGTGCA
 AATGCGGGATTACCACCGGCAAGCCATGGCTGAAAGTTAACCCGAATTACACCGAAATCAATGCCCG
 GCGTGAAATTGGAGATCCCAAGAGCGTCTATAGCTTTTATAGAAACCTGATTTCTATACGTCACGAAA
 CACCGGCGTTGTCAACTGGGTCTTATCGAGACATTGATCCTAGCAACGCCGATGTGTACGCGTATACG
 CGGTCACAGGACGGAGAGACATATTTGGTAGTCGTTAATTTTAAAGGCAGAGCCCCGGTCTTTTACCTT
 ACCGGATGGTATGCATATTGCGGAAACACTGATTGAGTCTCTTCCCGGCGGCTCCGGCAGCCGGCG
 CAGCCTCTCTGGAGCTCCAGCCGTGGCAGAGTGGTATTTACAAGGTTAAATAA

SEQ ID NO: 11

MLEKNHKKISLSGKSLLMGTLSTAAIVLSASTANAATINADNVNENQTVEVTASSVNNENNKQVTEKDS
 ADKSTSDVAEDANTKKSNENTETTEKNTQTVVTNAPVSDVKNTNTVTAETPVDKVVNSDQKTTNAATT
 DTKKDDVKQVEKKDSVDKTNAEENKDSVVKPAENATKAELKGQVKDIVEESGVDTSKLTNDQINELNKN
 FSKEAKSGTQLTYNDFKKIAKTLIEQDARYAIPFFNASKIKNMPAAKTLDAQSGKVEDLEIWDSWPVQDAK
 TGYVSNWNGYQLVIGMMGVPNVNDNHIYLLYNKYGDNDFNHWNAGPIFGLGTPVIQQWSGSATLNKD
 GSIQLYYTKVDTSDNNTNHQKLASATVYLNLEKDQDKISIAHVDNDHIVFEGDGYHYQTYDQWKETNKG
 ADNIA MRDAHVDDDNNGNRYLVFEASTGTENYQDQIYQWLNYYGGTNKDNLDGDFQILSNSDIKDRAK
 WSNAIIGIILNDDVKNPSVAKVYSPLISAPMVSDIEPRPDVVKLGKNKYLF AATRLNRGSND DAWMATN
 KAVGDNVAMIGYVSDNLTHGYVPLNESGVVL TASVPANWRTATYSYYAVPVEGRDDQLLITSYITNRGEV
 AGKGMHATWAPSFLQINPDNTTTLAKMTNQGDWIWDDSSENPDMMGVLEKDAPNSAALPGEWGKPV
 DWDLIGGYNLKPHQPVPTIPNVPTTPETPTTPDKPEVPTTPEVPTTPEPTTPEAPKNPVKKTSSQSKLPKAGDK
 NSFAAVVLGAVSSILGAVLGTGVSKRKRNN

SEQ ID NO: 12

ATGTTAGAAAAACAAGAATCATAAGAAGATTAGCCTTAGTGAAAAATCACTTTTGATGGGAACTTTATC
 AACAGCCGCTATTGTACTGAGCGCCAGCACAGCAAACGCTGCAACAATTAATGCAGATAACGTTAACG
 AAAATCAAACCGTTGAAGTAACGGCATCCTCCGTGAATAACGAAAAATAATAAGCAGGTGACTGAAAA
 AGATAGCGCGGATAAATCAACTTCAGACGTTGCCGAGGACGCCAACACAAAAAATCTAATGAAAAAC
 ACAGAAACGACAGAGAAAAAATACACAACTGTGTAAACAAACGCTCCGGTGAGTGATGTAAGAACA
 CGAATACGGTTACAGCAGAAACACCGGTTGATAAGGTGGTTAACTCTGACCAGAAAACTAA
 TGCAGCGACAACAGATACAAAAAAGACGACGTGAAACAAGTAGAGAAAAAAGATTCCGTGGATAA
 GACCAACGCTGAAGAAAAATAAGGACTCCTCAGTAAAGCCTGCTGAAAACGCTACAAAGGCAGAATTG
 AAAGGGCAGGTAAAAGATATCGTTGAGGAAAGCGGAGTCGACACGAGTAACTTACCAATGATCAAA
 TTAACGAGCTGAACAAAATTAACCTTAGCAAAGAGGCAAAAAAGCGGAACGCAATTGACATATAATGA
 TTTTAAAAAGATCGCCAAAACCTTATTGAACAAGATGCCCGCTATGCCATCCCTTTTTTAAATGCAAG
 TAAAATTAAAAATATGCCAGCAGCTAAAACACTCGACGCACAAAAGCGGCAAAAGTGGAAGATCTTGAA
 ATTTGGGATTCTTGGCCGGTGACAGACGCGAAAACGGGCTACGTGAGTAACTGGAATGGGTACCAGCT
 GGTCATTGGAATGATGGGAGTTCCGAATGTCAACGATAATCATATTTACCTCCTGTATAATAAATATG
 GTGACAACGACTTTAATCATTGGAAAAACGCGGGCCCTATTTTCGGCCTGGGAACACCAGTGATCCAA
 CAATGGTCCGGATCAGCCACACTTAATAAAGACGGATCGATACAGCTTTACTACACTAAGGTAGACAC
 AAGCGATAATAATACAAATCATCAAAACTGGCCAGTGCTACAGTCTACTTAAATTTGAAAAAGATC

AAGACAAGATAAGTATTGCACACGTGGATAATGACCACATCGTGTTTGAAGGAGACGGTTACCACTAT
CAGACATACGATCAATGGAAGGAAACCAATAAAGGCGCAGACAATATCGCAATGCGCGATGCACACG
TCATAGACGACGATAACGGGAATCGCTATCTCGTGTTTGAAGCGAGTACCGGAACAGAAAACCTATCA
GGGTGATGACCAGATTTATCAATGGTTGAATTATGGAGGCACTAACAAAGATAACCTTGGTGACTTTT
TTCAAATCTTAAGCAACTCAGATATCAAAGATCGCGCAAAATGGTCCAATGCAGCCATCGGCATCATT
AAATTAATGATGATGTCAAGAATCCTTCTGTGCGCAAGGTCTATTACCGTTGATTTACGGCCTATG
GTATCTGATGAGATCGAACGGCCGGATGTCGTGAAGTTAGGAAATAAATATTATCTGTTTGTGCCAC
GCGGTTAAACAGAGGCAGCAACGATGATGCTTGGATGGCAACAAACAAAGCAGTGGGGGACAATGTG
GCAATGATTGGGTATGTATCTGACAACCTTACCCATGGCTATGTACCGCTTAATGAATCTGGAGTAGTC
CTTACCGCGTCAGTTCCTGCGAATTGGCGCACGGCCACCTACTCTTATTATGCGGTCCCTGTGCAAGGG
CGCGATGATCAATTGCTGATCACATCTTATATTACCAATAGAGGTGAGGTTGCCGGTAAAGGAATGCA
TGCCACGTGGGCACCCAGCTTTTGTCTCAGATCAACCCGGACAATACGACCACAGTTCTGGCAAAGA
TGACGAACCAGGGGGACTGGATATGGGACGATAGCAGTGAACCCCTGACATGATGGGCGTATTAGA
AAAAGACGCCCCTAATTCAGCAGCACTCCCGGGCGAGTGGGGCAAACCAGTTGACTGGGATCTGATTG
GTGGGTACAACCTCAAACCACATCAACCTGTCACACCTATACCGAACGTGCCAACTACGCCGGAAACA
CCTACAACGCCGTGATAAACCAGAAGTACCAACTACCCAGAAAGTCCCGACAACACCGGAAACACCGA
CCCCGGAAGCCCCATAAAACCCTGTCAAAAAGACGTACAGTCGAAACTTCCTAAAGCGGGTGATAA
GAATTCCTTTGCCGCGGTTGTTTATAGGGCAGTCTCAAGCATTCTGGGCGCCGTGGGCCTCACGGGCGT
TTCTAAAAGAAAACGTAACAATTAA

SEQ ID NO: 13

MAPNLSKAKDDLIGDVVAVDGLIKPPRFTLK GKDLAVDGHFPFLDVPANIRLTPASTLVPNSDVPAAAAGS
FLGFDAPAKDRHVVPPIGLRDRFMSIFRFK VWWTTHWVG TNGRDVENETQMMLDQSGTKSSPTGPRP
YVLLLPIVEGKCAEDYVHVMVLESSTVRGSVFRSAVYLHAGDDPDFL VKDAMRVVRAHLGT
FRLMEEKTPPPIVDKFGWCTWDAFYLK VHPEGVWEGVRRLLADGGCPPGLVLIDDGWQSICHDDDLGSG
AEGMNRTSAGEQMPCLIKFQENYKFREYKGGMGGFVREMKAAPFTVEQVYVWHALCGYWGGLRPGAP
GLPPAKVVAPRLSPGLQRTMEDLAVDKIVNNGVGLVDPRRARELYEGLHSHLQASGIDGVKVDVIHLLM
VCEEYGGRVELAKAYFAGLTESVRRHFNGNGVIASMEHCNDFMLLGTEAVALGRVGDDFWCTDPSGDPD
GTFWLQGCHMVHCAYNLSLWMGAFIHPDWMFQSTHPCAAFHAASRAVSGGPVYVSDAVGCHDFDLRR
LALPDGTLRCERYALPTRDCLFADPLHDGKTM LKIWNVNKFSGVLGAFNCQGGWSREARRNMCAAGF
SVPVTARASPADVEWSHGGGGDRFAVYFVEARKLQLLRDESVELTLEPFTYELLVVAPVRAIVSPELGI
GFAPIGLANMLNAGGAVQGF EAARKDGDVA AEVAVK GAGEMVAYSSARPR LCKVNGQDAEFKYEDGIV
TVDVPWTGSSKKLSRVEYFY

SEQ ID NO: 14

ATGGCTCCGAATCTGAGTAAAGCAAAGGACGATTTAATTGGAGACGTCGTCGAGTGATGGCCTTAT
CAAACCTCCTCGGTTTACACTCAAAGGAAAAGACCTCGCCGTTGATGGTCACCCGTTTCTGCTGGACGT
GCCGGCGAATATCCGGTTAACACCAGCATCAACTCTTGTCCTCCAACTCGGATGTTCCGGCTGCAGCTG
CTGGGAGCTTTCTTGCTTCGATGCACCCGCGGCAAAAGATCGTCATGTCGTTCCGATTGGGAAACTG
CGCGATACAAGATTTATGTCTATCTTTCGTTTTAAAGTCTGGTGGACGACACATTGGGTTGGGACAAAT
GGACGGGATGTAGAGAACGAGACCCAGATGATGATCTTGATCAGTCCGGAACAAAGTCTAGCCCGA
CTGGACCGCGCCCGTATGTTCTGCTTCTTCCGATCGTCGAAGGCCCGTTTAGAGCCTGTTTAGAATCGG
GGAAAGCCGAAGATTATGTCCATATGGTGTCTGAAAGTGGGTCTAGCACCGTACGTGGTTTCACTGTTT
AGAAGCGCTGTGTATCTTCATGCTGGCGATGACCCCTTTGATTTGGTGAAAGATGCAATGCGCGCTCGTA
CGTGCGCATTTAGGAACGTTCCGGTTGATGGAGGAAAAGACACCGCCTCCGATCGTTGATAAGTTTGG
CTGGTGACCTGGGATGCCTTTTATTTAAAGGTACATCCAGAAGGAGTGTGGGAAGGAGTACGGAGAT
TAGCAGATGGCGGCTGCCACCGGGGCTGGTATTAATTGACGATGGATGGCAATCGATTTGCCATGAT
GATGACGATCTTGGCTCAGGAGCTGAAGGAATGAACAGAACAAAGCGCGGGCGAGCAAATGCCTTGTA
GGCTTATTAAATTTAGGAAAACTATAAATTCAGGGAATACAAAGGTGGCATGGGAGGCTTCGTGCGT
GAGATGAAAGCAGCATTTCCGACCGTGGAGCAAGTCTATGTTTGGCACGCACTGTGCGGATATTGGGG
AGGTCTCCGTCCGGGTGCACCGGGTCTGCCGCCAGCAAAAGTGGTAGCACCGCGATTGTCCCCGGGCT
TACAGCGGACGATGGAAGATCTTGGGTAGATAAAATTGTTAACAATGGCGTGGGCCTTGTGGATCCT
CGCAGAGCTAGAGAACTGTATGAAGGCCTTCATCCACCTCCAGGCCAGCGGTATAGATGGAGTGAA
AGTAGATGTTATACACCTTCTTGAGATGGTTTTCGAGGAATACGGGGGGAGAGTAGAATTGGCTAAAG
CGTATTTTCGAGGACTCACTGAGTCAGTTCGGAGGCATTTTAAACGGAAATGGTGTCAATTCGCTCAATG

GAGCACTGCAACGATTTTATGCTGCTGGGTACAGAAGCTGTAGCATTAGGCCGGGTGGGAGATGATTT
 CTGGTGTACAGATCCGTCTGGTGACCCAGATGGTACGTTCTGGTTACAGGGATGCCACATGGTACACT
 GTGCGTATAATAGTTTATGGATGGGCGCCTTTATACATCCGGACTGGGACATGTTTCAATCTACACACC
 CCTGCGCGCGCTTCCATGCTGCGAGCAGAGCCGCTCTCCGGTGGACCAGTCTATGTATCAGATGCGGTT
 GGCTGCCATGATTTTCGATCTGCTTCGTAAGGCTCGCACTGCCGGATGGCACAAATTTGCGTTGCGAACGC
 TATGCACTGCCGACAAGAGATTGTCTTTTTCGGGATCCCTTGCATGATGGCAAAACAATGTTGAAAAT
 ATGGAACGTAAACAAATTTTCAGGAGTTTTAGGAGCATTTAATTGCCAGGGCGGAGGATGGTCACGCG
 AAGCTCGTCGCAACATGTGCGCGGCTGGTTTTAGTGTTCCGGTGA CTGCTCGGGCTTACCTGCAGACG
 TTGAGTGGTCACACGGAGGCGGTGGTGGAGATCGTTTTGCCGTTTATTTTCGTGGAGGCTCGCAAACTTC
 AATTACTGCGTCGCGACGAAAGCGTGGAAGTACATTAGAACCGTTTACATATGAGCTGTTAGTCGTG
 GCCCCGTGAGAGCTATCGTGTCCAGAGAATTAGGTATCGGCTTCGCACCGATTGGCCTCGCTAATAT
 GCTGAACGCGCGGGCGGGTTCAGGGTTTTGAAGCCGCCGTAAGGATGGAGATGTGGCTGCTGAG
 GTGGCTGTCAAAGGGCGGGTGAAATGGTCGCATACTCTAGTGCGAGACCAAGGTTATGCAAGGTAA
 ACGGACAGGACGCCGAATTTAAATATGAAGACGGAATTGTCACAGTAGACGTCCTTGGACAGGATC
 CTCAAAAAACTTAGCAGGGTAGAATATTTCTATTAG

