

(12) STANDARD PATENT
(19) AUSTRALIAN PATENT OFFICE

(11) Application No. **AU 2015259088 B2**

(54) Title
Alpha (1,2) fucosyltransferase syngenes for use in the production of fucosylated oligosaccharides

(51) International Patent Classification(s)
C12N 1/21 (2006.01) **C12P 19/18** (2006.01)
C12N 15/63 (2006.01)

(21) Application No: **2015259088** (22) Date of Filing: **2015.05.14**

(87) WIPO No: **WO15/175801**

(30) Priority Data

(31)	Number	(32)	Date	(33)	Country
	61/993,742		2014.05.15		US

(43) Publication Date: **2015.11.19**

(44) Accepted Journal Date: **2021.11.11**

(71) Applicant(s)
Glycosyn LLC

(72) Inventor(s)
McCoy, John M.;Heidtman, Matthew Ian;Merighi, Massimo

(74) Agent / Attorney
Griffith Hack, Level 10 161 Collins St, MELBOURNE, VIC, 3000, AU

(56) Related Art
Glycobiology (Epub 2013-Nov-17) vol.24, no.2, p.170-178
US 20140031541 A1
AC B8GDY9, SubName: Full=Glycosyl transferase family 11 [Methanosphaerula
palustris (strain ATCC BAA
US 7521212 B1
WO 2015150328 A1



(51) International Patent Classification:

C12N 1/21 (2006.01) *C12P 19/18* (2006.01)
C12N 15/63 (2006.01)

(21) International Application Number:

PCT/US2015/030823

(22) International Filing Date:

14 May 2015 (14.05.2015)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

61/993,742 15 May 2014 (15.05.2014) US

(71) Applicant: GLYCOSYN LLC [US/US]; 890 Winter Street, Suite 208, Waltham, MA 02451 (US).

(72) Inventors: MCCOY, John, M.; 56 Howard Street, Reading, MA 01867 (US). HEIDTMAN, Matthew, Ian; 15 South Street, Apt. 2, Brighton, MA 02135 (US). MERIGHI, Massimo; 38 Benton Road, Apt. 2, Somerville, MA 02143 (US).

(74) Agents: BEATTIE, Ingrid, A. et al.; Mintz Levin Cohn Ferris Glovsky and Popeo, P.C., One Financial Center, Boston, MA 02111 (US).

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

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

Published:

— with international search report (Art. 21(3))

(54) Title: ALPHA (1,2) FUCOSYLTRANSFERASE SYNGENES FOR USE IN THE PRODUCTION OF FUCOSYLATED OLIGOSACCHARIDES

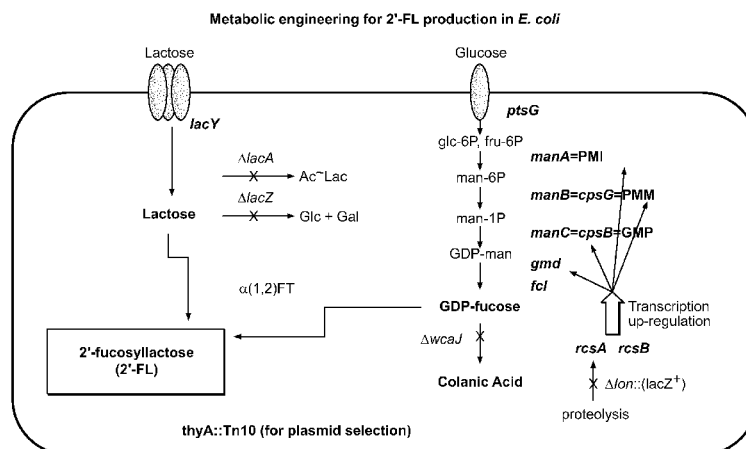


FIG. 2

(57) Abstract: The invention provides compositions and methods for engineering *E. coli* or other host production bacterial strains to produce fucosylated oligosaccharides, and the use thereof in the prevention or treatment of infection.

ALPHA (1,2) FUCOSYLTRANSFERASE SYNGENES FOR USE IN THE PRODUCTION OF FUCOSYLATED OLIGOSACCHARIDES

FIELD OF THE INVENTION

The invention provides compositions and methods for producing purified oligosaccharides, in particular certain fucosylated oligosaccharides that are typically found in human milk.

BACKGROUND OF THE INVENTION

Human milk contains a diverse and abundant set of neutral and acidic oligosaccharides. More than 130 different complex oligosaccharides have been identified in human milk, and their structural diversity and abundance is unique to humans. Although these molecules may not be utilized directly by infants for nutrition, they nevertheless serve critical roles in the establishment of a healthy gut microbiome, in the prevention of disease, and in immune function. Prior to the invention described herein, the ability to produce human milk oligosaccharides (HMOS) inexpensively was problematic. For example, their production through chemical synthesis was limited by stereo-specificity issues, precursor availability, product impurities, and high overall cost. As such, there is a pressing need for new strategies to inexpensively manufacture large quantities of HMOS.

It is to be understood that if any prior art publication is referred to herein, such reference does not constitute an admission that the publication forms a part of the common general knowledge in the art in Australia or any other country.

SUMMARY OF THE INVENTION

The invention features an efficient and economical method for producing fucosylated oligosaccharides. Such production of a fucosylated oligosaccharide is accomplished using an isolated nucleic acid comprising a sequence encoding a lactose-utilizing α (1,2) fucosyltransferase gene product (e.g., polypeptide or protein), which is operably linked to one or more heterologous control sequences that direct the production of the recombinant fucosyltransferase gene product in a host production bacterium such as *Escherichia coli* (*E. coli*).

A first aspect provides a method for producing a fucosylated oligosaccharide in a bacterium comprising providing bacterium comprising an exogenous lactose-utilizing α (1,2) fucosyltransferase enzyme, wherein the amino acid sequence of said enzyme

comprises at least 90% identity to FutW (SEQ ID NO: 17), FutP (SEQ ID NO: 11), FutO (SEQ ID NO: 10), FutQ (SEQ ID NO: 12), FutR (SEQ ID NO: 13), FutS (SEQ ID NO: 14), FutU (SEQ ID NO: 15), FutV (SEQ ID NO: 16), FutX (SEQ ID NO: 18), FutZ (SEQ ID NO: 20) or FutZA (SEQ ID NO: 21), wherein said fucosylated oligosaccharide comprises 2'-fucosyllactose (2'-FL), lactodifucotetraose (LDFT), Lacto-N-fucopentaose I (LNF I), or lacto-N-difucohexaose I (LDFH I).

A second aspect provides a nucleic acid construct comprising an isolated nucleic acid operably linked to one or more heterologous control sequences, wherein the nucleic acid encodes a lactose-utilizing $\alpha(1,2)$ fucosyltransferase enzyme for the production of said enzyme in a host bacteria production strain, wherein the amino acid sequence of said enzyme comprises at least 90% identity to FutW (SEQ ID NO: 17), FutP (SEQ ID NO: 11), FutO (SEQ ID NO: 10), FutQ (SEQ ID NO: 12), FutR (SEQ ID NO: 13), FutS (SEQ ID NO: 14), FutU (SEQ ID NO: 15), FutV (SEQ ID NO: 16), FutX (SEQ ID NO: 18), FutZ (SEQ ID NO: 20) or FutZA (SEQ ID NO: 21), wherein the lactose-utilizing $\alpha(1,2)$ fucosyltransferase enzyme is capable of producing a fucosylated oligosaccharide comprising 2'-fucosyllactose (2'-FL), lactodifucotetraose (LDFT), Lacto-N-fucopentaose I (LNF I), or lacto-N-difucohexaose I (LDFH I), and wherein said one or more heterologous control sequences direct the production of the enzyme in the bacterium.

A third aspect provides an isolated bacterium comprising an isolated nucleic acid encoding a lactose-accepting $\alpha(1,2)$ fucosyltransferase enzyme, wherein the amino acid sequence of said enzyme encoded by said nucleic acid comprises at least 90% identity to FutW (SEQ ID NO: 17), FutP (SEQ ID NO: 11), FutO (SEQ ID NO: 10), FutQ (SEQ ID NO: 12), FutR (SEQ ID NO: 13), FutS (SEQ ID NO: 14), FutU (SEQ ID NO: 15), FutV (SEQ ID NO: 16), FutX (SEQ ID NO: 18), FutZ (SEQ ID NO: 20) or FutZA (SEQ ID NO: 21), wherein the lactose-utilizing $\alpha(1,2)$ fucosyltransferase enzyme is capable of producing a fucosylated oligosaccharide comprising 2'-fucosyllactose (2'-FL), lactodifucotetraose (LDFT), Lacto-N-fucopentaose I (LNF I), or lacto-N-difucohexaose I (LDFH I).