SEQ ID NO: 15

MAHVRRKVATLNMALAGSLLMVLGAQSALAQGNFSRQEAARMHRPGVMPRGGPLFPGRSLAGVPGFP
 LPSIHTQQAYDPQSDFTARWTRADALQIKAHSDATVAAGQNSLPAQLTMPNIPADFPVINPDVWVWDTWT
 LIDKHADQFSYNGWEVIFCLTADPNAGYGDDRHHVARIGFFYRRAGIPASRRPVNGGWYTGHLFPDGA
 SAQVYAGQTYTNQAEWSGSSRLMQIHGNTVSIFYTDVAFNRDANANNITPPQAIITQLGRIHADFNHVV
 FTGFTAHTPLLQPDGVLYQNGAQNEFFNFRDPFTFEDPKHPGVNYMVFEGNTAGQRGVANCTEADLGFRP
 NDPNAETLQEVLDGAYYQKANIGLAIAITDSTLSKWKFLSPLISANCVNDQTERPQVYLHNGKYYIFTISHR
 TTFAAGVDGPDGVYGFVGDGIRSDFPMPNYGSGLTMGNPDTLNTAAGTDFDPSDPQNPRAFQSYSHYVM
 PGGLVESFIDTVENRRGGTLAPTVRVRIAQNASAVDLRYGNGGLGGYGDIPANRADVNIAFGIQLDFGQPT
 SGLAAQASTNNAQVLAQVRQFLNQ

SEQ ID NO: 16

MLNKAGIAEPSLWTRADAMKVHTDDPTATMPTIDYDFPVMTDKYWVWDTWPLRDINGQVVSFQGWSVI
 FALVADRTKYGWHNRNDGARIGYFYSRGGSNWIFGGHLLKDGANPRSEWSGCTIMAPGTANSVEVFT
 SVNDTPSESVPAQCKGYIYADDKSVWFDGDKVTDLFQADGLYYADYAENNFWDPRDPHFVFINPEDGKT
 YALFEGNVAMERGTVAVGEEIGPVPPKTEPTDGARYCAAIGIAQALNEARTEWKLLPPLVTAFGVNDQ
 TERPHVVFQNGLTYLFTISHHSTYADGLSGPDGVYGFVSENGIFGPYEPLNGSGLVLGNPSSQPYQAYSHYV
 MTNGLVTSFIDTIPSSDPNVYRYGGTLAPTIKLELVGHRFSVTEVKGYGYIPPQIEWLAEDESSNSAAALSLL
 NK

SEQ ID NO: 17

MNKLKIVKCILIGSMICSGIITQQTFASTNDMNYKETYGVSHITRYNMSKIPMEQNDLKFKVPQFNASTLKN
 ASAKGYDKNGNLIDLVDWSWPLQNGDGTVANYHGYHIVFALAGDPKNQDDTSIYMFYQKIGENSIDSW
 KNAGKVFKDSDKYVANDPYLKYQTQEWSGSATLTSQGQVRLFYTDFSGVAKDGGTDASNQVITTTQVNL
 SQPDSNTINIDSVDHKSVDGNGTIYQNVQQFIDEKGWSSGDNHTRLRDPHYVEDNGRKYLVFEANTGTN
 DGYQGDTSLLNKAFYGRSQSFFKTEKDQLLIDTNKKHDASLANGALGIELNNDYTLKKEMKPLIASNTVT
 DEIERANVFKMNGRWYLFDTSRGSKMTINGISSKDIYMLGFSNSLTGPYKPLNKTGLVLNLDPTDLTFT
 YSHFAPQTNKGNNVITSYITNRGMYSDDHSSFAPSFLNLIKGTSTVISNSILQQQLTIDNY

SEQ ID NO: 18

MNIKKFAKQATVLTFTTALLAGGATQAFAKETNQKPYKETYGISHITRHDMLQIPEQQKNEKYQVPEFDSS
 TIKNISSAKGLDVWDSWPLQNDGTVANYHGYHIVFALAGDPKNADDTSIYMFYQKVGETSIDSWKNAG
 RVFKDSDKFDANDSILKDQTQEWSGSATFTSDGKIRLFYTDGSGKHYGKQTLTTAQVNVASDSSLNINGV
 EDYKSIFDGDGKTYQNVQQFIDEGNYSSGDNHTRLRDPHYVEDKGHKYLVFEANTGTEDGYQGEESLNFKA
 YYGKSTSFFRQESQKLLQSDKKRTAELANGALGMIELNDDYTLKKVMKPLIASNTVTDEIERANVFKMNG
 KWYLFDTSRGSKMTIDGITSNDIYMLGYVSNLSLTGPYKPLNKTGLVLKMDLDPNDVTFTYSHFAPQAKG
 NNVVITSYMTNRGFYADKQSTFAPSFLNLIKGGKTSVVKDSILEQQQLTVNK

SEQ ID NO: 19

ATGAACATAAAGAAGTTTGC GAAGCAGGCGACAGTATTAACGTTACCCACGGCACTCTTGGCTGGGGG
CGCAACCCAGGCTTTTGCTAAAGAGACCAACCAGAAGCCGTATAAGGAAACGTATGGGATTTCCACACA
TTACAAGACATGATATGCTGCAGATCCCAGAACACAGAAAAACGAAAAATACCAAGTCCCGGAATT
TGATTCTTCGACCATTAAAAACATTTCTCAGCTAAAGGGCTGGACGTATGGGATTTCTGGCCCCCTTCA
GAATGCAGATGGAACGTGTGCCAATTACCATGGCTATCATATTGTCTTTGCGCTGGCGGGGGATCCGA
AAAACGCAGATGATACGTCCATCTACATGTTTTACCAAAAAGTAGGGGAAACAAGCATTGATTCTCTGG
AAAAACGCTGGCCGCGTCTTCAAGGATTCAGACAAATTTGATGCGAATGATAGTATACTGAAAGATCA
GACGCAAGAGTGGAGCGGTTCCGCAACTTTCACGAGCGACGGAAAAATCCGTCTGTCTACACCGATT
TTAGCGGCAAGCATTACGGGAAGCAGACTCTGACGACGGCCCAAGTCAATGTATCTGCATCGGACTCT
TCCCTGAATATTAACGGAGTTGAAGATTACAAATCTATATTGACGGTGATGGCAAAACCTACCAAAA
CGTCCAGCAGTTCATAGACGAAGGCAATTATAGTTTCAGGGGACAATCATACTTTACGTGACCCACATT
ATGTCGAAGATAAAGGACATAAATATCTGGTTTTTGAGGCGAACACTGGCACAGAGGACGGGTATCA
GGGGGAAGAGTCCCTGTTCAATAAAGCTTATTACGGCAAGTCAACATCTTTTTTCGCCAAGAATCAC
AAAAGCTGCTGCAATCAGATAAGAAAAAGGACTGCTGAGTTAGCGAACGGCGCATTAGGCATGATTGA
ACTGAATGACGACTACACACTTAAAAAAGTTATGAAACCATTAAATTGCGTCGAACACCGTTACTGATG
AAATTGAAAGGGCGAATGTTTTTAAGATGAACGGGAAATGGTATCTCTTTACGGATAGTCGCGGATCA
AAGATGACCATCGATGGTATCACGTCCAATGACATTTACATGCTTGGATATGTTAGCAATTCCTTAACT
GGACCTTATAAACCGCTTAACAAAAACAGGCTTGGTACTGAAAATGGATTTGGATCCAAATGACGTCAC
ATTTACATACTCTCATTTGCTGTTCCGCGAGGCAAAAGGGCAACAATGTTGTGATTACATCTTATATGAC
CAACCGCGGCTTTTATGCGGATAAGCAATCAACCTTCGCACCATCATTCTGCTTAACATCAAGGGCA
AAAAACGAGCGTAGTCAAAGATTCTATCTTGAACAGGGTCAGCTGACGGTCAACAAATAA

SEQ ID NO: 20

ATGAACATCAAAAAGTTTGC AAAACAAGCAACAGTATTAACCTTTACTACCGCACTGCTGGCAGGAGG
CGCAACTCAAGCGTTTGCGAAAGAAACGAACCAAAAAGCCATATAAGGAAACATACGGCATTTCCTCAT
ATTACACGCCATGATATGCTGCAAATCCCTGAACAGCAAAAAATGAAAAATATCAAGTTCCTGAATT
CGATTCTGTCACAATTA AAAATATCTCTTCTGCAAAAAGGCTGGACGTTTGGGACAGCTGGCCATTAC
AAAACGCTGACGGCACTGTGCAAACTATCACGGCTACCACATCGTCTTTGCATTAGCCGGAGATCCT
AAAAATGCGGATGACACATCGATTTACATGTTCTATAAAAAGTCGGCGAAACTTCTATTGACAGCTG
GAAAAACGCTGGCCGCGTCTTTAAAGACAGCGACAAATTCGATGCAAAATGATTCTATCTTAAAGACC
AAACACAAGAATGGTCAGGTTACGCCACATTTACATCTGACGGAAAAATCCGTTTATTCTACACTGAT
TTCTCCGGTAAACATTACGGCAAAACAAACACTGACAAGTGAACGTTAACGTATCAGCATCAGACAG
CTCTTTGAACATCAACGGTG TAGAGGATTATAAATCAATCTTTGACGGTGACGGAAAAACGTATCAAAA
ATGTACAGCAGTTTCATCGATGAAGGCAACTACAGCTCAGGCGACAACCATACGCTGAGAGATCCTCAC
TACGTAGAAGATAAAGGCCACAAATACTTAGTATTTGAAGCAAAACACTGGAAGTGAAGATGGCTACC
AAGGCGAAGAATCTTTATTAAACAAAGCATACTATGGCAAAAAGCACATCATTCTTCCGTCAAGAAAGT
CAAAAACCTTCTGCAAAGCGATAAAAAACGCACGGCTGAGTTAGCAAACGGCGCTCTCGGTATGATTG
AGCTAAACGATGATTACACACTGAAAAAAGTGATGAAACCGCTGATTGCATCTAACACAGTAACAGA
TGAAATTGAACGCGCGAACGTCTTTAAATGAACGGCAAAATGGTACCTGTTCACTGACTCCCGCGGAT
CAAAAATGACGATTGACGGCATTACGTCTAACGATATTTACATGCTTGGTTATGTTTCTAATTCTTTAA
CTGGCCCATAAAGCCGCTGAACAAAACCTGGCCTTGTTGTTAAAAATGGATCTTGATCCTAACGATGTA
ACCTTTACTTTACTCACACTTCGCTGTACCTCAAGCGAAAGGAAACAATGTCGTGATTACAAGCTATATG
ACAAACAGAGGATTCTACGCAGACAAAACAATCAACGTTTGCGCCAAGCTTCCTGCTGAACATCAAAAG
CAAGAAAACATCTGTTGTCAAAGACAGCATCCTTGAACAAGGACAATTAACAGTTAACAATAA

SEQ ID NO: 21

MNIKKFAKQATVLTFTTALLAGGATQAFAKETNQKPYKETYGISHITRHDMLQIPEQQKNEKYQVPEFDSS
TIKNISSAKGLDVWDSWPLQNADGTVANYHGYHIVFALAGDPKNADDTSIYMFYQKVGETSIDSWKNAG
RVFKDSDKFDANDSILKDQTQEWSGSATFTSDGKIRLFYTDGSGKHGKQTLTTAQVNVSASDSSLNINGV
EDYKSIFDGDGKTYQNVQQFIDEGNYSSGDNHTLRDPHYVEDKGHKYLVFEANTGTEDGYQGEESLFNKA
YYGKSTSFFRQESQKLLQSDKKRTAELANGALGMIELNDDYTLKKVMKPLIASNTVTDEIERANVFKMNG
KWYLFDTDSRGSKMTIDGITSNDIYMLGYVNSLTGPYKPLNKTGLVLKMDLDPNDVFTTYSHFAVPQAKG
NNVVITSYMTNRGFYADKQSTFAPSFLNLIKGGKTSVVKDSILEQQQLTVNK

SEQ ID NO: 22

ATGAGAAATTAACCACAATATTGCAGCGCTTAACACACTGAACCGTTTGTCTTCAAACAACAGTGCGAG
CCAAAAGAACATGGAGAAACTTTCTTCAGGTCTTCGCATCAACCGTGCGGGAGATGACGCAGCAGGTC
TTGCGATCTCTGAAAAAATGAGAGGACAAATCAGAGGTCTTGAAATGGCTTCTAAAACTCTCAAGAC
GGAATCTCTCTTATCCAAACAGCTGAGGGTGCATTAACCTGAACTCATGCGATCCTTCAACGTGTTCTGT
GAGCTAGTTGTTCAAAGCTGGAAACACTGGAATCAGGACAAAAGCAACTGATTTGCAATCTATTCAAGA
TGAAATTTACAGCTTTAACAGATGAAATCGATGGTATTTCAAATCGTACAGAATTCATGGTAAGAAAT
TGCTCGATGGCACTTACAAAGTTGACACAGCTACTCCTGCAAATCAAAGAAGCTTGGTATTCCAAATC
GGAGCAAATGCTACACAGCAAATCTCTGTAAATATTGAGGATATGGGTGCTGACGCTCTTGGAATTAA
AGAAGCTGATGGTTCAATTGCAGCTCTTCATTAGTTAATGATCTTGACGTAACAAAATTCGCAGATA
ATGCAGCAGATACTGCTGATATCGGTTTCGATGCTCAATTGAAAGTTGTTGATGAAGCGATCAACCAA
GTTTCTTCTCAACGTGCTAAGCTTGGTGCGGTACAAAATCGTCTAGAGCACACAATTAACAACCTTAAG
CGCTTCTGGTGAAAACTTGACAGCTGCTGAGTCTCGTATCCGTGACGTTGACATGGCTAAAGAGATGA
GCGAATTCACAAAGAACAACATTCTTCTCAGGCTTCTCAAGCTATGCTTGCTCAAGCAAACCAACAG
CCGCAAAACGTACTTCAATTATTACGTTAA

SEQ ID NO: 23

GGAATTGACGCCCCAAAGCATATTGATATTCACAGGAAAGAAATTTACTTGACCATTTCAGGAAGAAAA
TAACCGTGACGAGCGTTATCCAGCGATGTGATCTCCGCATTATCCTCACAAAAAAGTGAGGATTTT
TTTATTTTTGTATTAACAAAATCAGAGACAATCCGATATTAATGATGTAGCCGGGAGGAGGCGCAAAA
GACTCAGCCAGTTACAAAATAAGGGCACAAAGGACGTGCCTTAACAACATATTCAGGGAGGAACAAAA
CA

SEQ ID NO: 24

GCACAAGGACGT

SEQ ID NO: 25

ATTCAGGGAGGAA

SEQ ID NO: 26

AGGGAGGA

SEQ ID NO: 27

ATGCTAGTTTATCGCGGAAAAATAACGAAGCGATTCAAATAGGTGCTGATATTGAAGTAAAAAGTGAT
TGCGGTTGAAGGGGATCAAGTGAAGCTTGGAATTGACGCCCCAAAGCATATTGATATTCACAGGAAA
GAAATTTACTTGACCATTTCAGGAAGAAAAATAACCGTGACGAGCGTTATCCAGCGATGTGATCTCCGC
ATTATCCTCACAAAAAAGTGA

SEQ ID NO: 28

TAAGGGCACAAAGGACGTGCCTTA

SEQ ID NO: 29

GCACAAGAACGT

SEQ ID NO: 30

ATTTAGGGAGGAA

SEQ ID NO: 31

ATGAAAATCAATCAATTTGGAACACAATCCGTTAATCCATATCAAAAAAATTATGATAAGCAAGCGGT
GCAAAAAACTGTTGCACAACCTCAAGATAAAATTGAAATTTATCACAGGCTAAAGAAATGCAACAT
GCATCCGACGCAGTCACTGGTTCACGACAGGAAAAAATTGCGCAGCTTAAAGCGCAAATTGAAAACG
GGTCATACAAAGTAGACGCAAATCATATTGCGAAAAATATGATTAATTTTTATAAAAAGCAATAA