Also disclosed is a novel $\alpha(1,2)$ fucosyltransferases (also referred to herein as $\alpha(1,2)$ FTs) that utilize lactose and catalyzes the transfer of an L-fucose sugar from a GDP-fucose donor substrate to an acceptor substrate in an α -1,2-linkage. In a preferred embodiment, the acceptor substrate is an oligosaccharide. The $\alpha(1,2)$ fucosyltransferases identified and described herein are useful for expressing in host bacterium for the production of human milk oligosaccharides (HMOS), such as fucosylated oligosaccharides. Exemplary fucosylated oligosaccharides produced by the methods described herein include 2'-fucosyllactose (2'FL), lactodifucotetraose (LDFT), lacto-N-fucopentaose I (LNF I), or lacto-N-difucohexaose I (LDFH I). The " $\alpha(1,2)$ fucosyltransferases" disclosed herein encompasses the amino acid

sequences of the $\alpha(1,2)$ fucosyltransferases and the nucleic acid sequences that encode the $\alpha(1,2)$ fucosyltransferases, as well as variants and fragments thereof that exhibit $\alpha(1,2)$ fucosyltransferase activity. Also within the disclosure is a nucleic acid construct comprising an isolated nucleic acid encoding a lactose-accepting $\alpha(1,2)$ fucosyltransferase enzyme, said nucleic acid being optionally operably linked to one or more heterologous control sequences that direct the production of the enzyme in a host bacteria production strain.

The amino acid sequence of the lactose-accepting $\alpha(1,2)$ fucosyltransferases described herein is at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95% identity to *Helicobacter pylori* 26695 alpha-(1,2) fucosyltransferase (futC or SEQ ID NO: 1). Preferably, the lactose-accepting $\alpha(1,2)$ fucosyltransferases described herein is at least 22% identical to *H. pylori* FutC, or SEQ ID NO: 1.

In another embodiment, the amino acid sequence of the lactose-accepting $\alpha(1,2)$ fucosyltransferases described herein is at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95% identity to *Bacteroides vulgatus* alpha-(1,2) fucosyltransferase (FutN or SEQ ID NO: 3). Preferably, the lactose-accepting $\alpha(1,2)$ fucosyltransferases described herein is at least 25% identical to *B. vulgatus* FutN, or SEQ ID NO: 3.

Alternatively, the exogenous $\alpha(1,2)$ fucosyltransferase preferably comprises at least at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95% identity to any one of the novel $\alpha(1,2)$ fucosyltransferases disclosed herein, for example, to the amino acid sequences in Table 1.

Exemplary $\alpha(1,2)$ fucosyltransferases include, but are not limited to, *Prevotella melaninogenica* FutO, *Clostridium bolteae* FutP, *Clostridium bolteae* +13 FutP, *Lachnospiraceae* sp. FutQ, *Methanosphaerula palustris* FutR, *Tannerella* sp. FutS, *Bacteroides caccae* FutU, *Butyrivibrio* FutV, *Prevotella* sp. FutW, *Parabacteroides johnsonii* FutX, *Akkermansia muciniphilia* FutY, *Salmonella enterica* FutZ, *Bacteroides* sp. FutZA. For example, the $\alpha(1,2)$ fucosyltransferases comprise the amino acid sequences comprising any one of the following: *Prevotella melaninogenica* FutO (SEQ ID NO: 10), *Clostridium bolteae* FutP (SEQ ID NO: 11), *Clostridium bolteae* +13 FutP (SEQ ID NO: 292), *Lachnospiraceae* sp. FutQ (SEQ ID NO: 12), *Methanosphaerula palustris* FutR (SEQ ID

NO: 13), *Tannerella* sp. FutS (SEQ ID NO: 14), *Bacteroides caccae* FutU (SEQ ID NO: 15), *Butyrivibrio* FutV (SEQ ID NO: 16), *Prevotella* sp. FutW (SEQ ID NO: 17), *Parabacteroides johnsonii* FutX (SEQ ID NO: 18), *Akkermansia muciniphilia* FutY (SEQ ID NO: 19), *Salmonella enterica* FutZ (SEQ ID NO: 20), and *Bacteroides* sp. FutZA (SEQ ID NO: 21), or a functional variant or fragment thereof. Other exemplary $\alpha(1,2)$ fucosyltransferases include any of the enzymes listed in Table 1, or functional variants or fragments thereof.

The present disclosure also relates to a method for producing a fucosylated oligosaccharide in a bacterium by providing bacterium that express at least one exogenous lactose-utilizing $\alpha(1,2)$ fucosyltransferase. The amino acid sequence of the exogenous lactose-utilizing $\alpha(1,2)$ fucosyltransferase is preferably at least 22% identical to *H. pylori* FutC or at least 25% identical to *B. vulgatus* FutN. In one embodiment, the bacterium also expresses one or more exogenous lactose-utilizing $\alpha(1,3)$ fucosyltransferase enzymes and/or one or more exogenous lactose-utilizing $\alpha(1,4)$ fucosyltransferase enzymes. The combination of fucosyltransferases expressed in the production bacterium is dependent upon the desired fucosylated oligosaccharide product. The method disclosed herein further includes retrieving the fucosylated oligosaccharide from said bacterium or from a culture supernatant of said bacterium.

Examples of suitable $\alpha(1,3)$ fucosyltransferase enzymes include, but are not limited to *Helicobacter pylori* 26695 *futA* gene (GenBank Accession Number HV532291 (GI:365791177), incorporated herein by reference), *H. hepaticus* Hh0072, *H. pylori* 11639 FutT, and *H. pylori* UA948 FucTa (e.g., GenBank Accession Number AF194963 (GI:28436396), incorporated herein by reference)(Rasko, D. A., Wang, G., Palcic, M. M. & Taylor, D. E. J Biol Chem 275, 4988-4994 (2000)). Examples of suitable $\alpha(1,4)$ fucosyltransferase enzymes include, but are not limited to *H. pylori* UA948 FucTa (which has relaxed acceptor specificity and is able to generate both $\alpha(1,3)$ - and $\alpha(1,4)$ -fucosyl linkages). An example of an enzyme possessing only $\alpha(1,4)$ fucosyltransferase activity is given by the FucT III enzyme from *Helicobacter pylori* strain DMS6709 (e.g., GenBank Accession Number AY450598.1 (GI:40646733), incorporated herein by reference) (S. Rabbani, V. Miksa, B. Wipf, B. Ernst, *Glycobiology* **15**, 1076-83 (2005).)

The disclosure also relates to a nucleic acid construct or a vector comprising a nucleic acid encoding at least one $\alpha(1,2)$ fucosyltransferase or variant, or fragment thereof, as described herein. The vector can further include one or more regulatory elements, e.g., a heterologous promoter. By “heterologous” is meant that the control sequence and protein-

encoding sequence originate from different bacterial strains. The regulatory elements can be operably linked to a gene encoding a protein, a gene construct encoding a fusion protein gene, or a series of genes linked in an operon in order to express the fusion protein. In yet another embodiment, the disclosure comprises an isolated recombinant cell, e.g., a bacterial cell containing an aforementioned nucleic acid molecule or vector. The nucleic acid is optionally integrated into the genome of the host bacterium. In some embodiments, the nucleic acid construct also further comprises one or more $\alpha(1,3)$ fucosyltransferases and/or $\alpha(1,4)$ fucosyltransferases. Alternatively, the $\alpha(1,2)$ fucosyltransferase also exhibits $\alpha(1,3)$ fucosyltransferase and/or $\alpha(1,4)$ fucosyltransferase activity.

The bacterium utilized in the production methods described herein is genetically engineered to increase the efficiency and yield of fucosylated oligosaccharide products. For example, the host production bacterium is characterized as having a reduced level of β -galactosidase activity, a defective colanic acid synthesis pathway, an inactivated ATP-dependent intracellular protease, an inactivated *lacA*, or a combination thereof. In one embodiment, the bacterium is characterized as having a reduced level of β -galactosidase activity, a defective colanic acid synthesis pathway, an inactivated ATP-dependent intracellular protease, and an inactivated *lacA*.

As used herein, an “inactivated” or “inactivation of a” gene, encoded gene product (*i.e.*, polypeptide), or pathway refers to reducing or eliminating the expression (*i.e.*, transcription or translation), protein level (*i.e.*, translation, rate of degradation), or enzymatic activity of the gene, gene product, or pathway. In the instance where a pathway is inactivated, preferably one enzyme or polypeptide in the pathway exhibits reduced or negligible activity. For example, the enzyme in the pathway is altered, deleted or mutated such that the product of the pathway is produced at low levels compared to a wild-type bacterium or an intact pathway. Alternatively, the product of the pathway is not produced. Inactivation of a gene is achieved by deletion or mutation of the gene or regulatory elements of the gene such that the gene is no longer transcribed or translated. Inactivation of a polypeptide can be achieved by deletion or mutation of the gene that encodes the gene product or mutation of the polypeptide to disrupt its activity. Inactivating mutations include additions, deletions or substitutions of one or more nucleotides or amino acids of a nucleic acid or amino acid sequence that results in the reduction or elimination of the expression or activity of the gene or polypeptide. In other embodiments, inactivation of a polypeptide is achieved through the addition of exogenous sequences (*i.e.*, tags) to the N or C-terminus of

the polypeptide such that the activity of the polypeptide is reduced or eliminated (*i.e.*, by steric hindrance).

A host bacterium suitable for the production systems described herein exhibits an enhanced or increased cytoplasmic or intracellular pool of lactose and/or GDP-fucose. For example, the bacterium is *E. coli* and endogenous *E. coli* metabolic pathways and genes are manipulated in ways that result in the generation of increased cytoplasmic concentrations of lactose and/or GDP-fucose, as compared to levels found in wild type *E. coli*. Preferably, the bacterium accumulates an increased intracellular lactose pool and an increased intracellular GDP-fucose pool. For example, the bacteria contain at least 10%, 20%, 50%, or 2X, 5X, 10X or more of the levels of intracellular lactose and/or intracellular GDP-fucose compared to a corresponding wild type bacteria that lacks the genetic modifications described herein.