SEQ ID NO: 32

ATGCAATCCTTGAATTATGAAGATCAGGTGCTTTGGACGCGCTGGAAAGAGTGGAAAGATCCTAAAGC
CGGTGACGACTTAATGCGCCGTTACATGCCGCTTGTCACATATCATGTAGGCAGAAATTTCTGTCGGACT
GCCGAAATCAGTGCATAAAGACGATCTTATGAGCCTTGGTATGCTTGGTTTATATGATGCCCTTAAAA
AATTTGACCCCGACCCGGGACTTAAATTTGATACCTACGCCTCGTTTGAATTCGCGGCGCAATCATA
GACGGGCTTCGTAAAGAAGATTGGCTGCCAGAACCTCGCGCGAAAAACAAAAAGGTTGAAGCAG
CAATTGAAAAGCTTGAACAGCGGTATCTTCGGAATGTATCGCCCGCGGAAATTGCAGAGGAACTCGGA
ATGACGGTACAGGATGTCGTGTCAACAATGAATGAAGGTTTTTTTTGCAAATCTGCTGTCAATTGATGA
AAAGCTCCATGATCAAGATGACGGGGAAAAACATTCAAGTCATGATCAGAGATGACAAAAATGTTCCG
CCTGAAGAAAAGATTATGAAGGATGAACTGATTGCACAGCTTGGCGAAAAAATTCACGAACTCTCTGA
AAAAGAACAGCTGGTTGTCAGTTTGTCTACAAAGAGGAGTTGACACTGACAGAAATCGGACAAAGTAT
TAAATCTTTCTACGTCCCGCATATCTCAGATCCATTCAAAGGCATTATTTAAATTAAGAATCTGCTGG
AAAAAGTGATACAATAA

SEQ ID NO: 33

MSRLVVVSNRIAPPDEHAASAGGLAVGILGALKAAAGGLWFGWSGETGNEDQPLKKVKKGNITWASFNLS
EQDLDEYYNQFSNAVLWPAFHYRLDLVQFQRPWDGYLRVNALLADKLLPLLQDDDIWIHDYHLLPFAH
ELRRKRGVNNRIGFFLHIPFPTPEIFNALPTYDTLLEQLCDYDLLGFQTENDRLAFLDCLSNLTRVTTTSAKSH
TAWGKAFRTEVYPIGIEPKEIAKQAAGPLPPKLAQLKAEKKNVQNIFSVRLDYSKGLPERFLAYEALLEKY
PQHHGKIRYTIAPTSRGDVQAYQDIRHQLENEAGRINKYQQLGWTPLYYLNQHFDRKLLMKIFRYSDV
GLVTPLRDGMNLVAKEYVAAQDPANPGVLVLSQFAGAANELTSALIVNPYDRDEVAAALDRALTMSLAE
RISRHAEMLDVIVKNDINHWQECFISDLKQIVPRSAESQQRDKVATFPKLA

SEQ ID NO: 34

ATGTCCCGTTTAGTAGTGGTGTCCAATCGTATCGCGCCGCCGCGACGAACATGCTGCCTCAGCTGGCGG
CCTGGCCGTAGGAATCCTGGGCGCCCTCAAGGCAGCTGGAGGATTATGGTTCCGGCTGGTCTGGCGAGA
CAGGAAATGAGGATCAACCACTTAAGAAAGTGAAAAAAGGCAATATCACATGGGCTTCCTTCAACCT
CAGCGAGCAAGACCTGGATGAGTATTATAACCAAGTTTAGCAATGCTGTGCTTTGGCCGGCTTTTCATTA
CAGGTTAGATTTGGTTTCAGTTTCAAAGACCCGCATGGGATGGATACCTCCGAGTGAATGCGTTGTTGG
CAGATAAACTTCTCCCGCTCCTTCAGGATGATGATATTATCTGGATTGACTACCATCTTCTCCCTTT
CGCCACGAATTGCGCAAACGGGGCGTGAACAATAGGATAGGTTTTTTTTGCACATTCTTTTCCAC
ACCGGAAATTTTCAACGCGCTTCCGACATACGATACTTTACTTGAACAGCTGTGTGATTACGATCTTCT
CGGCTTCCAACTGAAAATGACAGACTTGCCTTTTTGGATTGCCTCTCAAATTTGACGAGGGTTACGAC
TAGAAGTGCCAAGAGCCATACAGCGTGGGGAAAAGCATTACAGGACAGAAGTTTATCCTATTGGGATC
GAGCCTAAGGAAATTGCGAAACAAGCGGCAGGGCCATTACCTCCGAAACTTGGCGAACTCAAAGCGG
AATTAAGAACGTACAAAACATTTTTAGCGTCGAAAGACTTGATTATTCTAAGGGTCTCCCGGAAAGA
TTCTTAGCCTACGAGGCATTGCTTGAAAAATATCCACAGCATCATGGGAAAATTCTGTTATACGCAAAT
CGCTCCGACTAGCAGAGGCGACGTCCAAGCGTATCAGGACATACGCCACCAACTTGAAAAATGAAGCG
GGTAGAATCAATGGCAAATATGGACAACTGGGGTGGACACCTCTTTATTACTTGAATCAACATTTTCA
TAGAAAATTGTTGATGAAAATCTTTCGTTATTCTGACGTCGGACTGGTGACACCGCTGAGAGATGGCA
TGAACCTAGTTGCCAAGGAATATGTAGCTGCGCAAGACCCTGCTAATCCGGGAGTACTGGTGCTCTCA
CAATTTGCAGGGGCCGCAATGAACTTACATCAGCTCTCATCGTTAATCCGTATGACAGGGATGAAGT
CGCAGCGGCGCTTGACCGGGCTTTACAATGCTTACGCGGAGAGAATTAGCAGACATGCTGAAATGC
TGGATGTGATTGTGAAGAACGATATCAATCATTTGGCAAGAATGTTTCATTTCTGACTTAAAGCAAAT
GTCCCGCGTTACAGCTGAGTCACAGCAACGGGATAAAGTCGCCACATTTCCCAAATGGCATAA

SEQ ID NO: 35

MDIVFAADDNYAAYLCVAAKSVEAAHPDTEIRFHVLDAGISEANRAAVAANLRGGGNIRFIDVNPEDFAG
FPLNIRHISITTYARLKLGEYIADCDKVLVLDIDVLVRDSLKPLWDTDLGDNWLGACIDLFVERQNA YKQKI
GMADGEYYFNAGVLLINLKKWRQHDIFKMACEWVEQYKDV MQYQDQDILNGLFKGGVCYANSRNFMP
TNDAFMANRFASTRHTDPLYRDRYTTAMPVAVSHYCGPAKPWHRDCTAWGAERFTELAGSLTSVPEEWRG
KLAVPHRVFPTKRMLQRWRRKLSARFLRKIY

SEQ ID NO: 36

ATGGATATTGTATTTGCGGCAGATGATAATTATGCTGCGTATCTCTGTGTTGCCGCCAAATCCGTGGAA
GCCGCTCACCCCGATACAGAGATCCGGTTCATGTACTTGACGCTGGCATTTCGGAAGCAAACCGGGC
AGCAGTTGCAGCCAATCTCCGCGGTGGCGGTAATATTCGTTTTATAGATGTGAACCCGGAAGACTTTG
CCGGCTTTCCATTGAATATTCGCCATATTAGTATAACAACCTTACGCACGTCTGAAACTTGGGGAGTACA
TTGCGGATTGTGACAAAGTACTTTATCTCGATATAGATGTACTTGTGCGTGATAGTCTTAAACCGTTAT
GGGATACAGATCTTGGAGATAATTGGCTTGGTGCATGCATCGACTTATTTGTAGAGAGACAAAACGCT
TATAAGCAAAAAATTGGCATGGCTGACGGTGAATATTACTTTAACGCAGGGGTGCTCCTGATTAACCT
TAAGAAGTGGCGGCAACACGATATCTTCAAATGGCTTGCGAATGGGTTGAGCAGTATAAGGATGTA
ATGCAGTACCAGGACCAAGACATCCTGAATGGATTATTTAAAGGTGGAGTGTGTTACGCTAATAGCAG
ATTCATTTTCATGCCGACCAATGATGCTTTTCATGGCAAACCGGTTTGCTTCACGCCATACGGACCCATT
GTATAGAGATAGAACGTATACAGCAATGCCTGTGGCCGTTTCGCATTATTGTGGTCCGGCGAAACCGT
GGCATCGCATTTGCACAGCATGGGGCGCAGAAAGATTTACAGAACTTGCAGGAAGTCTGACATCAGTT
CCGGAGGAATGGCGCGGAAAACCTTGCAGTGCCTCATCGGGTGTTCGGGACTAAACGTATGCTGCAAG
ATGGAGACGCAAACTCTCCGCTCGATTCTGAGAAAAATCTATTGA

SEQ ID NO: 37

MIIVHLCGGLGNQMFQYAAGLAAHRIGSEVKFDTHWFDATCLHQGLELRRVFGLELPEPSSKDLRKVLG
ACVHPAVRRLLAGHFLHGLRPKSLVIQPHFYWTGFEHLPDENVYLEGYWQSERYFSNIADIIRQQFRFVEPL
DPHNAALMDEMQSGVSVSLHIRRGDYFNNPQMRRVHGVLDSEYYPAAVATMIEKTNAERFYVFSDDPQW
VLEHLKLPVSYTVVDHNRGAASYRDMQLMSACRHIIANSTFSWWGAWLNPRPDKVVIAPRHWFNVDVF
DTRDLYCPGWIVL

SEQ ID NO: 38

ATGATTATTGTGCACCTTTGCGGTGGTTTAGGAAACCAGATGTTCCAATATGCCGCAGGCCTTGCCGCG
GCGCACAGAATTGGCAGCGAAGTTAAATTCGATACGCATTGGTTCGATGCGACATGTCTGCATCAGGG
TTTAGAGTTGCGCAGAGTTTTGGTTTGGAGCTTCCGGAACCCAGTTCCAAGGACCTGAGAAAAGTTTT
AGGCGCTTGCGTTCATCCTGCAGTGAGACGGCTTTGGCTGGTCACTTTCTGCATGGCTTACGGCCGAA
GTCGCTCGTTATCCAACCCATTTTCACTACTGGACGGGCTTTGAACATTTACCGGACAACGTATACTT
AGAGGGCTACTGGCAGTCTGAAAGATACTTTTCAAATATTGCTGATATTATAAGACAACAGTTTCGTTT
TGTCGAACCGTTAGATCCGCATAATGCGGCCCTCATGGATGAAATGCAGTCCGGGGTTAGTGTTCAC
TTCATATTGCGCGTGAGACTACTTCAACAACCCGCAAAATGCGTCGTGTCCACGGCGTAGACTTGAGC
GAGTACTACCCGGCGGCAGTCGCGACGATGATCGAAAAACCAACGCAGAACGGTTTTACGTGTTTTTC
AGATGACCCACAGTGGGTATTGGAACATCTGAAACTGCCTGTTTCTTACACCGTGGTGGACCATAACC
GAGGCGCTGCCTCGTATAGGGATATGCAACTTATGTCTGCTTGCCGACACCATATCATTGCGAATTCAA
CATTAGCTGGTGGGGAGCTTGGCTTAACCTCGGCCGACAAGGTCGTGATCGCTCCACGCCATTGG
TTTAACGTAGATGTATTCGACACACGCGATTATATTGCCAGGGTGGATAGTACTGTAA

SEQ ID NO: 39

MNVLSSICYGGDYNPEQWPPEIWIYEDAKLMQKAGVNLVSLGIFSWSKIEPSDGVDFEFLDKVIDILYDHG
VYINLGTATATTPAFVFKKYPDSLPIDESGVLSFGSRQHYCPNHPQLITHIKRLVRAIAERYKNHPALKMW
HVNNEYACHVSKFCENCVAFRKWLKERYKTIDELNERWGTNFWGQRYNHWDEINPPRKAPTFINPSQE
LDYYRFMNDLSILKFLTEKEILREVTDPDIPVSTNFMGSKPLNYFQWAQHVDIVTWDSYDPREGPLIQHAM
MNDLMRSLRKGPFFILMEQVTSNVNWRDINVPKPPGVMRLWSYATARGADGIMFFQWRQSRAGAEKFH
GAMVPHFLNENNRIYREVTQLGQELKKLDCLVGSRIKAEVAIIFDWWENWWAVELSSKPHNKLRYIPIVEAY
YRELYKRNIAVDVFRPSDDLTKYKVVIAPMLYMVKEGEDENLRQFVANGGTLIVSFFSGIVDENDRVHLG
GYPGPLRDILGIFVEEFVPYPETKVNKIYSNDGEYDCTTWADIIRLEGAELATFKGDWYAGLPVTRNCYG
KGEIYVGTYPDSNYLGRLLLEQVFAKHHPINPIEVAENVEVQQRETDEWKYLIINHNDEYVTLSPEDKIY
QNMIDGKCFRGGELRIQGVDAVLREHDEAGKV

SEQ ID NO: 40

ATGAACGTTTTGTCGTCTATTTGCTACGGTGGGGATTATAACCCGGAGCAATGGCCGGAGGAGATCTG
GTATGAAGATGCTAAATTAATGCAAAAAGCGGGGGTCAATTTAGTGAGCTTAGGCATTTTTTCTTGGT
CTAAAATCGAACCGTCTGATGGGGTTTTTGACTTTGAATGGTTAGATAAAGTCATTGATATCCTTTACG
ATCACGGCGTGTATTAATCTGGGAACAGCGACGGCTACTACTCCGGCGTGGTTTGTAAAAAAGTAT