Increased intracellular concentration of lactose in the host bacterium compared to wild-type bacterium is achieved by manipulation of genes and pathways involved in lactose import, export and catabolism. In particular, described herein are methods of increasing intracellular lactose levels in *E. coli* genetically engineered to produce a human milk oligosaccharide by simultaneous deletion of the endogenous β -galactosidase gene (*lacZ*) and the lactose operon repressor gene (*lacI*). During construction of this deletion, the *lacIq* promoter is placed immediately upstream of (contiguous with) the lactose permease gene, *lacY*, *i.e.*, the sequence of the *lacIq* promoter is directly upstream and adjacent to the start of the sequence encoding the *lacY* gene, such that the *lacY* gene is under transcriptional regulation by the *lacIq* promoter. The modified strain maintains its ability to transport lactose from the culture medium (via LacY), but is deleted for the wild-type chromosomal copy of the *lacZ* (encoding β -galactosidase) gene responsible for lactose catabolism. Thus, an intracellular lactose pool is created when the modified strain is cultured in the presence of exogenous lactose.

Another method for increasing the intracellular concentration of lactose in *E. coli* involves inactivation of the *lacA* gene. A inactivating mutation, null mutation, or deletion of *lacA* prevents the formation of intracellular acetyl-lactose, which not only removes this molecule as a contaminant from subsequent purifications, but also eliminates *E. coli*'s ability to export excess lactose from its cytoplasm (Danchin A. Cells need safety valves. *Bioessays* 2009, Jul;31(7):769-73.), thus greatly facilitating purposeful manipulations of the *E. coli* intracellular lactose pool.

The disclosure also relates to methods for increasing intracellular levels of GDP-fucose in a bacterium by manipulating the organism's endogenous colanic acid biosynthesis pathway. This increase is achieved through a number of genetic modifications of endogenous *E. coli* genes involved either directly in colanic acid precursor biosynthesis, or in overall control of the colanic acid synthetic regulon. Particularly preferred is inactivation of the genes or encoded polypeptides that act in the colanic acid synthesis pathway after the production of GDP-fucose (the donor substrate) and before the generation of colanic acid. Exemplary colanic acid synthesis genes include, but are not limited to: a *wcaJ* gene, (e.g., GenBank Accession Number (amino acid) BAA15900 (GI:1736749), incorporated herein by reference), a *wcaA* gene (e.g., GenBank Accession Number (amino acid) BAA15912.1 (GI:1736762), incorporated herein by reference), a *wcaC* gene (e.g., GenBank Accession Number (amino acid) BAE76574.1 (GI:85675203), incorporated herein by reference), a *wcaE* gene (e.g., GenBank Accession Number (amino acid) BAE76572.1 (GI:85675201), incorporated herein by reference), a *wcaI* gene (e.g., GenBank Accession Number (amino acid) BAA15906.1 (GI:1736756), incorporated herein by reference), a *wcaL* gene (e.g., GenBank Accession Number (amino acid) BAA15898.1 (GI:1736747), incorporated herein by reference), a *wcaB* gene (e.g., GenBank Accession Number (amino acid) BAA15911.1 (GI:1736761), incorporated herein by reference), a *wcaF* gene (e.g., GenBank Accession Number (amino acid) BAA15910.1 (GI:1736760), incorporated herein by reference), a *wzxE* gene (e.g., GenBank Accession Number (amino acid) BAE77506.1 (GI:85676256), incorporated herein by reference), a *wxC* gene, (e.g., GenBank Accession Number (amino acid) BAA15899 (GI:1736748), incorporated herein by reference), a *wcaD* gene, (e.g., GenBank Accession Number (amino acid) BAE76573 (GI:85675202), incorporated herein by reference), a *wza* gene (e.g., GenBank Accession Number (amino acid) BAE76576 (GI:85675205), incorporated herein by reference), a *wzb* gene (e.g., GenBank Accession Number (amino acid) BAE76575 (GI:85675204), incorporated herein by reference), and a *wzc* gene (e.g., GenBank Accession Number (amino acid) BAA15913 (GI:1736763), incorporated herein by reference).

Preferably, a host bacterium, such as *E. coli*, is genetically engineered to produce a human milk oligosaccharide by the inactivation of the *wcaJ* gene, which encoding the UDP-glucose lipid carrier transferase. The inactivation of the *wcaJ* gene can be by deletion of the gene, a null mutation, or inactivating mutation of the *wcaJ* gene, such that the activity of the encoded *wcaJ* is reduced or eliminated compared to wild-type *E. coli*. In a *wcaJ* null background, GDP-fucose accumulates in the *E. coli* cytoplasm.

Over-expression of a positive regulator protein, RcsA (e.g., GenBank Accession Number M58003 (GI:1103316), incorporated herein by reference), in the colanic acid synthesis pathway results in an increase in intracellular GDP-fucose levels. Over-expression of an additional positive regulator of colanic acid biosynthesis, namely RcsB (e.g., GenBank Accession Number E04821 (GI:2173017), incorporated herein by reference), is also utilized, either instead of or in addition to over-expression of RcsA, to increase intracellular GDP-fucose levels.

Alternatively, colanic acid biosynthesis is increased following the introduction of a mutation into the *E. coli lon* gene (e.g., GenBank Accession Number L20572 (GI:304907), incorporated herein by reference). Lon is an adenosine-5'-triphosphate (ATP)-dependant intracellular protease that is responsible for degrading RcsA, mentioned above as a positive transcriptional regulator of colanic acid biosynthesis in *E. coli*. In a *lon* null background, RcsA is stabilized, RcsA levels increase, the genes responsible for GDP-fucose synthesis in *E. coli* are up-regulated, and intracellular GDP-fucose concentrations are enhanced. Mutations in *lon* suitable for use with the methods presented herein include null mutations or insertions that disrupt the expression or function of *lon*.

A functional lactose permease gene is also present in the bacterium. The lactose permease gene is an endogenous lactose permease gene or an exogenous lactose permease gene. For example, the lactose permease gene comprises an *E. coli lacY* gene (e.g., GenBank Accession Number V00295 (GI:41897), incorporated herein by reference). Many bacteria possess the inherent ability to transport lactose from the growth medium into the cell, by utilizing a transport protein that is either a homolog of the *E. coli* lactose permease (e.g., as found in *Bacillus licheniformis*), or a transporter that is a member of the ubiquitous PTS sugar transport family (e.g., as found in *Lactobacillus casei* and *Lactobacillus rhamnosus*). For bacteria lacking an inherent ability to transport extracellular lactose into the cell cytoplasm, this ability is conferred by an exogenous lactose transporter gene (e.g., *E. coli lacY*) provided on recombinant DNA constructs, and supplied either on a plasmid expression vector or as exogenous genes integrated into the host chromosome.

As described herein, in some embodiments, the host bacterium preferably has a reduced level of β -galactosidase activity. In the embodiment in which the bacterium is characterized by the deletion of the endogenous β -galactosidase gene, an exogenous β -galactosidase gene is introduced to the bacterium. For example, a plasmid expressing an exogenous β -galactosidase gene is introduced to the bacterium, or recombined or integrated

into the host genome. For example, the exogenous β -galactosidase gene is inserted into a gene that is inactivated in the host bacterium, such as the *lon* gene.

The exogenous β -galactosidase gene is a functional β -galactosidase gene characterized by a reduced or low level of β -galactosidase activity compared to β -galactosidase activity in wild-type bacteria lacking any genetic manipulation. Exemplary β -galactosidase genes include *E. coli lacZ* and β -galactosidase genes from any of a number of other organisms (e.g., the *lac4* gene of *Kluyveromyces lactis* (e.g., GenBank Accession Number M84410 (GI:173304), incorporated herein by reference) that catalyzes the hydrolysis of β -galactosides into monosaccharides. The level of β -galactosidase activity in wild-type *E. coli* bacteria is, for example, 6,000 units. Thus, the reduced β -galactosidase activity level encompassed by engineered host bacterium of the present invention includes less than 6,000 units, less than 5,000 units, less than 4,000 units, less than 3,000 units, less than 2,000 units, less than 1,000 units, less than 900 units, less than 800 units, less than 700 units, less than 600 units, less than 500 units, less than 400 units, less than 300 units, less than 200 units, less than 100 units, or less than 50 units. Low, functional levels of β -galactosidase include β -galactosidase activity levels of between 0.05 and 1,000 units, e.g., between 0.05 and 750 units, between 0.05 and 500 units, between 0.05 and 400 units, between 0.05 and 300 units, between 0.05 and 200 units, between 0.05 and 100 units, between 0.05 and 50 units, between 0.05 and 10 units, between 0.05 and 5 units, between 0.05 and 4 units, between 0.05 and 3 units, or between 0.05 and 2 units of β -galactosidase activity. For unit definition and assays for determining β -galactosidase activity, see Miller JH, Laboratory CSH. Experiments in molecular genetics. Cold Spring Harbor Laboratory Cold Spring Harbor, NY; 1972; (incorporated herein by reference). This low level of cytoplasmic β -galactosidase activity is not high enough to significantly diminish the intracellular lactose pool. The low level of β -galactosidase activity is very useful for the facile removal of undesired residual lactose at the end of fermentations.

Optionally, the bacterium has an inactivated *thyA* gene. Preferably, a mutation in a *thyA* gene in the host bacterium allows for the maintenance of plasmids that carry *thyA* as a selectable marker gene. Exemplary alternative selectable markers include antibiotic resistance genes such as BLA (beta-lactamase), or *proBA* genes (to complement a *proAB* host strain proline auxotrophy) or *purA* (to complement a *purA* host strain adenine auxotrophy).

In one embodiment, the *E. coli* bacterium comprises the genotype $\Delta ampC::P_{trp}^B cI$, $\Delta(lacI-lacZ)::FRT$, $P_{lacIq} lacY^+$, $\Delta wcaJ::FRT$, $thyA::Tn10$, $\Delta lon:(npt3, lacZ^+)$, $\Delta lacA$, and also comprises any one of the exogenous $\alpha(1,2)$ fucosyltransferases described herein.

The bacterium comprising these characteristics is cultured in the presence of lactose. In some cases, the method further comprises culturing the bacterium in the presence of tryptophan and in the absence of thymidine. The fucosylated oligosaccharide is retrieved from the bacterium (*i.e.*, a cell lysate) or from a culture supernatant of the bacterium.

The disclosure also relates to a purified fucosylated oligosaccharide produced by the methods described herein. The fucosylated oligosaccharide is purified for use in therapeutic or nutritional products, or the bacterium is used directly in such products. The fucosylated oligosaccharide produced by the engineered bacterium is 2'-fucosyllactose (2'-FL) or lactodifucotetraose (LDFT). The new α 1,2-fucosyltransferases are also useful to synthesize HMOS of larger molecular weight bearing α 1,2 fucose moieties, e.g., lacto-N-fucopentaose (LNF I) and lacto-N-difucohexaose (LDFH I). For example, to produce LDFT, the host bacterium is engineered to express an exogenous α (1,2) fucosyltransferase that also possesses α (1,3) fucosyltransferase activity, or an exogenous α (1,2) fucosyltransferase and an exogenous α (1,3) fucosyltransferase. For the production of LNF I and LDFH I, the host bacterium is engineered to express an exogenous α (1,2) fucosyltransferase that also possesses α (1,3) fucosyltransferase activity and/or α (1,4) fucosyltransferase activity, or an exogenous α (1,2) fucosyltransferase, an exogenous α (1,3) fucosyltransferase, and an exogenous α (1,4) fucosyltransferase.

A purified fucosylated oligosaccharide produced by the methods described above is also within the invention. The purified oligosaccharide (2'-FL) obtained at the end of the process is a white/slightly off-white, crystalline, sweet powder. For example, an engineered bacterium, bacterial culture supernatant, or bacterial cell lysate according to the invention comprises 2'-FL, LDFT, LNF I or LDFH I produced by the methods described herein, and does not substantially comprise an other fucosylated oligosaccharides prior to purification of the fucosylated oligosaccharide products from the cell, culture supernatant, or lysate. As a general matter, the fucosylated oligosaccharide produced by the methods contains a negligible amount of 3-FL in a 2'-FL-containing cell, cell lysate or culture, or supernatant, e.g., less than 1% of the level of 2'-FL or 0.5% of the level of 2'-FL. Moreover, the fucosylated oligosaccharide produced by the methods described herein also have a minimal amount of contaminating lactose, which can often be co-purified with the fucosylated

oligosaccharide product, such as 2'-FL. This reduction in contaminating lactose results from the reduced level of β -galactosidase activity present in the engineered host bacterium.

A purified oligosaccharide, e.g., 2'-FL, LDFT, LNF I, or LDFH I, is one that is at least 90%, 95%, 98%, 99%, or 100% (w/w) of the desired oligosaccharide by weight. Purity is assessed by any known method, e.g., thin layer chromatography or other chromatographic techniques known in the art. The disclosure also relates to a method of purifying a fucosylated oligosaccharide produced by the genetically engineered bacterium described above, which method comprises separating the desired fucosylated oligosaccharide (e.g., 2'-FL) from contaminants in a bacterial cell lysate or bacterial cell culture supernatant of the bacterium.

The oligosaccharides are purified and used in a number of products for consumption by humans as well as animals, such as companion animals (dogs, cats) as well as livestock (bovine, equine, ovine, caprine, or porcine animals, as well as poultry). For example, a pharmaceutical composition comprises purified 2'-FL and a pharmaceutically-acceptable excipient that is suitable for oral administration. Large quantities of 2'-FL are produced in bacterial hosts, e.g., an *E. coli* bacterium comprising an exogenous α (1,2) fucosyltransferase gene.

A method of producing a pharmaceutical composition comprising a purified human milk oligosaccharide (HMOS) is carried out by culturing the bacterium described above, purifying the HMOS produced by the bacterium, and combining the HMOS with an excipient or carrier to yield a dietary supplement for oral administration. These compositions are useful in methods of preventing or treating enteric and/or respiratory diseases in infants and adults. Accordingly, the compositions are administered to a subject suffering from or at risk of developing such a disease.

The disclosure also relates to methods of identifying an α (1,2) fucosyltransferase gene capable of synthesizing fucosylated oligosaccharides in a host bacterium, *i.e.*, 2'-fucosyllactose (2'-FL) in *E. coli*. The method of identifying novel lactose-utilizing, α (1,2)fucosyltransferase enzyme comprises the following steps:

- 1) performing a computational search of sequence databases to define a broad group of simple sequence homologs of any known, lactose-utilizing α (1,2)fucosyltransferase;
- 2) using the list from step (1), deriving a search profile containing common sequence and/or structural motifs shared by the members of the list;

3) searching sequence databases, using a derived search profile based on the common sequence or structural motif from step (2) as query, and identifying a candidate sequences, wherein a sequence homology to a reference lactose-utilizing $\alpha(1,2)$ fucosyltransferase is a predetermined percentage threshold;

4) compiling a list of candidate organisms, said organisms being characterized as expressing $\alpha(1,2)$ fucosyl- glycans in a naturally-occurring state;

5) selecting candidate sequences that are derived from candidate organisms to generate a list of candidate lactose-utilizing enzymes;

6) expressing the candidate lactose- utilizing enzyme in a host organism; and

7) testing for lactose- utilizing $\alpha(1,2)$ fucosyltransferase activity, wherein detection of the desired fucosylated oligosaccharide product in said organism indicates that the candidate sequence comprises a novel lactose- utilizing $\alpha(1,2)$ fucosyltransferase. In another embodiment, the search profile is generated from a multiple sequence alignment of the amino acid sequences of more than one enzyme with known $\alpha(1,2)$ fucosyltransferase activity. The database search can then be designed to refine and iteratively search for novel $\alpha(1,2)$ fucosyltransferases with significant sequence similarity to the multiple sequence alignment query.

The disclosure also relates to a method of treating, preventing, or reducing the risk of infection in a subject comprising administering to said subject a composition comprising a purified recombinant human milk oligosaccharide, wherein the HMOS binds to a pathogen and wherein the subject is infected with or at risk of infection with the pathogen. In one embodiment, the infection is caused by a Norwalk-like virus or *Campylobacter jejuni*. The subject is preferably a mammal in need of such treatment. The mammal is, e.g., any mammal, e.g., a human, a primate, a mouse, a rat, a dog, a cat, a cow, a horse, or a pig. In a preferred embodiment, the mammal is a human. For example, the compositions are formulated into animal feed (e.g., pellets, kibble, mash) or animal food supplements for companion animals, e.g., dogs or cats, as well as livestock or animals grown for food consumption, e.g., cattle, sheep, pigs, chickens, and goats. Preferably, the purified HMOS is formulated into a powder (e.g., infant formula powder or adult nutritional supplement powder, each of which is mixed with a liquid such as water or juice prior to consumption) or in the form of tablets, capsules or pastes or is incorporated as a component in dairy products such as milk, cream, cheese, yogurt or kefir, or as a component in any beverage, or combined in a preparation containing live microbial cultures intended to serve as probiotics, or in

prebiotic preparations to enhance the growth of beneficial microorganisms either *in vitro* or *in vivo*.

Polynucleotides, polypeptides, and oligosaccharides of the invention are purified and/or isolated. Purified defines a degree of sterility that is safe for administration to a human subject, e.g., lacking infectious or toxic agents. Specifically, as used herein, an "isolated" or "purified" nucleic acid molecule, polynucleotide, polypeptide, protein or oligosaccharide, is substantially free of other cellular material, or culture medium when produced by recombinant techniques, or chemical precursors or other chemicals when chemically synthesized. For example, purified HMOS compositions are at least 60% by weight (dry weight) the compound of interest. Preferably, the preparation is at least 75%, more preferably at least 90%, and most preferably at least 99%, by weight the compound of interest. Purity is measured by any appropriate standard method, for example, by column chromatography, thin layer chromatography, or high-performance liquid chromatography (HPLC) analysis. For example, a "purified protein" refers to a protein that has been separated from other proteins, lipids, and nucleic acids with which it is naturally associated. Preferably, the protein constitutes at least 10, 20, 50, 70, 80, 90, 95, 99-100% by dry weight of the purified preparation.