CCTGATTCACTCCCGATCGATGAATCCGGCGTCATCTTATCTTTTGGTTCTCGGCAACATTATTGCCCCG
AACCACCCTCAATTGATCACACATATTAAACGTCTTGTTCGGGCTATCGCCGAACGCTATAAAAACCA
TCCAGCGCTTAAGATGTGGCACGTCAATAACGAGTATGCATGTGTCATGTGTCAAAATGCTTTTGCGAGA
ATTGTGCTGTGGCGTTTCGAAAATGGTTAAAAAGAACGTTATAAAACAATCGACGAACCTAACGAAAAGA
TGGGGGCTAAATTTCTGGGGCCAGAGATATAATCACTGGGACGAAATTAACCCGCCGCGAAAAGCAC
CGACATTCATCAATCCTAGCCAAGAACTTGACTATTATCGATTTATGAACGATTCTATCCTGAAAACGT
TTTTAACAGAAAAAGAAATTTTAAGAGAAGTGACACCTGACATCCCTGTCTCAACAAATTTTCATGGGA
AGCTTCAAGCCGTTGAACTATTTTCAATGGGCACAACACGTGGATATTGTTACGTGGGATTCTTATCCG
GACCCCTCGCAAGGTCTTCCGATCCAACATGCCATGATGAATGATTTAATGCGTTCACTCAGGAAGGG
CCAACCGTTTATTCTGATGGAGCAGGTAACAAGCCATGTGAATTGGAGAGATATCAACGTCCCTAAAC
CGCCAGGCGTAATGAGGCTGTGGTCTATGCAACTATTGCGCGGGGAGCAGATGGCATTATGTTTTTT
CAGTGGCGTCAATCACGAGCTGGGGCTGAAAAATTTACGGAGCGATGGTGCCGCATTTTTTAAATGA
AAACAATAGAATTTATCGCGAAGTTACACAGCTTGGACAGGAAGTGAAGAACTTGATTGTTTGGTCG
GGTCACGTATAAAAGCCGAAGTCGCGATAATATTTGATTGGGAAAACCTGGTGGGCCGTAGAGTTATCC
TCGAAACCACACAACAACTCCGATACATTCCCATTTGTTGAAGCATACTATCGCGAAGTGTATAAAAG
AAATATTGCAGTGGATTTTGTGCGGCCATCTGATGACTTGACGAAGTATAAGGTTGTGATCGCACCAA
TGTTATATATGGTAAAAAGAAGCCGAAGATGAAAACCTTACGACAATTCGTGGCTAATGGAGGCACATTG
ATCGTCTCATTCTTTAGCGGAATTGTGGACGAAAACGATCGCGTTTATTTAGCGGTTATCCTGGCCCT
CTCCGCGATATCCTTGGCATATTTGTTGAGGAGTTTGTACCTTACCCAGAACTAAGGTGAATAAAATT
TATTCCAACGATGGGGAATATGACTGTACAACATGGGCGGACATTATTCGTTTAGAAGGTGCGGAGCC
GTTAGCTACCTTTAAAGGCGATTGGTATGCGGGCCTGCCAGCGGTCACGAGGAATTGCTATGGTAAAG
GCGAAGGTATTTATGTGGGCACTTACCCTGATAGCAATTATTTAGGCAGACTGTTGGAACAAGTGTTT
GCCAAACATCACATTAATCCTATCCTCGAAGTGGCGGAGAATGTGCAAGTCCAACAGAGGGGAAACAG
ATGAATGGAATATCTGATTATTATTAACCAACGATTATGAAGTAACCTGTCTTACCGGAAGAT
AAGATCTACCAAAATATGATAGACGGAATAATGTTTTAGGGGCGGAGAGCTGCGTATCCAAGGGGTGG
ACGTTGCAGTCTTACGGGAACACGACGAAGCAGGAAAAAGTTAA

SEQ ID NO: 41

MKKAISCVFILSALILSSFQVPVQGGQAMSKTSSAAGNSVSYDGERRVNFNENWRFQRETNNGSIAGAQNPGF
DDSSWRKLNLPDWSIELDFNKNSLATHEGGYLDGGIGWYRKFTFTIPESMKGKRISLDFDGVYMNSTTYL
NGEVLGTYPFGYNAFSYDISDKLYKDGGRANVLVVKVNNTQPPSRWYSGSGIYRNVYLTVDPIHVARYGT
FVTTNPLEKSIKEDRADVNIKTISNDAAEAKQVQIKSTIYDGAGNTVQTVETEEKTAAAGTVTPFEQNTVI
KQPKLWSIDKPYRYNLVTEVIVGGQTVDTYETKFGVRYFKFDENEGFSLNGEYMKLHGVSMHDLGALG
AATNARGVERQMQIMKDMGVNAIRVTHNPASPELLEAANKLGLFIIIEAFDSWAQSKKPYDYGRFFNAW
AEHDIKEMVDRGKNEPAIIMWSIGNEIYDITNAAGVETARNLVGWVKEIDTTRPTTIGEDKTRGDKVNVTP
INSYIKEIFNIVDVVGLNYSNNYDGYHKQNPWSKLYGSETSSATRSRGVYTHPYQYNQSTKYADLQSSY
DNDYVGVWGRTAEDAWKYDRDLKHIAGQFIWTGFDYIGEPTPYNSYPKSSYFGAVDTAGFPKDIFYYYQ
SQWKKEPMVHLLPHWNWKEGKVRVLAYTNASKVELVLNGESLGEKNYDNKQTSWGAPYKETKDGMT
YLEWAVPFKPGKLEAVAKDENGKVIARDQVVTAGEPASVRLTADRKVVKADGTDLSFITADIVDSKGIVV
PDADHLITFNVTTGQELAGVDNNGNASSVERYKDNKRKAFSGKALAIQSSKLSGKITVHASVAGLSSDSTS
VFTVTPADHDKKIVAGIDDVNLTVDVNEAPKLPEIKVYYSDSAAAKNVTWDEVDPKQYSTVGFTVEG
SVEGTSCLKAKAFVIVKGIVAVKPYSTATKVGVPVLPEKATLLYSDGTTKGATVTWDEIPEDKLAKEGRFT
VEGSVEGTDLKANVYVRVTNEVKSVMIMLQEQGSAYPKLEATFTNPADNLQHLNDGIKSYTNNPVNRWT
NWTRTPRDAGDSITVNFQKKHVINNLDFVFTDSGTVVPEKAEVQYWDGTAWKDVENLTQPSPYVVEKN
ELTFDAVATEKLKFHLTPSVKGKFLALTEAEVYADQIVMGETAKLQSITVNGKALEGFDHAKKNYELVLP
YGSELPKIEAAAADNATVTILPAFSYPGTAKLFVTSDEGKVTTEYSIGVSTEEPKLVSALSDAKTNVMEDD
IIDLKVIGLFESKEKIDVTDQSPTYEFDQIIEGKNKLYALETGNVKVKVTVTYKGVSVTTPALEFTIAKNPA
PKYITSLEPVTVVVKKGEAPELPATVVAHYNRGIPRDVKVKWERINPSKYQQLGEFTVSGMVEGTDIKAQA
KVIVKGAVAVEDIRMAVLLKQMPQLPGKVTVYYSDGAEQRAVKWEEIPQEELENVGEFKVKGDVNGVK
LKATATIRVTDEVGGEQNISRANKNGEYYPKAEASFTNNNGPGSSDRIEAINDDVISYEANPHNRWNTWQPPV
RAGDWVSITFGDYEPTEYDVDSMEIHWFADHGTSYPERFQIEYKSGDSWKEVTSLSKSDPASPALGKANVY
SFDVRKTSAIRVKMTAQAGKSLAITELKVFSKWPKAGTEPEVTDIKVGGKSILEDFEQKGDHYEVTIDAGD
ANVMPKINVAKAQDQTSITVPAVTSPSTAKVIAKSEDGKKVKVYSIHYK

SEQ ID NO: 42

ATGAAGAAAGCGATCAGCTGCGTTTTTCTGATATCTGCACTCATCTTAAGCAGCTTTCAAGTACCGGTG
CAGGGGCAAGCGATGTCTAAACTACGTACGCGGCCGGAACAGCGTTTCATACGATGGTGAACGGA
GAGTTAATTTCAATGAAAATTGGCGGTTCCAGCGGGAACCAATGGTAGCATTGCAGGTGCACAAAAT
CCAGGGTTTGACTCTTCTGGAGAAAGCTTAATTTACCGCATGACTGGAGCATAGAAGTTGATTTT
AATAAAAACTCATTGGCGACGCACGAAGGGGGGTATTTGGACGGCGGCATCGGGTGGTACCGAAAA
CGTTCACCATACCTGAATCAATGAAAGGCAAAAGAATCTCTCTGGATTTTCGATGGTGTATATGAAC
TCAACAACATATTTAAACGGGGAGGTATTAGGTACATATCCGTTTGGCTATAACGCCTTTTCTTACGAC
ATTTCCGATAAACTGTATAAAGATGGAAGAGCGAACGTGCTTGTAGTGAAGGTCAATAATACGCAACC
ATCTAGTCGATGGTATTCAGGAAGCGGCATATATCGTAATGTGTATTTAACTGTAACGGATCCCATTCA
TGTTGCGCGGTATGGCACATTTGTCAACACCTAACTTGGAGAAATCAATTAAGAAGATCGGGCAG
ACGTCAACATTTAAACGAAAATTTCTAACGATGCGGCGGAAGCAAAACAAGTCAAAATCAAATCAAC
AATATATGACGGAGCTGGTAACACAGTACAGACGGTTGAAACGGAGGAAAAACAGCAGCAGCAGGT
ACCGTAACCCCGTTTGAGCAAAACACGGTAATAAAACAGCCTAACTGTGGTCCATCGACAAACCGTA
CCGTTATAACCTGGTAACGGAGGTTATTGTTGGCGGCCAAACGGTGGATACTTATGAAACTAAATTCG
GAGTCAGGTATTTTAAAGTTTGACGAGAATGAGGGGTTTTCTCTGAACGGAGAGTACATGAAGCTTCAT
GGCGTCTCCATGCATCAGATTTAGGAGCTTTAGGCGCGGCACAAACGCTAGAGGGGTTGAAAGGC
AGATGCAGATTATGAAAGACATGGGAGTCAATGCCATCAGAGTAACACATAACCCGGCAAGCCCTGA
ACTTCTGGAAGCCGCGAACAAGCTCGGCCTGTTTATTATTGAGGAAGCTTTTGATTTCGTGGGCTCAGA
GTAAGAAACCCTATGACTATGGAAGATTTTTCAATGCGTGGGCGGAGCACGATATAAAAGAAATGGTT
GACAGAGGGAAAAACGAACCCGCAATCATCATGTGGAGTATTGGGAATGAAATCTATGATACGACTA
ATGCAGCGGGCGTGGAACCCGCGAGGAATCTTGTGGGCTGGGTCAAAGAAATTGACACTACACGGCC
CACAACAATCGGCGAAGACAAGACTAGGGGCGACAAAGTAAACGTGACGCCAATTAACAGTTATATC
AAAGAAATTTTCAACATTGTTGACGTAGTTGGGCTGAACACTACTCGGAAAACAATTATGGATGATCA
CAAACAGAATCCCTCTTGGAACTGTATGGATCCGAAACAAGCAGTGCGACAAGATCAAGGGGTGTTT
ATACACACCCATATCAGTATAACCAATCTACAAAATACGCCGATCTGCAACAATCCTCTTATGACAAC
GACTATGTTGGCTGGGGTAGAACAGCGGAAGACGCGTGGAAGTACGACAGAGATTTGAAACATATCG
CGGGACAGTTTATTTGGACGGGATTTCGATTACATTGGTGAACCAACCCCGTACTATAATAGCTATCCTG
CGAAAAGCAGCTACTTCGGAGCAGTGGATACAGCTGGATTCCCTAAAGATATCTTTTACTATTATCAG
TCGCAATGGAAAAAGGAGCCGATGGTACATCTTCTCCGATTGGAAGTGGAACTGGAAGGAAGGAGAAGG
TCAGAGTACTGGCCCTACACCAATGCTAGCAAAGTCAAACTGTACTGAATGGCGAAATCCCTCGGAGAG
AAAAATTATGATAATAAAACAAACGTCCTGGGGAGCTCCTTATAAGGAAACAAAAGACGGAAGACAT
ATCTGGAATGGGCGGTCCCTTTCAAGCCTGGCAAACTGGAAGCCGTGGCAAAAGATGAAAACGGCAA
GGTAATTGCAAGAGATCAAGTAGTGACGGCGGGAGAGCCCGTAGCGTTAGATTGACAGCTGATAGG
AAAGTTGTGAAAGCAGATGGAACAGATTTAAGCTTTATTACCGCAGATATCGTAGATTCTAAGGGCAT
TGTTGTCCCGGATGCTGATCATTTGATTACATTCAATGTGACTGGACAAGGAGAAGTGGTGGAGTAG
ATAATGGTAATGCATCATCCGTTGAACGGTACAAAGATAATAAAAGAAAGGCCTTTTCAGGTAAGGCC
TTGGCAATCGTACAGTCATCAAACTTAGCGGTAAAATCACCGTACACGCCTCGGTTGCAGGTCTTTC
ATCCGATAGTACAAGTGTCTTCACAGTCACACCGGCGGATCACGATAAAAAAATTGTAGCAGGTATTG
ACGATGTGAACCTTACGGTTGACGTCAACGAAGCGCCAAAACCTCCGTCTGAAATTAAGTCTATTAT
TCAGACGAATCTGCGGCCGCGAAAAATGTTACATGGGATGAAAGTGACCCCTAAACAGTATTCTACAGT
TGGCGAGTTTACAGTCGAAGGCTCGGTTGAGGGCACAAAGCTTAAAGGCGAAAGCGTTTGTATAGTTA
AAGGTATCGTCGCCGTTAAACCCTATTCTACGGGACCAAAAGTTGGAGTTCAACCTGTTTACCTGAGA
AAGCAACATTACTTTATAGCGATGGCACTACTAAAGGGGCCACAGTGACATGGGATGAAATCCCGAGA
AGATAAATTGGCGAAAGAGGGCCGTTTACGGTAGAAGGTTTCAGTAGAGGGCACAGACTTGAAAGCG
AATGTTTATGTGCGAGTTACGAATGAGGTAAAGTCTGTTAACATTATGCTTCAGGAGCAGGGCAGCGC
ATATCCAAAGCTTGAGGCGACCTTTACAAATCCGGCGGACAATCTGCAGCACCTGAATGATGGTATCA
AATCTTATACAAACAATCCGGTGAATAGATGGACGAATTGGACAAGAACCAAGGGACGCGGGAGA
CAGTATCACGGTGAACCTTTGGGAAAAAACAGTAATCAACAATCTTGACCTTTTTGTATTTACAGATA
GTGGTACGGTTGTCCCTGAGAAAGCTGAAGTTCAATATTGGGATGGAAGTGCCTGGAAAGACGTAGAG
AATTTAACACAGCCTTACCCTGACGTTGTGGAAAAAACAGAGTTGACTTTCGATGCAGTCGCGACAGA
AAAATTTAAATTTTCTGACACCGTCTGTCAAAGGTAAATTTCTGGCCCTGACAGAAGCAGAAGTAT
ACGCAGACCAGATTGTAATGGGAGAAACAGCTAACTCCAGTCCATAACGGTCAATGGTAAGGCACT
CGAAGGGTTTGATCATGCGAAAAAGAATTATGAGCTCGTACTGCCGTACGGGTCCGAATTGCCTAAGA
TTGAAGCAGCTGCGGCTGACAATGCTACCGTGACGATTTTACCTGCCCTTTCTTATCCAGGAACGGCGA