Similarly, by "substantially pure" is meant an oligosaccharide that has been separated from the components that naturally accompany it. Typically, the oligosaccharide is substantially pure when it is at least 60%, 70%, 80%, 90%, 95%, or even 99%, by weight, free from the proteins and naturally-occurring organic molecules with which it is naturally associated.

By "isolated nucleic acid" is meant a nucleic acid that is free of the genes which, in the naturally-occurring genome of the organism from which the DNA of the invention is derived, flank the gene. The term covers, for example: (a) a DNA which is part of a naturally occurring genomic DNA molecule, but is not flanked by both of the nucleic acid sequences that flank that part of the molecule in the genome of the organism in which it naturally occurs; (b) a nucleic acid incorporated into a vector or into the genomic DNA of a prokaryote or eukaryote in a manner, such that the resulting molecule is not identical to any naturally occurring vector or genomic DNA; (c) a separate molecule such as a cDNA, a genomic fragment, a fragment produced by polymerase chain reaction (PCR), or a restriction fragment; and (d) a recombinant nucleotide sequence that is part of a hybrid gene, i.e., a gene encoding a fusion protein. Isolated nucleic acid molecules according to the present invention

further include molecules produced synthetically, as well as any nucleic acids that have been altered chemically and/or that have modified backbones.

A "heterologous promoter" is a promoter which is different from the promoter to which a gene or nucleic acid sequence is operably linked in nature.

The term "overexpress" or "overexpression" refers to a situation in which more factor is expressed by a genetically-altered cell than would be, under the same conditions, by a wild type cell. Similarly, if an unaltered cell does not express a factor that it is genetically altered to produce, the term "express" (as distinguished from "overexpress") is used indicating the wild type cell did not express the factor at all prior to genetic manipulation.

The terms "treating" and "treatment" as used herein refer to the administration of an agent or formulation to a clinically symptomatic individual afflicted with an adverse condition, disorder, or disease, so as to effect a reduction in severity and/or frequency of symptoms, eliminate the symptoms and/or their underlying cause, and/or facilitate improvement or remediation of damage. The terms "preventing" and "prevention" refer to the administration of an agent or composition to a clinically asymptomatic individual who is susceptible to a particular adverse condition, disorder, or disease, and thus relates to the prevention of the occurrence of symptoms and/or their underlying cause.

By the terms "effective amount" and "therapeutically effective amount" of a formulation or formulation component is meant a nontoxic but sufficient amount of the formulation or component to provide the desired effect.

The transitional term "comprising," which is synonymous with "including," "containing," or "characterized by," is inclusive or open-ended and does not exclude additional, unrecited elements or method steps. By contrast, the transitional phrase "consisting of" excludes any element, step, or ingredient not specified in the claim. The transitional phrase "consisting essentially of" limits the scope of a claim to the specified materials or steps "and those that do not materially affect the basic and novel characteristic(s)" of the claimed invention.

The host organism used to express the lactose-accepting fucosyltransferase gene is typically the enterobacterium *Escherichia coli* K12 (*E. coli*). *E. coli* K-12 is not considered a human or animal pathogen nor is it toxicogenic. *E. coli* K-12 is a standard production strain of bacteria and is noted for its safety due to its poor ability to colonize the colon and establish infections (see, e.g., epa.gov/oppt/biotech/pubs/fra/fra004.htm). However, a variety of bacterial species may be used in the oligosaccharide biosynthesis methods, e.g., *Erwinia herbicola* (*Pantoea agglomerans*), *Citrobacter freundii*, *Pantoea citrea*, *Pectobacterium*

carotovorum, or *Xanthomonas campestris*. Bacteria of the genus *Bacillus* may also be used, including *Bacillus subtilis*, *Bacillus licheniformis*, *Bacillus coagulans*, *Bacillus thermophilus*, *Bacillus laterosporus*, *Bacillus megaterium*, *Bacillus mycoides*, *Bacillus pumilus*, *Bacillus lentus*, *Bacillus cereus*, and *Bacillus circulans*. Similarly, bacteria of the genera *Lactobacillus* and *Lactococcus* may be modified using the methods of this invention, including but not limited to *Lactobacillus acidophilus*, *Lactobacillus salivarius*, *Lactobacillus plantarum*, *Lactobacillus helveticus*, *Lactobacillus delbrueckii*, *Lactobacillus rhamnosus*, *Lactobacillus bulgaricus*, *Lactobacillus crispatus*, *Lactobacillus gasseri*, *Lactobacillus casei*, *Lactobacillus reuteri*, *Lactobacillus jensenii*, and *Lactococcus lactis*. *Streptococcus thermophiles* and *Propionibacterium freudenreichii* are also suitable bacterial species for the invention described herein. Also included as part of this invention are strains, modified as described here, from the genera *Enterococcus* (e.g., *Enterococcus faecium* and *Enterococcus thermophiles*), *Bifidobacterium* (e.g., *Bifidobacterium longum*, *Bifidobacterium infantis*, and *Bifidobacterium bifidum*), *Sporolactobacillus* spp., *Micromonospora* spp., *Micrococcus* spp., *Rhodococcus* spp., and *Pseudomonas* (e.g., *Pseudomonas fluorescens* and *Pseudomonas aeruginosa*). Bacteria comprising the characteristics described herein are cultured in the presence of lactose, and a fucosylated oligosaccharide is retrieved, either from the bacterium itself or from a culture supernatant of the bacterium. The fucosylated oligosaccharide is purified for use in therapeutic or nutritional products, or the bacteria are used directly in such products. A suitable production host bacterial strain is one that is not the same bacterial strain as the source bacterial strain from which the fucosyltransferase-encoding nucleic acid sequence was identified.

Other features and advantages of the disclosure will be apparent from the following description of the preferred embodiments thereof, and from the claims. Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, suitable methods and materials are described below. All published foreign patents and patent applications cited herein are incorporated herein by reference. Genbank and NCBI submissions indicated by accession number cited herein are incorporated herein by reference. All other published references, documents, manuscripts and scientific literature cited herein are incorporated herein by reference. In the case of conflict, the present specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a schematic illustration showing the synthetic pathway of the major neutral fucosyl-oligosaccharides found in human milk.

FIG. 2 is a schematic demonstrating metabolic pathways and the changes introduced into them to engineer 2' -fucosyllactose (2' -FL) synthesis in *Escherichia coli* (*E. coli*). Specifically, the lactose synthesis pathway and the GDP-fucose synthesis pathway are illustrated. In the GDP-fucose synthesis pathway: *manA* = phosphomannose isomerase (PMI), *manB* = phosphomannomutase (PMM), *manC* = mannose-1-phosphate guanylyltransferase (GMP), *gmd* = GDP-mannose-4,6-dehydratase, *fcl* = GDP-fucose synthase (GFS), and $\Delta wcaJ$ = mutated UDP-glucose lipid carrier transferase.

FIG. 3A and FIG. 3B show the sequence identity and a multiple sequence alignment of 4 previously known lactose-utilizing $\alpha(1,2)$ -fucosyltransferase protein sequences. FIG. 3A is a table showing the sequence identity between the 4 known lactose-utilizing $\alpha(1,2)$ -fucosyltransferases: *H. pylori* futC (SEQ ID NO: 1), *H. mustelae* FutL (SEQ ID NO: 2), *Bacteroides vulgatus* futN (SEQ ID NO: 3), and *E. coli* 0126 wbgL (SEQ ID NO: 4). FIG. 3B shows multiple sequence alignment of the 4 known $\alpha(1,2)$ -fucosyltransferases. The ovals highlight regions of particularly high sequence conservation between the four enzymes in the alignment.

FIG. 4 shows the sequence alignment of the 12 identified $\alpha(1,2)$ -fucosyltransferase syngenes identified, along with the 4 previously known lactose-utilizing $\alpha(1,2)$ -fucosyltransferase protein sequences. The 4 known lactose-utilizing $\alpha(1,2)$ -fucosyltransferases are boxed and include *H. pylori* futC (SEQ ID NO: 293), *H. mustelae* FutL (SEQ ID NO: 2), *Bacteroides vulgatus* futN (SEQ ID NO: 3), and *E. coli* 0126 wbgL (SEQ ID NO: 294). The 12 identified $\alpha(1,2)$ -fucosyltransferase are as follows: *Prevotella melaninogenica* FutO (SEQ ID NO: 10), *Clostridium bolteae* + 13 FutP (SEQ ID NO: 292), *Lachnospiraceae* sp. FutQ (SEQ ID NO: 12), *Methanosphaerula palustris* FutR (SEQ ID NO: 13), *Tannerella* sp. FutS (SEQ ID NO: 14), *Bacteroides caccae* FutU (SEQ ID NO: 15), *Butyrivibrio* FutV (SEQ ID NO: 16), *Prevotella* sp. FutW (SEQ ID NO: 17), *Parabacteroides johnsonii* FutX (SEQ ID NO: 18), *Akkermansia muciniphilia* FutY (SEQ ID NO: 19), *Salmonella enterica* FutZ (SEQ ID NO: 20), *Bacteroides* sp. FutZA (SEQ ID NO: 21). The sequence for *Clostridium bolteae* FutP (without the 13 additional amino acids in the N-terminus) (SEQ ID NO: 11) is also shown in the alignment.

FIG. 5A and FIG. 5B are two pictures of gels showing the construction of the syngenes for each of the 12 novel $\alpha(1,2)$ -fucosyltransferases. FIG. 5A shows post-Gibson assembly PCR. FIG. 5B shows gel-purified RI/XhoI syngene fragments.

FIG. 6A and FIG. 6B are two photographs showing thin layer chromatograms of fucosylated oligosaccharide products produced in *E. coli* cultures using the 12 novel $\alpha(1,2)$ -fucosyltransferase syngenes. FIG. 6A shows fucosylated oligosaccharide products from 2 μ l of culture supernatant. FIG. 6B shows fucosylated oligosaccharide products from 0.2 OD₆₀₀ cell equivalents of whole cell heat extracts.

FIG. 7 is a graph showing the growth curve of the host bacterium expressing plasmids containing the $\alpha(1,2)$ fucosyltransferase genes WbgL, FutN, FutO, FutQ, and FutX after tryptophan induction in the presence of lactose in the culture medium (i.e. lac + trp).

FIG. 8 is a photograph of a SDS-PAGE gel showing the proteins produced from host bacterium expressing $\alpha(1,2)$ fucosyltransferase genes WbgL, FutN, FutO, FutQ, and FutX after induction.

FIG. 9A and FIG. 9B are two photographs of thin layer chromatograms showing the production of fucosylated oligosaccharide products from in *E. coli* cultures expressing select $\alpha(1,2)$ -fucosyltransferase syngenes WbgL, FutN, FutO, FutQ, and FutX at 7 hours or 24 hours after induction. FIG. 9A shows fucosylated oligosaccharide products from 2 μ l of culture supernatant. FIG. 9B shows fucosylated oligosaccharide products from 0.2 OD₆₀₀ cell equivalents of whole cell heat extracts.

FIG. 10A and FIG. 10B are two photographs of thin layer chromatograms showing the fucosylated oligosaccharide products after two different 1.5L fermentation runs from *E. coli* expressing FutN: FIG. 10A) 36B and FIG. 10B) 37A. The culture yield for run 36B was 33g/L while the yield for run 37A was 36.3 g/L.

FIG. 11 is a plasmid map of pG217 carrying the *B. vulgatus* FutN gene.

FIG. 12 is a schematic diagram showing the insertion of the *LacIq* promoter, the functional *lacY* gene, and the deletion of *lacA*.

FIG. 13 is a schematic diagram showing the deletion of the endogenous *wcaJ* gene using FRT recombination.

FIG. 14 is a schematic diagram of the *E. coli* W3110 chromosome, showing the insertion of a DNA fragment carrying kanamycin resistance gene (derived from transposon Tn5) and wild-type *lacZ* into the *lon* gene.

DETAILED DESCRIPTION OF THE INVENTION

While some studies suggest that human milk glycans could be used as antimicrobial anti-adhesion agents, the difficulty and expense of producing adequate quantities of these agents of a quality suitable for human consumption has limited their full-scale testing and perceived utility. What has been needed is a suitable method for producing the appropriate glycans in sufficient quantities at reasonable cost. Prior to the invention described herein, there were attempts to use several distinct synthetic approaches for glycan synthesis. Some chemical approaches can synthesize oligosaccharides (Flowers, H. M. *Methods Enzymol* 50, 93-121 (1978); Seeberger, P. H. *Chem Commun (Camb)* 1115-1121 (2003)), but reactants for these methods are expensive and potentially toxic (Koeller, K. M. & Wong, C. H. *Chem Rev* 100, 4465-4494 (2000)). Enzymes expressed from engineered organisms (Albermann, C., Piepersberg, W. & Wehmeier, U. F. *Carbohydr Res* 334, 97-103 (2001); Bettler, E., Samain, E., Chazalet, V., Bosso, C., et al. *Glycoconj J* 16, 205-212 (1999); Johnson, K. F. *Glycoconj J* 16, 141-146 (1999); Palcic, M. M. *Curr Opin Biotechnol* 10, 616-624 (1999); Wymer, N. & Toone, E. J. *Curr Opin Chem Biol* 4, 110-119 (2000)) provide a precise and efficient synthesis (Palcic, M. M. *Curr Opin Biotechnol* 10, 616-624 (1999)); Crout, D. H. & Vic, G. *Curr Opin Chem Biol* 2, 98-111 (1998)), but the high cost of the reactants, especially the sugar nucleotides, limits their utility for low-cost, large-scale production. Microbes have been genetically engineered to express the glycosyltransferases needed to synthesize oligosaccharides from the bacteria's innate pool of nucleotide sugars (Endo, T., Koizumi, S., Tabata, K., Kakita, S. & Ozaki, A. *Carbohydr Res* 330, 439-443 (2001); Endo, T., Koizumi, S., Tabata, K. & Ozaki, A. *Appl Microbiol Biotechnol* 53, 257-261 (2000); Endo, T. & Koizumi, S. *Curr Opin Struct Biol* 10, 536-541 (2000); Endo, T., Koizumi, S., Tabata, K., Kakita, S. & Ozaki, A. *Carbohydr Res* 316, 179-183 (1999); Koizumi, S., Endo, T., Tabata, K. & Ozaki, A. *Nat Biotechnol* 16, 847-850 (1998)). However, prior to the invention described herein, there was a growing need to identify and characterize additional glycosyltransferases that are useful for the synthesis of HMOS in metabolically engineered bacterial hosts.

Human Milk Glycans

Human milk contains a diverse and abundant set of neutral and acidic oligosaccharides (Kunz, C., Rudloff, S., Baier, W., Klein, N., and Strobel, S. (2000). *Annu Rev Nutr* 20, 699-722; Bode, L. (2006). *J Nutr* 136, 2127-130). More than 130 different complex oligosaccharides have been identified in human milk, and their structural diversity and abundance is unique to humans. Although these molecules may not be utilized directly

by infants for nutrition, they nevertheless serve critical roles in the establishment of a healthy gut microbiome (Marcobal, A., Barboza, M., Froehlich, J. W., Block, D. E., et al. *J Agric Food Chem* 58, 5334-5340 (2010)), in the prevention of disease (Newburg, D. S., Ruiz-Palacios, G. M. & Morrow, A. L. *Annu Rev Nutr* 25, 37-58 (2005)), and in immune function (Newburg, D. S. & Walker, W. A. *Pediatr Res* 61, 2-8 (2007)). Despite millions of years of exposure to human milk oligosaccharides (HMOS), pathogens have yet to develop ways to circumvent the ability of HMOS to prevent adhesion to target cells and to inhibit infection. The ability to utilize HMOS as pathogen adherence inhibitors promises to address the current crisis of burgeoning antibiotic resistance. Human milk oligosaccharides produced by biosynthesis represent the lead compounds of a novel class of therapeutics against some of the most intractable scourges of society.

One alternative strategy for efficient, industrial-scale synthesis of HMOS is the metabolic engineering of bacteria. This approach involves the construction of microbial strains overexpressing heterologous glycosyltransferases, membrane transporters for the import of precursor sugars into the bacterial cytosol, and possessing enhanced pools of regenerating nucleotide sugars for use as biosynthetic precursors (Dumon, C., Samain, E., and Priem, B. (2004). *Biotechnol Prog* 20, 412-19; Ruffing, A., and Chen, R.R. (2006). *Microb Cell Fact* 5, 25). A key aspect of this approach is the heterologous glycosyltransferase selected for overexpression in the microbial host. The choice of glycosyltransferase can significantly affect the final yield of the desired synthesized oligosaccharide, given that enzymes can vary greatly in terms of kinetics, substrate specificity, affinity for donor and acceptor molecules, stability and solubility. A few glycosyltransferases derived from different bacterial species have been identified and characterized in terms of their ability to catalyze the biosynthesis of HMOS in *E. coli* host strains (Dumon, C., Bosso, C., Utile, J.P., Heyraud, A., and Samain, E. (2006). *Chembiochem* 7, 359-365; Dumon, C., Samain, E., and Priem, B. (2004). *Biotechnol Prog* 20, 412-19; Li, M., Liu, X.W., Shao, J., Shen, J., Jia, Q., Yi, W., Song, J.K., Woodward, R., Chow, C.S., and Wang, P.G. (2008). *Biochemistry* 47, 378-387). The identification of additional glycosyltransferases with faster kinetics, greater affinity for nucleotide sugar donors and/or acceptor molecules, or greater stability within the bacterial host significantly improves the yields of therapeutically useful HMOS. Prior to the invention described herein, chemical syntheses of HMOS were possible, but were limited by stereo-specificity issues, precursor availability, product impurities, and high overall cost (Flowers, H. M. *Methods Enzymol* 50, 93-121 (1978); Seeberger, P. H. *Chem Commun (Camb)* 1115-1121 (2003);

Koeller, K. M. & Wong, C. H. Chem Rev 100, 4465-4494 (2000)). The invention overcomes the shortcomings of these previous attempts by providing new strategies to inexpensively manufacture large quantities of human milk oligosaccharides (HMOS) for use as dietary supplements. Advantages include efficient expression of the enzyme, improved stability and/or solubility of the fucosylated oligosaccharide product (2'-FL, LDFT, LNF I, and LDFH I) and reduced toxicity to the host organism. The present invention features novel $\alpha(1,2)$ FTs suitable for expression in production strains for increased efficacy and yield of fucosylated HMOS compared to $\alpha(1,2)$ FTs currently utilized in the field.