AGTTGTTTCGTAACAAGTGAGGACGGAAAAAGTTACAACCGAATATAGCATTGGCGTTCGACTGAAGAG
CCAAAGCTGGTCTCAGCTGAGTTGAGCGCAGATAAGACCAATGTCATGGAAGATGATATAATTGACCT
GAAAGTGATAGGCCTCTTTGAGAGCAAGGAAAAAGATTGACGTCACAGACAGCCAACCAACCTACGAG
TTCGACCAGCAAATTATAAAGATCGAAGGAAATAAATTGTATGCGCTGGAAACCGGTAATGTCAAAGT
AAAGGTCACCTTACTTATAAAGGAGTCTGTTACGACACCAGCCCTTGAATTCACGATCGCGAAGA
ATCCTGCTCCGAAATATATTACATCGCTCGAGCCAGTCACAGTAGTTGTGAAAAAAGGCGAAGCTCCC
GAACTCCCTGCGACTGTTGTCGCTCATTATAACCGCGGTATCCCGCGGGATGTTAAAGTTAAATGGGA
ACGGATAAATCCATCAAAGTACCAACAATTAGGCGAATTTACCGTTTCGGGAATGGTGGAAGGAACTG
ATATTAAGGCCAAGCTAAAGTGATAGTGAAAGGAGCAGTTGCACTAGAGGATATTAGAATGGCAGT
CCTTTTGAAGCAAATGCCTCAGTTACCTGGAAAGGTTACGGTATACTATAGTGATGGTGCGGAAGAAC
AACGAGCTGTTAAATGGGAGGAGATACCGCAAGAAGAAGCTGAAAACGTGGGAGAATTTAAGGTGAA
AGGAGATGTTAATGGTGTTAAGCTGAAAGCAACCGCGACCATCCGGGTAACGGATGAAGTGGGGGGA
GAACAGAATATTTTCGCGGGCAAAAAATGGCTATGAGTACCCGAAGGCTGAGGCTAGTTTACAAATA
ATGGGCCCGGGTCTAGCGACCGTATTGAGGCTATCAATGACGACGTTATCTCATACGAAGCAAATCCG
CACAACAGGTGGACCAACTGGCAGCCGGTGCCTAGAGCTGGCGACTGGGTCTCCATTACCTTTGGTGA
TTATGAACCAACAGAATACGACGTTGACTCGATGGAAATCCATTGGTTTTCGGGACCACGGAACCTCTT
ACCCAGAGAGGTTTCAGATCGAATATAAAGCGGAGACTCGTGGAAGAAGTCACGAGCCTGAAGAG
TGACCCGGCCTCTCCAGCCCTGGGAAAAGCTAATGTTTATTTCATTCGACCGTGTGAAGACCTCTGCTAT
TAGGGTAAAAATGACAGCTCAGGCGGGTAAAAAGTCTTGCTATTACGGAATAAAAAGTCTTTTCAAAAT
GGCCTAAAGCAGGCACCGAACCTGAGGTCACCGATATCAAGGTTGGCGGCAAATCTATCCTCGAAGAT
TTTGAACAAAAAGGCGACCATTAAGTGACAATTGATGCGGGTGATGCGAATGTAATGCCGAAAA
TCAATGTGAAAGCGAAGGATCAAACGTCTATCACAATCGTCCCGCGGTTACTAGTCCGTCAACGGCA
AAAGTGATCGCTAAATCGGAAGATGGCAAAAAAGTGAAAGTATATTCAATACATTATAAATAA

SEQ ID NO: 43

MPEVRSSTQSESGMSQWMGKILSIRGAGLTIGVFGLCALIAATSVTLPEEQLIVAFVVCVFFIVGHKPSRR
SQIFLEVL SGLVSLRYLTWRLTETLSFDTWLQGLLGTMLLVAELYALMMLFLSYFQTIAPLHRAPLPLPPNP
DEWPTVDIFVPTYNEELSIVRLTVLGSGLGIDWPPEKVRVHILDDGRRPEFAFAAECCGANIARPTNEHAKA
GNLNYAIGHTDGDYILIFDCDHVPTRAFLQLTMGWMVEDPKIALMQTPHHFYSPPDFQRNLSAGYRTPPEG
NLFGYGVVDGNDFWDATFFCGSCAILRRRTAIEQIGGFATQTVTEDAHTALKMQRLLGWSTAYLRIPLAGGL
ATERLILHIGQVRVRWARGMLQIFRIDNPLFGRGLSWGRLCYLSAMTSFLFAVPRVIFLSSPLAFLFFGQNIIA
ASPLALLAYAIHPMFHAVGTASKINKGWRYSFVSEVYETTMALFLVRVTIVTLLSPSRGKFNVTDKGGLLE
KGYFDLGAVYPNIILGLIMFGGLARGVYELSFHGLDQIAERAYLLNSAWAML SLIILAAIAVGRETQQKR
SHRIPATIPVEVANADGSIIVTGVTEDLSMGAAVKMSWPAKLSGPTPVYIRTVL DGEELILPARIIRAGNGR
GIFIWTIDNLQQEFVIRLVFGRADAWVDWGNKYADRP LLSLMDMVL SVKGLFRSSGDIVHRSSTPKPLAG
NALSDDTNPNPSRKERV LKGTVMVSL LALLTFASSAQAASAPRAVAAPAHQPEASDLPLLPALLPATSG
AAQAGAGDAGANGPGSPTGQPLAADSADALVENAENTS DTVHNYTLKDLGAAGSITMRGLAPLQGIEF
GIPSDQLVTSARLVLSGMSMPNLRPETNSVTMTLNEQYIGTLRPDPAHPTFGPMSFEINPIFFVSGNRLNFNFA
SGSKGCS DITNDTLWATISQNSQLQITIALPPRRLSRLPQPFYDKNVRQHVTVPMLVAQTYDPQILKSAGI
LASWFGKQTD FLGVTFPVSS TIPQSGNAILIGVADELPTSLGRPQVNGPAVLELPNPSDANATILVVTGRDRD
EVITASKGIAFASAPLPTD SHMDVAPVDIAPRKPNDA PSFIAMDHPVRFGDLV TASKLQGTGFTSGVLSVPF
RIPPDLYTWRNRPYKMQVRFRSPAGEAKDVEKSRLDVGINEVYLHSYPLRETHGLVGAVLQGVGLARPAS
GMQVHDL DVPPWTVFGQDQLNFYFDAMPLARGICSGAANNAFHLGLDPDSTIDFSRAHHIAQMPNLAY
MATVGFPTTYADLSQTA VVLPEHPNAATVGAYLDLMGFMGAATWYPVAGVDIVSADHVS DVADRNL
VISTLATS GEIAPLLSRSSYEADGHLRTVSHASALDNAIKAVDDPLTAFRDRDSKPQD VDTPLTGGVGAMI
EAESPLTAGRTVLALLSSDGAGLNNLLQMLGERKKQANIQGD LVVAHGEDLSSYRTSPVYTIGTLPLWLW
PDWYMHNRPVRLVGLLGCILIVSVLARALARHATRRFKQLEDERRKS

SEQ ID NO: 44

ATGCCAGAGGTCCGATCTTCGACGCAGTCAGAGAGCGGGATGTCCCAGTGATGGGAAAAATCTTATC
AATAAGAGGGGGCAGGACTGACGATAGGAGTATTTGGCCTCTGTGCTCTTATAGCCGCGACAAGCGTTA
CACTTCCACCTGAACAACAACCTATTGTGGCGTTCGTCTGTGTTGTAATCTTTTTATCGTCGGACACA
AGCCATCACGGAGATCGCAAATCTTCTGGAAGTACTTCCGGACTTGCTCTCTCTTCGCTACCTTACCT
GGCGGTTAACGGAGACCCTGTCCTTTGATACCTGGTTACAGGGTCTTCTGGGCACGATGTTACTGGTCG
CGGAATTATATGCCCTGATGATGTTATTCTTGTCTACTTCCAGACAATAGCACCGCTGCATAGAGCGC

CCCTCCCACTCCCCCGAATCCTGATGAATGGCCTACAGTAGATATATTCGTCCCGACATACAATGAG
GAACTTTCAATCGTAAGACTGACAGTTTTGGGGAGTTTTGGGAATAGATTGGCCGCCGAAAAAGTTAG
AGTACATATCCTTGACGATGGGAGGCGTCCAGAAATTTGCCGCCTTTGCAGCCGAGTGCGGAGCAAAC
ATATTGCTAGACCTACAAATGAGCATGCGAAAGCTGGAAATTTAAATTATGCGATTGGCCATACAGAC
GGGACTATATATTAATTTTTGATTGTGACACGTACCAACACGCGCTTTCCTGCAGCTTACCATTGGGA
TGGATGGTTGAAGATCCAAAAATCGCGCTTATGCAAACTCCGCATCACTTTTACAGCCCTGACCCATT
CAACGCAATTTGTCTGCGGGGTATAGAACTCCGCCTGAAGGCAACTTATTTTACGGCGTAGTGCAGGA
TGGAATGATTCTGGGATGCTACTTTCTTCTGCGGTAGCTGCGCTATTCTTAGACGGACAGCCATAGA
ACAAATTGGGGGATTTCGCGACGCAGACCGTAACGGAGGACGCTCATACAGCCCTTAAATGCAAAGA
CTTGGTTGGTCAACAGCGTATTTGAGAATACCATTTGGCAGGTGGTCTTTCGACGGAAAGATTAATCCT
GCATATCGGACAACGCGTTTCGTTGGGCACGGGGCATGCTGCAAATCTTTCGCATAGATAACCCGCTCT
TTGGACGCGGCTTATCATGGGGGCAGAGGTTGTGTTATTTAAGTGCCATGACCTCCTTTCTTTTTGCGG
TGCCGCGTGTATTTTTTTAAGTTACCTCTGGCCTTTCTGTTTTTTGGACAAAACATAATTGCTGCAAG
CCCCTTGGCTTTACTGGCTTATGCGATCCCCATATGTTTACGCGCTTGGAACTGCATCTAAGATCAA
TAAAGGTTGGCGTTATTCGTTCTGGAGTGAAGTGTACGAGACTACAATGGCGCTGTTCTGGTGCGTGT
GACCATTTGTCACACTGCTCTCTCCTAGCAGAGGCAAGTTAATGTGACAGACAAAGGGGGTCTGCTCG
AAAAAGGTTACTTTGATCTTGGAGCGGTATATCCTAATATCATTCTTGGCTTGATCATGTTTGGAGGGT
TGGACGCGGTATACGAATTGTCGTTTGGACATCTTGATCAAATCGCCGAACGTGCTTATTTGCTGA
ATTCAGCGTGGGCAATGTTATCACTTATAATCATCTTAGCTGCGATTGCGGTAGGCAGGGAGACACAA
CAGAAAAGAAATTCTCACAGGATTCCGGCCACAATTCCGGTTCGAGGTGCGCAATGCTGACGGAAGCAT
AATTGTACAGGAGTACAGAAAGATTTGTCTATGGGTGGAGCGGCAGTTAAAATGAGCTGGCCGGCG
AAGCTGTCCGGCCCTACGCCGGTTTATATTAGAACAGTACTGGATGGGGAGGAATTAATCCTTCCCGC
GAGGATCATTAGGGCTGGGAACGGCCGAGGCATTTTTATCTGGACGATTGACAATTTGCAGCAAGAAT
TTCTGTCTCCGTCTGGTATTCGGGCGTGGCATGCGTGGGTTGACTGGGGAACTACAAGAGAT
CGTCCGCTGCTGTCTTAATGGATATGGTGTCTCGGTTAAAGGACTTTTTCTGTTCCAGCGCGATATT
GTTTCATCGATCTTCTCTACAAAACCGTTAGCGGGGAATGCCTTGAGTGACGATACAAATAACCCTAG
CAGAAAAGAACGTGTTTTGAAAGGAACGGTGAAGATGGTATCACTGCTTGCACTGTTGACATTTCGAA
GCTCGGCACAGGCCGCTTCCGCTCCTAGAGCGGTGGCAGCTAAGGCTCCGGCTCATCAACCGGAAGCT
TCCGATCTGCCCCGCTGCCTGCCTTACTTCCGGCGACAAGCGGAGCAGCACAAAGCAGGCGCCGGAGA
CGCGGGAGCAAATGGACCGGGGTCTCAACCGGACAACCCCTTGGCGCAGATTCCGCAGATGCGCTTG
TTGAGAATGCCGAGAACACGTCCGATACGGCCACAGTTTCACAATTACACACTTAAGGACCTGGGAGCT
GCTGGCTCCATTACTATGAGGGGACTGGCACCCCTGCAGGGAATTGAGTTCGGCATCCCAAGCGATCA
ACTGGTAACCTCGGCTCGCTTGTACTTTAGGAAGCATGTCACCTAACCTCCGACCGGAAACGAATT
CTGTTACAATGACCCTCAACGAGCAATATATTGGTACATTGAGACCTGACCCAGCACACCCGACGTTT
GGGCCAATGAGTTTCGAAATCAATCCGATTTTTTTGTGTCTGGAAACAGATTAATTTCAATTTCTGCT
TCGGGGTCAAAAGGCTGCTCTGACATTACTAATGACACGCTCTGGGCAACAATCTCGCAAAATAGTCA
ACTCCAGATAACAACAATTGCCCTTCCCCGAGAAGATTGTTAAGCAGGCTGCCCAACCTTTCTATG
ATAAAAATGTACGGCAACATGTCACGGTCCCTATGGTACTTGCCAGACCTATGATCCCCAGATTCTTA
AGAGCGCAGGCATTCTCGCTTCTTGGTTTGGCAAGCAGACAGACTTTCTTGGTGTAACGTTTCCGGTTT
CTTCAACGATCCCCGAGTCCGGAACGCGATCCTTATCGGGGTGGCGGATGAATTACCGACAAGCTTG
GGCAGGCCTCAAGTTAACGGTCCGGCAGTACTGGAATTGCCGAATCCTTCCGATGCTAACGCCACAAT
CCTTGTGTGACCGGGCGTGATCGAGATGAAGTAATTACAGCCAGCAAAGGGATCGTTTTGCCTCGG
CACCCTTCTTACGGATAGCCATATGGATGTCGCGCAGTTGATATCGCCCCACGTAAGCCTAATGAT
GCGCCCAGTTTATCGCAATGGATCATCCTGTACGGCTTTGGAGACCTCGTTACAGCATCCAAATTACAA
GGCACAGGGTTTACGTACGGCGTTTTGAGCGTTCCATTCCGGATCCCTCCTGACTTATACACCTGGCGG
AATCGTCTTACAAAATGCAAGTTCGCTTTCGTTTACCAGCGGGCGAAGCGAAAGATGTGGAAAAATC
TCGGCTTGATGTCGGAATCAACGAAGTTTATTTGCATAGTTATCCCTTACGAGAGACACACGGCCTGGT
CGGTGCCGTTTTACAAGGGGTTGGCCTTGCTAGACCAGCATCTGGTATGCAAGTGCACGACCTGGATG
TCCCTCCATGGACCGTGTGTTGGACAAGATCAATTAATTTTTATTTTATGCGGATGCCGCTTGTCTCGG
GGATTTGTGAGTCCGGAGCAGCAAATAATGCGTTCATCTGGGATTGGACCCAGACTACTATAGAC
TTTTCTCGGGCCCATCATATAGCGCAGATGCCTAACCTTGGCTACATGGCAACCGTTGGCTTCCATTT
ACTACCTATGCCGATTTGAGCCAGACAGCTGTCTCTGCCGAGCATCCCAATGCGGCAACAGTTGG
AGCTTATCTTGACCTGATGGGGTTTATGGGCGCGGCCACATGGTATCCGGTGGCGGGTGTGGATATTG
TCAGTGCGGATCACGTTAGCGATGTCGCGGACCGGAATCTCCTTGTTATTTCTACACTGGCTACATCTG
GTGAGATCGCACCGCTTCTTCTCGTTCCAGCTATGAGGTAGCTGATGGACATTTGCGTACCGTTAGCC

ATGCTAGCGCGTTAGATAACGCTATTAAGGCAGTGGATGATCCTTTAACCGCCTTTAGAGACCGGGAC
TCAAAACCACAAGACGTGGATACGCCACTTACTGGCGGGGTCGGTGCTATGATCGAAGCAGAATCACC
GCTCACCGCTGGACGTACGGTGTAGCCCTTTTGTTCATCCGATGGCGCTGGCCTGAATAATCTCTTGCA
GATGCTGGGCGAAAGGAAAAAGCAGGCCAATATACAAGGGGATCTTGTGTAGCCACGGAGAAGAT
TTGTTCATCTACCGTACATCTCCGGTCTATACCATCGGCACGCTGCCATTATGGCTGTGGCCGGACTGG
TACATGCACAATCGGCCAGTGAGGGTACTTCTTGTGGGGCTTCTGGGTTCATTCTCATAGTTTCCGTG
CTTGCGCGCGCTTTAGCTAGACATGCGACTAGACGGTTTAAACAACCTGGAGGACGAGAGACGCAAATC
GTAA