As described in detail below, *E. coli* (or other bacteria) is engineered to produce selected fucosylated oligosaccharides (*i.e.*, 2'-FL, LDFT, LDHF I, or LNF I) in commercially viable levels. For example, yields are >5 grams/liter in a bacterial fermentation process. In other embodiments, the yields are greater than 10 grams/liter, greater than 15 grams/liter, greater than 20 grams/liter, greater than 25 grams/liter, greater than 30 grams/liter, greater than 35 grams/liter, greater than 40 grams/liter, greater than 45 grams/liter, greater than 50 grams/liter, greater than 55 grams/liter, greater than 60 grams/liter, greater than 65 grams/liter, greater than 70 grams/liter, or greater than 75 grams/liter of fucosylated oligosaccharide products, such as 2'-FL, LDFT, LDHF I, and LNF I.

Role of Human milk glycans in infectious disease

Human milk glycans, which comprise both unbound oligosaccharides and their glycoconjugates, play a significant role in the protection and development of the infant gastrointestinal (GI) tract. Neutral fucosylated oligosaccharides, including 2'-fucosyllactose (2'-FL), protect infants against several important pathogens. Milk oligosaccharides found in various mammals differ greatly, and the composition in humans is unique (Hamosh M., 2001 *Pediatr Clin North Am*, 48:69-86; Newburg D.S., 2001 *Adv Exp Med Biol*, 501:3-10). Moreover, glycan levels in human milk change throughout lactation and also vary widely among individuals (Morrow A.L. et al., 2004 *J Pediatr*, 145:297-303; Chaturvedi P et al., 2001 *Glycobiology*, 11:365-372). Approximately 200 distinct human milk oligosaccharides have been identified and combinations of simple epitopes are responsible for this diversity (Newburg D.S., 1999 *Curr Med Chem*, 6:117-127; Ninonuevo M. et al., 2006 *J Agric Food Chem*, 54:7471-7480).

Human milk oligosaccharides are composed of 5 monosaccharides: D-glucose (Glc), D-galactose (Gal), N-acetylglucosamine (GlcNAc), L-fucose (Fuc), and sialic acid (N-acetylneuraminic acid, Neu5Ac, NANA). Human milk oligosaccharides are usually divided into

two groups according to their chemical structures: neutral compounds containing Glc, Gal, GlcNAc, and Fuc, linked to a lactose (Gal β 1-4Glc) core, and acidic compounds including the same sugars, and often the same core structures, plus NANA (Charlwood J. et al., 1999 Anal Biochem, 273:261-277; Martín-Sosa et al., 2003 J Dairy Sci, 86:52-59; Parkkinen J. and Finne J., 1987 Methods Enzymol, 138:289-300; Shen Z. et al., 2001 J Chromatogr A, 921:315-321).

Approximately 70-80% of oligosaccharides in human milk are fucosylated, and their synthetic pathways are believed to proceed as shown in FIG. 1. A smaller proportion of the oligosaccharides are sialylated or both fucosylated and sialylated, but their synthetic pathways are not fully defined. Understanding of the acidic (sialylated) oligosaccharides is limited in part by the ability to measure these compounds. Sensitive and reproducible methods for the analysis of both neutral and acidic oligosaccharides have been designed. Human milk oligosaccharides as a class survive transit through the intestine of infants very efficiently, being essentially indigestible (Chaturvedi, P., Warren, C. D., Buescher, C. R., Pickering, L. K. & Newburg, D. S. Adv Exp Med Biol 501, 315-323 (2001)).

Human milk glycans inhibit binding of enteropathogens to their receptors

Human milk glycans have structural homology to cell receptors for enteropathogens and function as receptor decoys. For example, pathogenic strains of *Campylobacter* bind specifically to glycans containing H-2, i.e., 2'-fucosyl-N-acetyllactosamine or 2'-fucosyllactose (2'FL); *Campylobacter* binding and infectivity are inhibited by 2'-FL and other glycans containing this H-2 epitope. Similarly, some diarrheagenic *E. coli* pathogens are strongly inhibited *in vivo* by human milk oligosaccharides containing 2-linked fucose moieties. Several major strains of human caliciviruses, especially the noroviruses, also bind to 2-linked fucosylated glycans, and this binding is inhibited by human milk 2-linked fucosylated glycans. Consumption of human milk that has high levels of these 2-linked fucosyloligosaccharides was associated with lower risk of norovirus, *Campylobacter*, ST of *E. coli*-associated diarrhea, and moderate-to-severe diarrhea of all causes in a Mexican cohort of breastfeeding children (Newburg D.S. et al., 2004 Glycobiology, 14:253-263; Newburg D.S. et al., 1998 Lancet, 351:1160-1164). Several pathogens utilize sialylated glycans as their host receptors, such as influenza (Couceiro, J. N., Paulson, J. C. & Baum, L. G. Virus Res 29, 155-165 (1993)), parainfluenza (Amonsén, M., Smith, D. F., Cummings, R. D. & Air, G. M. J Virol 81, 8341-8345 (2007)), and rotoviruses (Kuhlenschmidt, T. B., Hanafin, W. P., Gelberg, H. B. & Kuhlenschmidt, M. S. Adv Exp Med Biol 473, 309-317 (1999)). The

sialyl-Lewis X epitope is used by *Helicobacter pylori* (Mahdavi, J., Sondén, B., Hurtig, M., Olfat, F. O., et al. Science 297, 573-578 (2002)), *Pseudomonas aeruginosa* (Scharfman, A., Delmotte, P., Beau, J., Lamblin, G., et al. Glycoconj J 17, 735-740 (2000)), and some strains of noroviruses (Rydell, G. E., Nilsson, J., Rodriguez-Diaz, J., Ruvoën-Clouet, N., et al. Glycobiology 19, 309-320 (2009)).

Identification of novel $\alpha(1,2)$ fucosyltransferases

The present invention provides novel $\alpha(1,2)$ fucosyltransferase enzymes. The present invention also provides nucleic acid constructs (*i.e.*, a plasmid or vector) carrying the nucleic acid sequence of a novel $\alpha(1,2)$ fucosyltransferases for the expression of the novel $\alpha(1,2)$ fucosyltransferases in host bacterium. The present invention also provides methods for producing fucosylated oligosaccharides by expressing the novel $\alpha(1,2)$ fucosyltransferases in suitable host production bacterium, as further described herein.

Not all $\alpha(1,2)$ fucosyltransferases can utilize lactose as an acceptor substrate. An acceptor substrate includes, for example, a carbohydrate, an oligosaccharide, a protein or glycoprotein, a lipid or glycolipid, *e.g.*, N-acetylglucosamine, N-acetyllactosamine, galactose, fucose, sialic acid, glucose, lactose, or any combination thereof. A preferred alpha (1,2) fucosyltransferase of the present invention utilizes GDP-fucose as a donor, and lactose is the acceptor for that donor.

A method of identifying novel $\alpha(1,2)$ fucosyltransferase enzymes capable of utilizing lactose as an acceptor was previously carried out (as described in PCT/US2013/051777, hereby incorporated by reference in its entirety) using the following steps: 1) performing a computational search of sequence databases to define a broad group of simple sequence homologs of any known, lactose-utilizing $\alpha(1,2)$ fucosyltransferase; 2) using the list of homologs from step 1 to derive a search profile containing common sequence and/or structural motifs shared by the members of the broad group, *e.g.* by using computer programs such as MEME (Multiple Em for Motif Elicitation available at <http://meme.sdsc.edu/meme/cgi-bin/meme.cgi>) or PSI-BLAST (Position-Specific Iterated BLAST available at ncbi.nlm.nih.gov/blast with additional information at cnx.org/content/m11040/latest/); 3) searching sequence databases (*e.g.*, using computer programs such as PSI-BLAST, or MAST (Motif Alignment Search Tool available at <http://meme.sdsc.edu/meme/cgi-bin/mast.cgi>);, using this derived search profile as query, and identifying “candidate sequences” whose simple sequence homology to the original lactose-accepting $\alpha(1,2)$ fucosyltransferase is 40% or less; 4) scanning the scientific literature and

developing a list of “candidate organisms” known to express $\alpha(1,2)$ fucosyl-glycans; 5) selecting only those “candidate sequences” that are derived from “candidate organisms” to generate a list of “candidate lactose-utilizing enzymes”; and 6) expressing each “candidate lactose-utilizing enzyme” and testing for lactose-utilizing $\alpha(1,2)$ fucosyltransferase activity.

The MEME suite of sequence analysis tools (meme.sdsc.edu/meme/cgi-bin/meme.cgi) can also be used as an alternative to PSI-BLAST. Sequence motifs are discovered using the program “MEME”. These motifs can then be used to search sequence databases using the program “MAST”. The BLAST and PSI-BLAST search algorithms are other well known alternatives.