SEQ ID NO: 45

MATPSAVGAACLLARAAWPAAVGDRARPRRLQVLRRLRRCVAELSREGPAPRPLPALLAPPLVPGFLAPP
AEPTGEPASTPPVPDAGLGLGLEPEGIAEGSIDNTVVVASEQDSEIVVGKEQARAKVTQSIVFVTGEASPY
AKSGGLGDVCGSLPVALAARGHRVMVMPRYLNGTSDKKNYANAFYTEKHIRIPCFGGEHEVTFHEYRDS
VDWVFDHPSYHRPNLYGDKFGAFGDNQFRYTLCLYAAACEAPLILELGGYIYGQNCMFVNDWHASLV
PVLLAAKYRPHYGVYKDSRSILVIHNLAHQGVPEASTYPDLGLPPEWYGALEWVFPEWARRHALDKGEAV
NFLKGAVVTADRVTVSKGYSWEVTTAEGGQGLNELLSSRKSVLNGIVNGIDINDWNPATDKICIPCHYSVD
DLSGKAKCKGALQKELGLPIRPDVPLIGFIGRLDYQKGIDLIQLIIPDLMREDVQFVMLGSGDPELEDWMRS
TESIFKDKFRGWVGFSPVSHRITAGCDILLMPSRFEPCLNQLYAMQYGTVPVHATGGLRDTVENFNPF
GENGEQGTGWAFAPLTENMLWTLRTAISTYREHKSSWEGLMKRGMKDFDTHDAAEQYEQIFQWAFID
RPYVM

SEQ ID NO: 46

ATGGCGACCCCGTCCGCAGTGGGCGCAGCCTGCCTTTTATTAGCGCGCGCAGCTTGGCCGGCAGCCGT
TGGAGACAGGGCAAGACCGCGGCGATTACAACGCGTGTGCGGCGGAGATGTGTAGCAGAACTTTCT
CGTGAAGGGCCAGCACCGAGGCCTTTACCTCCTGCCTTGCTTGCCCCCGTTAGTACCTGGTTTTCTG
GCACCCCGAGCGGAACCAACAGGCGAACCAGCGAGCACGCCTCCGCCGGTTCGGGACGCTGGACTTG
GCGACTTGGGATTAGAACCTGAAGGAATCGCGGAAGGTTCAATCGATAATACCGTCGTGGTGGCTTCT
GAACAGGATAGTGAGATCGTAGTTGGGAAAGAGCAGGCTCGCGCAAAAGTAACGCAATCAATTGTAT
TCGTAACCGGCGAGGCAAGCCCCTATGCGAAATCTGGAGGCCTGGGCGATGTTTGCGGAAGCCTTCCG
GTTGCGTTAGCTGCCAGAGGACATCGAGTCATGGTCGTCATGCCGCGGTATCTGAACGGAACGTCAGA
TAAAAATTATGCCAATGCCTTCTATACCGAGAAAGCATATCCGGATCCCTTGCTTTGGTGCGGAACACG
AAGTGACTTTTTTTCATGAATATCGTGACTCAGTCGACTGGGTTTTTGTGACCAACCCGAGCTATCATA
GACCGGGTAACCTGTACGGGGATAAATTTGGAGCGTTTGGCGATAATCAATTCCGGTATACCTTGCTG
TGTTATGCCGCATGCGAAGCCCCCTCATCTTGGAACCTCGGAGGCTATATTTATGGACAAAACCTGTATG
TTCGTAGTAAACGATTGGCACGCATCACTCGTACCAGTACTTCTCGCAGCGAAATATAGACCGTATGG
CGTTTACAAAGATTCCAGATCAATTTTAGTTATTACAACTTAGCTCACCAAGGCGTAGAACCGGCGTC
CAGATATCCAGATCTTGGAATTGCCGCCAGAGTGGTATGGAGCGCTTGAATGGGTCTTTCCCTGAATGGG
CTCGTCGACATGCGCTGGATAAAGGTGAAGCTGTCAATTTTCTCAAAGGTGCTGTGGTCACTGCCGAC
AGAATTGTAACAGTGAGCAAAGGCTATTCTGGGAAGTTACCACCGCTGAGGGTGGCCAAGGGCTCA
ATGAATTGCTGAGCAGCCGTAAGAGTGTGTTGAATGGTATAGTGAATGGTATCGACATCAATGATTGG
AACCCGGCAACAGACAAATGTATCCCCTGTCACTACTCCGTCGACGACCTTTCGGGAAAAGCAAAGTG
TAAAGGCGCGCTTCAAAAAGAGTTGGGCTTGCCGATTAGACCGGATGTGCCTCTTATTGGATTCAATTG
GCCGTTTGGATTATCAGAAGGGAATTGATCTGATCCAGCTGATTATTCCGGACTTGATGAGAGAAGAT
GTCCAGTTTGTGATGTTGGGCTCCGGCGATCCAGAAGTTGAAGATTGGATGCGGAGCACCGAATCAAT
CTTTAAGGATAAATTTAGAGGATGGGTCCGGTCTCTGTGCCTGTCTCACATCGCATTACGGCGGGCTG
CGATATCCTCCTTATGCCTTCTCGGTTCTGAACCGTGTGGTTTAAATCAACTTTATGCGATGCAGTACGG
CACTGTGCCGTTGTTTACGCGACTGGAGGGCTGCGAGATACTGTTGAGAATTTAATCCGTTTGGAG
AGAACGGTGAACAAGGAACAGGATGGGCCTTCGCACCACTGACTACGGAGAACATGCTGTGGACACT
TAGAACGGCCATCTCTACGTATAGGGAGCACAAGTCCTCGTGGGAGGGACTCATGAAACGGGGGAATG
AGTAAAGATTTACCTGGGATCACGCTGCAGAACAAATATGAGCAAATCTTTAGTGGGCGTTTATCGA
TCGCCCTACGTTATGTGA

SEQ ID NO: 47

MVSLSNQTRFSFHPNNLVVSEKRRLGISGVNFPRKIKLKITCFAAERPRQEKQKKKSQSQSTSDAEAGVDPV
GFLTRLGIADRIFAQFLRERHKALKDLKDEIFKRHFDFRDFASGFELLMHRHMEHRVDFMDWGPGRSYG

AIIGDFNGWSPTENAAREGLFGHDDYGYWFIILEDKLRERGEEDLYFQQYNYVDDYDKGDSGVSAEEIFQ
 KANDEYWEPGEDRFIKNRFEVPAKLYEQMFGPNSPQTLEELGDIPDAETRYKQWKEEHKDDPPSNLPPCDII
 DKGGQKPYDIFNVVTSPEWTKKFYEKEPPIPYWLETRKGRKAWLQKYIPAVPHGSKYRLYFNTPDGPLERV
 PAWATYVQPEDEGKQAYAIHWEPSPEAAKYWKYSKPKVPESLRIYECHEVVISGSEPKVSTFEFTKKVLPH
 VKRAGYNAIQILIGVPEHKDYFTVGYRVNFFAASSRYGTPDDFKRLVDEAHGLGLLVFLDIVHSYAAADQ
 MVGLSLFDGNSDCYFHYGKRGHKHWGTRMFKYGDLDVLHFLISNLNWWITEYQVDGYQFHSLSASMIYT
 HNGFASFNNDDDYCNQYVDRDALMYLILANEILHVQHPNIITIAEDATYYPGLCEPVSQGGGLGFDYYVNL
 SASEMWVSLLDNVPDNEWSMSKIVSTLVANKEYADKMLSYAENHNQSSISGGRSFAEILFGGVDNGSPGGK
 ELLDRGISLHKMIRLITFTSGGRAYLNFMGNEFGHPERVEFPTQSNNFSFSLANRRWDLLESVHHHLFSFD
 KELMDLDKSKGILSRGLPSIHVN DANMVISFSRGPFLFIFNFHPSNSYEKYDVGVVEEAGEYTMILNSDEVK
 YGGQGIVTEDHYLQRSISKRIDGQRNCLEVLPSRTAQVYKLTRILRI

SEQ ID NO: 48

ATGGTCTCTTTGTGCAATCAGACTAGATTTTCTTTCCATCCGAATAACCTGGTCGTGAGTGAGAAACGA
 CGTTTAGGAATCTCGGGCGTTAACTTCCCTCGAAAAGATTAAATTAATAATTACATGCTTTGCAGCGGA
 GAGACCGCGCCAAGAAAAGCAGAAAGAAAAGTCACAATCTCAAAGCACGTCCGATGCGGAAGCAGG
 AGTAGACCCGGTGGGCTTTTAAACACGCTTGGGCATAGCGGATAGGATTTTGCACAATTTCTTAGAG
 AAAGACATAAGGCTCTTAAAGACCTTAAGGACGAAATATTTAAACGGCATTTTGATTTTCGGGATTTT
 GCATCAGGCTTCGAACTGTTAGGAATGCACAGGCATATGGAGCATCGGGTTGATTTTATGGATTGGGG
 ACCGGGATCACGGTACGGCGCAATTATTGGTGATTTTAAACGGATGGTCTCCAACGGAGAATGCTGCGC
 GCGAAGGCCTCTTTGGCCATGATGACTATGGTTATTGGTTTATTATACTTGAAGATAAATTGAGAGAA
 GGAGAGGAGCCGGACGAGTTGTATTTTCAACAATATAACTATGTTGATGACTATGATAAAGGTGACTC
 AGGCGTGTCCGGCTGAAGAAATTTTCAAAAAGCAAATGATGAATATTGGGAGCCGGGTGAAGATAGG
 TTTATCAAAAATAGATTGTGAAGTGCCGGCTAAATTATATGAGCAAATGTTTGGAACCGAATTCACCGCA
 AACACTGGAGGAATTAGGTGATATCCCTGACGCGGAAAACAAGATACAAGCAGTGGAAGAAGAGCAT
 AAAGATGATCCTCCATCTAACCTGCCGCCTTGCGATATTATTGATAAAGGTCAAGGCAAACCGTATGA
 TATCTTTAATGTTGTACGTCCCCAGAATGGACAAAAAAATTTTATGAGAAAGAACCCCCGATCCCAT
 ATTGCTGAGACACGTAAAGGCAGAAAGGCGTGGCTCCAGAAATATATCCCGGCCGTCCCCACGG
 CTCCAAGTACCGCTTATACTTCAATACCCCGGATGGACCATTAGAAAGGGTTCGGGCTTGGGCGACCT
 ACGTACAGCCAGAAGATGAAGGCAAAACAGGCCTATGCTATTTCATTGGGAACCGAGCCCGAGGGTGC
 CTATAAATGGAAGTACTCAAAACCAAAAGTACCAGAATCTTTACGGATTTATGAATGCCATGTGGGGA
 TTAGCGGAAGCGAACCGAAAGTAAGCACTTTTGAAGAGTTTACAAAAAAGGTGCTGCCGCACGTCAA
 ACGAGCGGGATATAACGCGATCCAGTTGATCGGCGTGCCTGAGCATAAGGATTATTTTACGGTGGGT
 ATAGAGTGACCACTTTTTCGAGCATCTTCCCGCTACGGTACTCCTGACGATTTTAAAGACTTGTGG
 ATGAAGCTCATGGGCTGGGTCTCTGGTCTTTCTGGATATTGTCCACTCATATGCTGCGGCGGATCAGA
 TGGTTGGGCTGAGCTTGTGACGGTTCCAACGATTGCTACTTCCACTATGGGAAACGTGGCCATCATA
 AACATTGGGGGACCAGAATGTTCAAATATGGCGACCTTGATGTGCTTCACTTTTAAATTTCAAACCTAA
 ATTGGTGGATTACAGAATACCAGGTTGACGGATATCAATTCCACAGCTTGGCATCGATGATCTATACA
 CATAACGGGTTTCGAAGTTTCAATAATGATTGGACGATTATTGCAATCAGTATGTAGACCGGGATGC
 CCTATGTACCTGATTCTTGCGAACGAAATCCTTCATGTTTCAGCATCCGAACATCATTACTATTGCGGA
 AGATGCAACATACTACCCGGGCTGTGCGAACAGTTTCCCAAGGCGGACTGGGATTTGATTATTATG
 TCAATCTGTCTGCATCCGAAATGTGGGTTAGCTTACTTGATAACGTGCCTGATAATGAGTGGTCAATGT
 CAAAAATCGTTAGCACACTCGTTGCAAATAAAGATAACGCTGACAAAATGTTATCCTACGCAGAAAAAT
 CATAATCAGAGTATAAGCGGTGGTGGTCAATTTGCCGAGATCTTATTTGGCGGCGTGGACAACGGCAG
 TCCGCGCGGTAAAGAGTTATTAGATAGAGGTATCAGCCTTCACAAGATGATTGCGCTGATTACTTTTAC
 AAGTGGCGGAAGAGCTTATTTGAACTTCATGGGAAACGAGTTCGGACATCCTGAAAGAGTAGAATTCC
 CAACACAATCCAACAATTTCTCGTTCAGCTTAGCAAAACCGCGATGGGATCTGCTGGAGAGTGGGGTA
 CATCATCATTTGTTTTGTTTTGATAAAGAGCTGATGGACCTGGATAAAAGCAAAGGAATTCTGAGCCG
 AGGACTTCCGAGCATCCATCATGTTAATGACGCAAATATGGTCATTAGCTTTTCCCGGGGCCCTTCTCT
 TTTTATCTTTAATTTTCAACCTCAAAATTCATATGAAAAGTATGATGTGCGGCGTCGAGGAGGCGGGTGA
 GTATACCATGATCCTCAACTCAGATGAGGTGAAATATGGTGGCCAAGGGATAGTAACCGAAGATCATT
 ATTTACAAAGGTCTATAAGCAAGAGGATCGATGGACAGAGAACTGCTTGGAGGTCTTTTGGCCAGT
 AGGACAGCTCAGGTGTACAAGCTTACACGAATCCTTCGCATTTGA

SEQ ID NO: 49

MKILFAVSECTPFVKSGGLADVAGALPKALARLGNEVAVMLPKYSQIPEPWKKRMKKQAECTVAVGWRO
QYCGIEHMAENDVNYFYFIDNEYFNDRSLYGHYDDGERFAFFSRAVLEAAKVVNVQADIVHTHDWHATA
MVNYLLKEEYRKHPFYERMKSVLTIHNLQFQGIFFPDVTHDLLGLEMDHFHYERLECNGFVNFMKAGIAA
DHVTTVSPTYRNEIMTPYYGEQLEQVLQYREDDVTGILNGIDDTFYQPKSDPYIEAQYDSGDLACKLENKT
KLQQRMGLEPEKNLISMVTRLTKQKGLDLVRRIMHELLEEQDIQLVVLGTGEREFEDYFRYAEFAFHEK
CRAIYIGFDEPLAHQIYAGSDMFLMPSKFEPCLGLQILALQYGAIPVRETGGLYDTPVRAYQEEEGTGNGFTF
SAFNAHDLKFTIERALSFYCQQDVWKSIVKTAMNADYSWGKSAKEYQRIFEQVTRSGRDVLE

SEQ ID NO: 50

ATGAAGATTTTGTTCGCGGTTAGCGAGTGCACACCTTTTCGTAAAAATCCGGCGGATTAGCGGACGTTGC
GGGTGCTTTACCGAAAGCCTTAGCGCGCCTTGGAATGAAGTCGCTGTGATGCTGCCGAAATATAGTC
AAATTCGGAACCGTGGAAGAAAAGAATGAAAAACAGGCAGAATGCACAGTTGCGGTTCGGCTGGCG
CCAACAGTACTGCGGAATCGAACATATGGCTGAGAATGACGTGAACTATTATTTATAGATAACGAAT
ATTATTTTAACAGAGATTCTCTGTATGGACACTATGACGATGGAGAGAGGTTTTCGTTTTTATAGCCGGG
CTGTGCTCGAAGCCGCGAAAGTCGTGAATGTGCAGGCTGATATCGTTCATACGCATGACTGGCATAACC
GCGATGGTCAACTATTTGCTGAAAGAAGAATATCGGAAACATCCGTTTTATGAGCGCATGAAAAGCGT
TCTTACGATTCTAATCTCCAATTCAGGGTATCTTTCCACCCGATGTCACACATGACCTTTTAGGCTTA
GAAATGGATCATTTCATTACGAACGTTTGGAATGCAACGTTTCGTGAATTTATGAAGGCTGGAAT
CATCGCCGCAGATCATGTGACTACGGTCTCTCTACGTATCGTAATGAAATAATGACGCCATATTATGG
TGAACAGCTGGAGCAGGTTCTTCAGTATCGCGAAGATGATGTCACGGGAATTCTGAACGGCATTGATG
ACACGTTCTACCAACCTAAATCAGACCCATATATTGAAGCGCAGTACGATAGTGGCGATCTTGCCTGC
AAATTAGAAAATAAAACAAAGCTGCAACAACGCATGGGATTACCAGAGAAGAATGATATCCCGTTAA
TTTCAATGGTAACAGACTTACGAAGCAGAGGGGCTGGATTTGGTCAGACGGATAATGCATGAACTT
TTGAAGAGCAGGATATCCAGCTGGTCTGTGCTGGCACCCGAGAGAAAGAGAATTTGAGGATTACTTTTCG
CTACGCTGAATTTGCGTTTTATGAGAAAGTCCCGCGCCTACATTGGCTTTGACGAACCCCTTAGCGCACC
GATTTACGCCGGATCAGATATGTTTCTCATGCCGAGCAAGTTTGAACCTTGTGGACTTGGCCAGCTGAT
TGCATTACAATACGGCGCCATTCCTATTGTACGGGAGACCGGAGGCCTGTATGACACAGTGCAGCCT
ATCAGGAAGAAGAAGGTACAGGCAATGGCTTTACTTTTATGTGCGTTTAAATGCACATGATCTGAAATTC
ACAATAGAAAGAGCTTTAAGCTTTTATTGTCAACAGGATGTCTGGAAATCAATCGTAAAGACCGCTAT
GAATGCCGATTATTCATGGGGTAAATCTGCAAAAGAGTACCAACGTATCTTCGAACAAGTGACACGGT
CTGGGCGCGACGTCCTTGAATAA

SEQ ID NO: 51

MALKRGLSGVNRIRSGGGSRSVLVLLIFFCVFAPLCFFVGRGVYIDSSNDYSIVSVKQNLDWRERLAMQS
VRSLSFKEILDVIATSTADLGPLSLDSFKKNNLSASWRGTGVDPSFRHSENPATPDVKSNNLNEKRDSISKDS
IHQKVETPTKIHRRQLREKRREMRANELVQHNDTILKLENAIERSKSVDSAFLGKYSIWRENENDNSD
SNIRLMRDQVIMARVYSGIAKLKNKNDLLQELQARLKDSQRVLGEATSDADLPRSAHEKLRAMGQVLAK
AKMQLYDCKLVTGKLRLMLQTADQVRSLLKQSTFLAQLAAKTIPNPIHCLSMRLTIDYYLLSPEKRFPR
SENLENPNLYHYALFSDNVLAASVVVNSTIMNAKDPKSHVFHLVTDKLNFGAMNMWFLNPPGKATIHVE
NVDEFKWLNSSYCPVLRQLESAAMREYYFKADHPTSGSSNLKYRNPKYLSMLNHLRFYLPVYPKLNKIL
FLDDDIIVQKDLTPLWEVNLNGKVNGAVETCGESFHRFDKYLNFNPHIARNFNPACGWAYGMNMFDL
KEWKKRDITGIYHKWQNMNENRTLWKLGLTPPLITFYGLTHPLNKAWHVLGLGYNPSIDKKDIENAAVV
HYNGNMKPWLELAMSKYRPYWTYIKFDHPYLRRCNLHE

SEQ ID NO: 52

ATGGCCCTTAAGAGGGGGCTGAGTGGAGTGAACCGTATCAGAGGATCAGGAGGAGGCAGCCGTTTCAG
TCCTGGTTCTTCTTATCTTTTTTTCGCTGTTTCGACCGTTATGTTTCTTCGTTGGTCGCGAGTCTATATA
GACAGCTCGAACGACTACTCAATTGTTAGTGTAAGCAGAATTTAGACTGGAGAGAGCGCCTGGCAAT
GCAATCTGTAAGATCATTATTCTCGAAGGAAATTTAGATGTGATAGCCACGTCTACGGCAGACTTGG
GGCCGTTATCTTTAGATAGCTTCAAAAAGAACAAATTTATCAGCTAGCTGGAGAGGCACGGGCGTTGAC
CCTAGCTTTTCGCCACAGCGAAAACCCCTGCGACCCCGGATGTAAAGTCAAACAATTTAAATGAAAAAG
AGATTCTATTTCTAAAGATTCCATCCATCAGAAGGTTGAGACACCGACTAAAATCCACCGACGACAAC
TTCGTGAGAAAAGGAGAGAGATGCGCGCGAATGAACTTGTCCAGCACACGATGATACAATTTTAAA
GTTGGAGAATGCAGCCATTGAGCGTAGTAAAGTGATGATTCTGCAGTATTAGGAAAAACTCAATTT
GGCGCCGGGAAAATGAAAATGACAATAGCGATTCTAATATCAGGCTGATGAGGGACCAGGTCATTAT

GGCACGCGTATATTCGGGCATTGCAAAATTAACAAAAACAAAAACGACTTACTTCAAGAACTTCAGGCGC
 GGCTTAAGGATTCCCAGAGAGTGTGGGAGAAAGCTACATCCGATGCCGATTTGCCAAGGAGCGCGCA
 CGAGAACTTAGAGCGATGGGTCAAGTTTTGGCTAAAGCGAAAAATGCAGCTGTATGATTGCAAATTAG
 TCACTGGTAAACTGAGAGCAATGTTGCAGACAGCAGATGAGCAGGTACGAAGCTTAAAAAACAATC
 AACGTTTCTTGCACAGCTTGCAGCTAAACAATACCCAATCCAATCCATTGTTATCTATGAGATTGAC
 CATAGATTATTACCTTTTGTCTCCCGAGAAGAGGAAGTTTCCGCGATCTGAAAAATCTGGAAAAACCGTA
 ATCTGTATCATTACGCTCTGTTCTCAGACAATGTTTTAGCAGCATCTGTTGTCGTAAATAGCACAATCA
 TGAATGCGAAGGACCCGTCTAAACATGTGTTCCATCTTGTACGGATAAACTTAACTTTGGCGCCATG
 AATATGTGGTTTCTTCTGAATCCTCCGGGCAAAGCCACAATCCATGTGGAAAAATGTCGACGAATTTAA
 ATGGCTTAACTCCTCCTATTGCCCGGTTCTTCGGCAGTTAGAATCAGCTGCTATGAGAGAATACTATTT
 TAAAGCGGATCATCTACGTCCGGCAGCAGCAATTTAAAATACAGAAATCCGAAGTACTTATCTATGCT
 TTAACCATTTAAGATTTTATTTACCGGAAGTTTATCCGAAACTTAATAAAATCCTGTTTCTCGACGATG
 ACATTATTGTTCAAAAAAGATTTGACTCCGCTTTGGGAAGTGAACCTTAAATGGCAAAGTCAATGGGGCC
 GTGGAGACTTGGGGGAGTCATTTTCATAGATTTGACAAGTACTTAACTTTTCTAACCCGCACATTGCA
 CGTAATTTCAACCCAAATGCATGTGGATGGGCGTATGGTATGAATATGTTTGAATTTAAAGAATGGAA
 AAAACGCGATATAACAGGAATTTATCATAAGTGGCAAAACATGAACGAAAACAGGACACTGTGGAAA
 CTGGGAACGTTGCCGCTGGACTTATTACGTTTTATGGGTTGACGCACCCGCTTAATAAAGCCTGGCAT
 GTTTTAGGACTGGGATATAATCCTAGCATTGACAAGAAAGATATTGAAAACGCGGCGGTAGTGCATTA
 TAATGGGAATATGAAACCGTGGCTTGAAGTGGCTATGTCTAAATATAGACCGTACTGGACGAAATACA
 TTAATTTGACCATCCGTATTTACGACGGTGCAACCTGCATGAATAG

SEQ ID NO: 53

MSVVDVIGLLATAAYVTLASAYKVVQFINVSSVTDVAGLES DALPLTPRVDVIVPTFNENSSTLLECVASIC
 AQDYRGPIVTVDDGSTNKTFSHAVCDKYASDERFIFVELDQNK GKRAAQMEAIRRTDGDILNVDSDTV
 IDKDVVTKLASSMRAPNVGGVMGQLVAKNRERSWLTRLIDMEYWLACNEERIAQSRFGSVGCCGPCAM
 YRRSAITPLLAIEYHQTF LGRPSNFGEDRHL TILMLKAGFR TGYVPGAVARTL VPDGLAPYLRQQLRWARS
 TYRDTALALRIKKNL SKYITFEICAQNLGTALLVMTMISLSL TTSQSQTPVILGVVVGMSIIRCCSV ALIAK
 DFRFLYFIVHSALNVLIL TPLKLYALLTIRDSRWLSRESS

SEQ ID NO: 54

ATGTCAGTGGTCGACGTAATTGGCCTTCTGGCCACCGCCGCTATGTGACCCTGGCTTCCGCGTACAAA
 GTAGTCCAATTTATCAATGTCTCTAGCGTGACTGACGTAGCAGGCCCTTGAGAGTGATGCCCTCCCGCTT
 ACACCGCGCGTAGATGTAATCGTGCCACGTTTAAATGAAAATTCTTCCACGCTGTAGAAATGCGTCGC
 AAGCATTGTGCGCAAGACTACAGAGGCCCAATCACCATTGTGGTGGTAGACGACGGCTCGACTAATA
 AAAGTAGCTTTACGCGAGTTTGCACAAATATGCGTCGGATGAGCGCTTCATCTTCGTCGAACTTGATC
 AGAACAAAGGGGAAACGCGCTGCACAGATGGAGGCAATTCGCCGACGGACGGCGATCTTATCTTGAA
 TGTGATTTCTGATACCGTGATCGACAAGGATGTTGTTACGAAACTTGCTTCGAGCATGCGCGCACCAA
 ACGTGGGTGGAGTGATGGGGCAACTGGTTGCCAAAAATCGGGAACGATCATGGCTCACACGGCTTATC
 GATATGGAATATTGGTTAGCTTGTAAATGAGGAGCGCATCGCTCAAAGCCGCTTTGGTTCTGTCATGTGC
 TGTGTGGTCCATGCGCGATGTACAGACGTTCTGCAATCACGCCACTTCTGGCGGAATACGAACACCA
 GACATTCCTTGACGACCGTCTAATTTTGGCGAAGATCGGCACCTTAACGATTTTAAATGTTGAAAGCGG
 GATTTTCGTACGGGGTATGTCCAGGCGCAGTAGCTCGTACGTTAGTTCTGACGGATTGGCTCCTTATT
 TGCGTCAACAACTTCGGTGGGCGCGAGCACATACCGTGATACAGCTCTGGCATTGCGTATTAAAAAG
 AACTTATCAAAAATATATCACATTTGAAATCTGCGCGCAAAATCTGGGTACGGCGCTTTTGCTTGTGATG
 ACAATGATTTTCGTTATCTCTTACCACATCGGGAAGTCAAACCCCGGTAATCATTCTTGGAGTTGTGCTC
 GGCATGTCAATCATTAGATGTTGTTTCGGTTGCCCTTATCGCTAAAGATTTTAGATTTCTCTATTTATTG
 TACATTCGGCGCTTAATGTACTGATTCTCACTCCGCTTAAATTGTACGCTCTTCTTACCATTAGAGATA
 GCAGATGGTTATCGCGCAATCAAGCTAG

SEQ ID NOs: 55-70: (see Table 2)

SEQ ID NOs: 71-82: (see Example 4)

CLAIMS

What is claimed is:

1. A method of decreasing the amount or concentration of a carbohydrate in the gut of a subject comprising administering to the subject a bacterium engineered to express a heterologous nucleic acid sequence encoding a fiber-synthesizing enzyme, wherein the fiber-synthesizing enzyme synthesizes fiber from a carbohydrate substrate.
2. A method of increasing the amount or concentration of a fiber in the gut of a subject comprising administering to the subject a bacterium engineered to express a heterologous nucleic acid sequence encoding a fiber-synthesizing enzyme, wherein the fiber-synthesizing enzyme synthesizes fiber from a carbohydrate substrate.
3. A method of treating a subject in need of decreased amount or concentration of a carbohydrate in the gut or increased amount or concentration of a fiber in the gut comprising administering to the subject a bacterium engineered to express a heterologous nucleic acid sequence encoding a fiber-synthesizing enzyme, wherein the fiber-synthesizing enzyme synthesizes fiber from a carbohydrate substrate.
4. The method of claim 3, wherein the subject is suffering from a condition positively correlated or associated with consumption of carbohydrate.
5. The method of claim 3, wherein the subject is suffering from a condition negatively correlated or associated with consumption of fiber.
6. The method of any one of claims 3-5, wherein the subject is suffering from a condition selected from cardiovascular disease, heart disease, high blood pressure, high blood cholesterol, high blood glucose, diabetes, obesity, dysbiosis of the gut, inflammatory bowel disease, irritable bowel syndrome (IBS), diverticulitis, colorectal cancer, intestinal cancer, bloating, cramping, gas, hemorrhoids, and diarrhea.

7. The method of any one of claims 1-6, wherein the fiber is a soluble fiber.
8. The method of any one of claim 1-7, wherein the synthesized fiber comprises glucose, fructose, galacturonic acid, N-acetyl-D-glucosamine, and/or galactose monomers.
9. The method of any one of claims 1-8, wherein the carbohydrate substrate is selected from one or more of UDP-glucose, UDP-galactose, UDP-fucose, ADP- α -D-glucose, UDP-galacturonic acid, UDP-N-acetyl-alpha-D-glucosamine, galactinol, lactose, glucose, and sucrose and/or wherein the carbohydrate decreased in amount or concentration in the gut is selected from one or more of glucose, galactose, sucrose, fructose, and/or lactose.
10. The method of any one of claim 1-9, wherein the synthesized fiber is a fructooligosaccharide and/or is selected from laminaribiose, callose, curdlan, oat beta-glucan, laminarin, pleuran, lentinan, yeast beta glucan, trehalulose, trehalose, inulin, kestose, nystose, levan, raffinose, stachyose, verbascose, globotriose, human milk oligosaccharides (HMOs), cellobiose, cellulose, microcellulose, cotton, maltose, amylose, starch, glycogen, amylopectin, pectin, chitin.
11. The method of any one of claims 1-10, wherein the fiber-synthesizing enzyme is selected from a sucrase, an inulosucrase, a levansucrase, a 1,3-beta-glucan synthase, a 1,3;1,4-beta-D-glucan synthase, a 1,6-beta-glucan synthase, a sucrose isomerase, a 1,6-alpha-galactosyltransferase, a trehalulose synthase, a trehalose-6-phosphate synthase, an α -1,4-galactosyltransferase, an alpha-1,2-fucosyltransferase, a beta-galactosidase, a b-D-Galactosidase, a cellulose synthase, a maltose synthase, a starch synthase, a starch-branching enzyme, a glycogen synthase, a galacturonosyltransferase, and a chitinoligosaccharide synthase.
12. The method of any one of claims 1-11, wherein the fiber-synthesizing enzyme is operatively linked with a secretion polypeptide.

13. The method of any one of claims 1-12, wherein the bacterium is a spore-forming bacterium and/or is in a spore form.
14. The method of any one of claims 1-13, wherein the bacterium is a probiotic bacterium.
15. The method of any one of claims 1-14, wherein the bacterium is of a genus selected from *Bacillus*, *Bifidobacterium*, *Enterococcus*, *Escherichia*, *Lactobacillus*, *Lactococcus*, *Leuconostoc*, *Pediococcus*, and *Streptococcus*, optionally wherein the bacterium is of the genus *Bacillus*, optionally wherein the bacterium is of the species *B. subtilis*.
16. The method of any one of claims 1-15, wherein the bacterium is of a strain characterized in that it does not colonize the gut.
17. The method of any one of claims 1-16, wherein the method comprises administering the engineered bacterium to a subject that has consumed carbohydrate and/or the carbohydrate substrate within a preceding period, wherein the preceding period is a period of 24 hours 12 hours, 6 hours, 3 hours, or 1 hour.
18. The method of any one of claims 1-17, wherein the method comprises administering the engineered bacterium to a subject that has not consumed carbohydrate and/or the carbohydrate substrate.
19. The method of any one of claims 1-17, wherein the method comprises administering the engineered bacterium to a subject that has not consumed carbohydrate and/or the carbohydrate substrate within a preceding period, wherein the preceding period is a period of 24 hours 12 hours, 6 hours, 3 hours, or 1 hour.
20. The method of any one of claims 1-19, wherein the subject consumes carbohydrate and/or the carbohydrate substrate during a period subsequent to administration of the engineered

bacterium, wherein the subsequent period is a period of 24 hours 12 hours, 6 hours, 3 hours, or 1 hour.

21. The method of any one of claims 1-20, wherein the method prevents accumulation of sugar in the colon of the subject.

22. The method of any one of claims 1-21, wherein fiber-synthesizing enzymes and/or enzymes encoded by heterologous nucleic acid sequences comprised by the engineered bacterium consist of the fiber-synthesizing enzyme.

23. The method of any one of claims 1-22, wherein the administration comprises oral administration of a composition comprising the engineered bacterium.

24. The method of any one of claims 1-23, wherein the administration comprises administration of about 10^4 to about 10^{12} colony forming units of the engineered bacterium.

25. The method of claim any one of claims 1-24, wherein the nucleic acid sequence encoding the expression product is operatively linked with a constitutive promoter.

26. The method of claim any one of claims 1-25, wherein the nucleic acid sequence encoding the expression product is operatively linked with a flagellin gene promoter.

27. The method of claim 26, wherein the flagellin gene promoter comprises a mutation in a CsrA binding site, wherein the mutation in the CsrA binding site inhibits binding of CsrA to mRNA transcripts encoding the fiber-synthesizing enzyme but does not preclude expression of the fiber-synthesizing enzyme.

28. The method of claim 26 or claim 27, wherein the engineered bacterium comprises a mutation of an endogenous *flgM* gene that reduces inhibition of a sigma factor by FlgM.

29. A bacterium engineered to express a heterologous nucleic acid sequence encoding a fiber-synthesizing enzyme, wherein the fiber-synthesizing enzyme synthesizes fiber from a carbohydrate substrate.
30. The engineered bacterium of claim 29, wherein the fiber is a soluble fiber.
31. The engineered bacterium of claim 29 or claim 30, wherein the synthesized fiber comprises glucose, fructose, galacturonic acid, N-acetyl-D-glucosamine, and/or galactose monomers.
32. The engineered bacterium of any one of claims 29-31, wherein the carbohydrate substrate is selected from one or more of UDP-glucose, UDP-galactose, UDP-fucose, ADP- α -D-glucose, UDP-galacturonic acid, UDP-N-acetyl- α -D-glucosamine, galactinol, lactose, glucose, and sucrose.
33. The engineered bacterium of any one of claim 29-32, wherein the synthesized fiber is a fructooligosaccharide and/or is selected from laminaribiose, callose, curdlan, oat beta-glucan, laminarin, pleuran, lentinan, yeast beta glucan, trehalulose, trehalose, inulin, kestose, nystose, levan, raffinose, stachyose, verbascose, globotriose, human milk oligosaccharides (HMOs), cellobiose, cellulose, microcellulose, cotton, maltose, amylose, starch, glycogen, amylopectin, pectin, chitin.
34. The engineered bacterium of any one of claims 29-33, wherein the fiber-synthesizing enzyme is selected from a sucrase, an inulosucrase, a levansucrase, a 1,3-beta-glucan synthase, a 1,3;1,4-beta-D-glucan synthase, a 1,6-beta-glucan synthase, a sucrose isomerase, a 1,6-alpha-galactosyltransferase, a trehalulose synthase, a trehalose-6-phosphate synthase, an α -1,4-galactosyltransferase, an alpha-1,2-fucosyltransferase, a beta-galactosidase, a b-D-Galactosidase, a cellulose synthase, a maltose synthase, a starch synthase, a starch-branching enzyme, a glycogen synthase, a galacturonosyltransferase, and a chitinoligosaccharide synthase.

35. The engineered bacterium of any one of claims 29-34, wherein the fiber-synthesizing enzyme is operatively linked with a secretion polypeptide.
36. The engineered bacterium of any one of claims 29-35, wherein the bacterium is a spore-forming bacterium and/or is in a spore form.
37. The engineered bacterium of any one of claims 29-36, wherein the bacterium is a probiotic bacterium.
38. The engineered bacterium of any one of claims 29-37, wherein the bacterium is of a genus selected from *Bacillus*, *Bifidobacterium*, *Enterococcus*, *Escherichia*, *Lactobacillus*, *Lactococcus*, *Leuconostoc*, *Pediococcus*, and *Streptococcus*, optionally wherein the bacterium is of the genus *Bacillus*, optionally wherein the bacterium is of the species *B. subtilis*.
39. The engineered bacterium of any one of claims 29-38, wherein the bacterium is of a strain characterized in that it does not colonize the gut.
40. The engineered bacterium of any one of claims 29-39, wherein fiber-synthesizing enzymes and/or enzymes encoded by heterologous nucleic acid sequences comprised by the engineered bacterium consist of the fiber-synthesizing enzyme.
41. The engineered bacterium of claim any one of claims 29-40, wherein the nucleic acid sequence encoding the expression product is operatively linked with a constitutive promoter.
42. The engineered bacterium of claim any one of claims 29-41, wherein the nucleic acid sequence encoding the expression product is operatively linked with a flagellin gene promoter.
43. The engineered bacterium of claim 42, wherein the flagellin gene promoter comprises a mutation in a CsrA binding site, wherein the mutation in the CsrA binding site inhibits binding

of CsrA to mRNA transcripts encoding the fiber-synthesizing enzyme but does not preclude expression of the fiber-synthesizing enzyme.

44. The engineered bacterium of claim 42 or claim 43, wherein the engineered bacterium comprises a mutation of an endogenous *flgM* gene that reduces inhibition of a sigma factor by FlgM.

45. A composition comprising the engineered bacterium of any one of claims 29-44.

46. The composition of claim 45, wherein the composition is formulated for oral administration.

47. The formulation of claim 45 or 46, wherein the formulation comprises about 10^4 to about 10^{12} colony forming units of the engineered bacterium.

48. The composition of any one of claims 45-47, wherein the composition comprises a physiologically acceptable carrier.

49. The composition of claim 48, wherein the physiologically acceptable carrier is selected from a lactic acid fermented food, fermented dairy product, resistant starch, dietary fiber, carbohydrate, protein, glycosylated protein, water, capsule filler, and gummy material.

50. A bacterial cell culture comprising the engineered bacterium of any one of claims 29-44.

51. A method of decreasing the amount or concentration of a carbohydrate in the gut of a subject comprising administering to the subject a composition comprising a fiber-synthesizing enzyme, wherein the fiber-synthesizing enzyme synthesizes fiber from a carbohydrate substrate.

52. A method of increasing the amount or concentration of a fiber in the gut of a subject comprising administering to the subject a composition comprising a fiber-synthesizing enzyme, wherein the fiber-synthesizing enzyme synthesizes fiber from a carbohydrate substrate.

53. A method of treating a subject in need of decreased amount or concentration of a carbohydrate in the gut or increased amount or concentration of a fiber in the gut comprising administering to the subject a composition comprising a fiber-synthesizing enzyme, wherein the fiber-synthesizing enzyme synthesizes fiber from a carbohydrate substrate.

54. The method of claim 53, wherein the subject is suffering from a condition positively correlated or associated with consumption of carbohydrate.

55. The method of claim 53, wherein the subject is suffering from a condition negatively correlated or associated with consumption of fiber.

56. The method of any one of claims 53-55, wherein the subject is suffering from a condition selected from cardiovascular disease, heart disease, high blood pressure, high blood cholesterol, high blood glucose, diabetes, obesity, dysbiosis of the gut, inflammatory bowel disease, irritable bowel syndrome (IBS), diverticulitis, colorectal cancer, intestinal cancer, bloating, cramping, gas, hemorrhoids, and diarrhea.

57. The method of any one of claims 51-56, wherein the fiber is a soluble fiber.

58. The method of any one of claim 51-57, wherein the synthesized fiber comprises glucose, fructose, galacturonic acid, N-acetyl-D-glucosamine, and/or galactose monomers.

59. The method of any one of claims 51-58, wherein the carbohydrate substrate is selected from one or more of UDP-glucose, UDP-galactose, UDP-fucose, ADP- α -D-glucose, UDP-galacturonic acid, UDP-N-acetyl-alpha-D-glucosamine, galactinol, lactose, glucose, and sucrose

and/or wherein the carbohydrate decreased in amount or concentration in the gut is selected from one or more of glucose, galactose, sucrose, fructose, and/or lactose.

60. The method of any one of claim 51-59, wherein the synthesized fiber is a fructooligosaccharide and/or is selected from laminaribiose, callose, curdlan, oat beta-glucan, laminarin, pleuran, lentinan, yeast beta glucan, trehalulose, trehalose, inulin, kestose, nystose, levan, raffinose, stachyose, verbascose, globotriose, human milk oligosaccharides (HMOs), cellobiose, cellulose, microcellulose, cotton, maltose, amylose, starch, glycogen, amylopectin, pectin, chitin.

61. The method of any one of claims 51-60, wherein the fiber-synthesizing enzyme is selected from a sucrase, an inulosucrase, a levansucrase, a 1,3-beta-glucan synthase, a 1,3;1,4-beta-D-glucan synthase, a 1,6-beta-glucan synthase, a sucrose isomerase, a 1,6-alpha-galactosyltransferase, a trehalulose synthase, a trehalose-6-phosphate synthase, an α -1,4-galactosyltransferase, an alpha-1,2-fucosyltransferase, a beta-galactosidase, a b-D-Galactosidase, a cellulose synthase, a maltose synthase, a starch synthase, a starch-branching enzyme, a glycogen synthase, a galacturonosyltransferase, and a chitinoligosaccharide synthase.

62. The method of any one of claims 51-61, wherein the method comprises administering the composition to a subject that has consumed carbohydrate and/or the carbohydrate substrate within a preceding period, wherein the preceding period is a period of 24 hours 12 hours, 6 hours, 3 hours, or 1 hour.

63. The method of any one of claims 51-62, wherein the method comprises administering the composition to a subject that has not consumed carbohydrate and/or the carbohydrate substrate.

64. The method of any one of claims 51-62, wherein the method comprises administering the composition to a subject that has not consumed carbohydrate and/or the carbohydrate substrate within a preceding period, wherein the preceding period is a period of 24 hours 12 hours, 6 hours, 3 hours, or 1 hour.

65. The method of any one of claims 51-64, wherein the subject consumes carbohydrate and/or the carbohydrate substrate during a period subsequent to administration of the composition, wherein the subsequent period is a period of 24 hours, 12 hours, 6 hours, 3 hours, or 1 hour.
66. The method of any one of claims 51-65, wherein the method prevents accumulation of sugar in the colon of the subject.
67. The method of any one of claims 51-66, wherein fiber-synthesizing enzymes comprising the composition and/or enzymes comprising the composition consist of the fiber-synthesizing enzyme.
68. The method of any one of claims 51-67, wherein the administration comprises oral administration of the composition.
69. A composition comprising a fiber-synthesizing enzyme, wherein the fiber-synthesizing enzyme synthesizes fiber from a carbohydrate substrate and wherein the composition is formulated for oral administration.
70. The composition of claim 69, wherein the fiber is a soluble fiber.
71. The composition of claim 69 or claim 70, wherein the synthesized fiber comprises glucose, fructose, galacturonic acid, N-acetyl-D-glucosamine, and/or galactose monomers.
72. The composition of any one of claims 69-71, wherein the carbohydrate substrate is selected from one or more of UDP-glucose, UDP-galactose, UDP-fucose, ADP- α -D-glucose, UDP-galacturonic acid, UDP-N-acetyl- α -D-glucosamine, galactinol, lactose, glucose, and sucrose.

73. The composition of any one of claim 69-72, wherein the synthesized fiber is a fructooligosaccharide and/or is selected from laminaribiose, callose, curdlan, oat beta-glucan, laminarin, pleuran, lentinan, yeast beta glucan, trehalulose, trehalose, inulin, kestose, nystose, levan, raffinose, stachyose, verbascose, globotriose, human milk oligosaccharides (HMOs), cellobiose, cellulose, microcellulose, cotton, maltose, amylose, starch, glycogen, amylopectin, pectin, chitin.

74. The composition of any one of claims 69-73, wherein the fiber-synthesizing enzyme is selected from a sucrase, an inulosucrase, a levansucrase, a 1,3-beta-glucan synthase, a 1,3;1,4-beta-D-glucan synthase, a 1,6-beta-glucan synthase, a sucrose isomerase, a 1,6-alpha-galactosyltransferase, a trehalulose synthase, a trehalose-6-phosphate synthase, an α -1,4-galactosyltransferase, an alpha-1,2-fucosyltransferase, a beta-galactosidase, a b-D-Galactosidase, a cellulose synthase, a maltose synthase, a starch synthase, a starch-branching enzyme, a glycogen synthase, a galacturonosyltransferase, and a chitinoligosaccharide synthase.

75. The composition of any one of claims 69-74, wherein fiber-synthesizing enzymes comprising the composition and/or enzymes comprising the composition consist of the fiber-synthesizing enzyme.

76. The composition of any one of claims 69-75, wherein the composition comprises a physiologically acceptable carrier.

77. The composition of claim 76, wherein the physiologically acceptable carrier is selected from a lactic acid fermented food, fermented dairy product, resistant starch, dietary fiber, carbohydrate, protein, glycosylated protein, water, capsule filler, and gummy material.