To identify additional novel $\alpha(1,2)$ fucosyltransferases, a multiple sequence alignment query was generated using four previously identified lactose-utilizing $\alpha(1,2)$ fucosyltransferase protein sequences: *H. pylori* futC (SEQ ID NO: 1), *H. mustelae* FutL (SEQ ID NO: 2), *Bacteroides vulgatus* futN (SEQ ID NO: 3), and *E. coli* 0126 wbgL (SEQ ID NO: 4). These sequence alignment and percentage of sequence identity is shown in FIG. 3. An iterative PSI-BLAST was performed, using the FASTA-formatted multiple sequence alignment as the query, and the NCBI PSI-BLAST program run on a local copy of NCBI BLAST+ version 2.2.29. An initial position-specific scoring matrix file (.pssm) was generated by PSI-BLAST, which the program then used to adjust the score of iterative homology search runs. The process is iterated to generate an even larger group of candidates, and the results of each run were used to further refine the matrix.

This PSI-BLAST search resulted in an initial 2515 hits. There were 787 hits with greater than 22% sequence identity to FutC. 396 hits were of greater than 275 amino acids in length. Additional analysis of the hits was performed, including sorting by percentage identity to FutC, comparing the sequences by BLAST to existing $\alpha(1,2)$ fucosyltransferase inventory (of known $\alpha(1,2)$ fucosyltransferases), and manual annotation of hit sequences to identify those originating from bacteria that naturally exist in the gastrointestinal tract. An annotated list of the novel $\alpha(1,2)$ fucosyltransferases identified by this screen are listed in Table 1. Table 1 provides the bacterial species from which the candidate enzyme is found, the GenBank Accession Number, GI Identification Number, amino acid sequence, and % sequence identity to FutC.

Of the identified hits, 12 novel $\alpha(1,2)$ fucosyltransferases were further analyzed for their functional capacity: *Prevotella melaninogenica* FutO, *Clostridium bolteae* FutP, *Clostridium bolteae* +13 FutP, *Lachnospiraceae* sp. FutQ, *Methanosphaerula palustries*

FutR, *Tannerella* sp. FutS, *Bacteroides caccae* FutU, *Butyrivibrio* FutV, *Prevotella* sp.

FutW, *Parabacteroides johnsonii* FutX, *Akkermansia muciniphila* FutY, *Salmonella enterica*

FutZ, and *Bacteroides* sp. FutZA. For *Clostridium bolteae* FutP, the annotation named the

wrong initiation methionine codon. Thus, the present invention includes FutP with an

additional 13 amino acids at the N-terminus of the annotated FutP (derived in-frame from the

natural upstream DNA sequence), which is designated herein as *Clostridium bolteae* +13

FutP. The sequence identity between the 12 novel $\alpha(1,2)$ fucosyltransferases identified and

the 4 previously identified $\alpha(1,2)$ fucosyltransferases is shown in Table 2 below.

Table 2. Sequence Identity

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
<i>H. pylori</i> futC	1		70.10	21.99	20.82	27.68	27.36	23.56	23.28	23.62	25.75	23.72	24.05	22.29	24.19	22.92	22.29
<i>H. mustelae</i> futL	2	70.10		23.87	19.88	26.38	28.21	24.30	23.38	24.62	25.31	25.31	24.47	23.56	25.15	23.55	23.26
<i>Bacteroides vulgatus</i> futN	3	21.99	23.87		25.16	32.05	28.71	28.94	25.79	37.46	32.27	26.11	61.27	71.63	27.67	25.15	84.75
<i>E. coli</i> O126 wbgL	4	20.82	19.88	25.16		24.25	22.73	22.32	26.04	25.45	24.77	21.49	23.29	26.71	24.83	21.45	25.16
<i>Prevotella melanogenica</i> FutO YP_003814512.1	5	27.68	26.38	32.05	24.25		36.96	31.63	35.74	35.16	55.74	30.28	30.03	32.80	30.09	26.28	31.83
<i>Clostridium bolteae</i> + 13 FutP WP_002570768.1	6	27.36	28.21	28.71	22.73	36.96		37.87	35.10	33.77	36.91	35.74	29.58	31.39	27.67	26.33	29.13
<i>Lachnospiraceae</i> sp. FutQ WP_009251343.1	7	23.56	24.30	26.94	22.32	31.63	37.87		29.87	29.17	32.90	51.02	28.53	30.00	27.69	24.00	27.74
<i>Methanospiraerula palustris</i> FutR YP_002467213.1	8	23.28	23.38	25.79	26.04	35.74	35.10	29.87		28.71	38.24	31.41	25.39	28.08	30.65	23.93	25.55
<i>Tannerella</i> sp. FutS WP_021929367.1	9	23.62	24.62	37.46	25.45	35.16	33.77	29.17	28.71		34.41	30.03	35.71	36.27	26.48	21.75	36.60
<i>Bacteroides caccae</i> FutU WP_005675707.1	10	25.75	25.31	32.77	24.77	55.74	36.91	32.90	38.24	34.41		31.21	29.94	33.33	29.28	24.46	33.01
<i>Butyrivibrio</i> FutV WP_022727218.1	11	23.72	25.31	26.11	21.49	30.28	35.74	51.02	31.41	30.03	31.21		27.62	26.20	26.46	22.15	26.52
<i>Prevotella</i> sp. FutW WP_022481266.1	12	24.05	24.47	61.27	23.29	30.03	29.58	28.53	25.39	35.71	29.94	27.62		57.60	25.79	22.15	59.01
<i>Parabacteroides johnsonii</i> FutX WP_008155883.1	13	22.29	23.56	71.63	26.71	32.80	31.39	30.00	28.08	36.27	33.33	26.20	57.60		28.71	24.00	74.02
<i>Akkermansia muciniphila</i> FutY YP_001877555	14	24.19	25.15	27.67	24.63	30.09	27.67	27.69	30.65	26.48	29.28	26.46	25.79	28.71		21.45	28.08
<i>Salmonella enterica</i> FutZ WP_023214330	15	22.92	23.55	25.15	21.45	26.28	26.33	24.00	23.93	21.75	24.46	22.15	22.15	24.00	21.45		24.62
<i>Bacteroides</i> sp. FutZA WP_022161880.1	16	22.29	23.26	84.75	25.16	31.83	29.13	27.74	25.55	36.60	33.01	26.52	59.01	74.02	28.08	24.62	

Based on the amino acid sequences of the identified $\alpha(1,2)$ fucosyltransferases (*i.e.*, in Table 1), syngenes can be readily designed and constructed by the skilled artisan using standard methods known in the art. For example, the syngenes include a ribosomal binding site, are codon-optimized for expression in a host bacterial production strain (*i.e.*, *E. coli*), and have common 6-cutter restriction sites or sites recognized by endogenous restriction enzymes present in the host strain (*i.e.*, EcoK restriction sites) removed to ease cloning and expression in the *E. coli* host strain. In a preferred embodiment, the syngenes are constructed with the following configuration: EcoRI site – T7g10 RBS – $\alpha(1,2)$ FT syngene – XhoI site. The nucleic acid sequences of sample syngenes for the 12 identified $\alpha(1,2)$ fucosyltransferases are shown in Table 3. (the initiating methionine ATG codon is bolded)

Table 3. Nucleic acid sequences of 12 novel $\alpha(1,2)$ fucosyltransferase syngenes

Bacteria/ Gene name	Sequence	SEQ ID NO:
FutO	CAGTCAGTCAGAATTCAAGAAGGAGATATACAT ATG AAAATCGTCAAAAATCCTGGGCGGTCTGGGCAATCAGATGTTCCAGTATGCTCTGTACCTGAGCCTGCAAGAAAGTTTTCACAAAAGAACGTGTGGCCCTGGACCTGTCCTCCTTCCACGGCTATCACCTGCATAATGGCTTTGAGCTGGAGAACATTTTCTCCGTTACCGCTCAGAAAGCATCCGCCGAGATATCATGCGTATTGCTTATTACTACCCGAACATCTGCTGTGGCGCATTGGCAAACGTTTCTGCCGCGTCGTAAAGGTATGTGCCTGGAATCTAGCTCCCTGCGTTTCGATGAAAGCGTTTCTGCGTCAGGAAAGTAACCGTTATTTTGACGGTTACTGGCAAGACGAACGCTACTTCGCAGCCTATCGTGAA	276

	AAAGTGCCTGAAGGCTTTCACCTTCTCCTGCATTCAAACGCGCAGAAAACCTGAGCCTGCTG GAAAAACTGGACGAAAAACAGCATTGCTCTGCATGTTTCGTGCGGGTGATTACGTAGGTAAT AACCTGTACCAAGGCATCTGTGACCTGGACTACTACCGTACCGCTATCGAGAAAATGTGT GCACACGTTACTCCGTCTCTGTTTTGTATCTTTTCCAACGACATCACGTGGTGCCAGCAG CACCTGCAACCGTACCTGAAGGCCCTGTGGTGACGTTACTTGGAACACCGGTGTTGAA TCCTACCGCGATATGCAGCTGATGTCCTGCTGCGCACATAACATCATCGCGAATAGCTCC TTCTCTTGGTGGGGTGCTTGGCTGAATCAGAACCGTGAAAAAGTTGTTATCGCCCCGAAA AAATGGCTGAACATGGAAGAATGTCACCTCACGCTGCCGGCAAGCTGGATCAAAATTTAG CTCGAGTGACTGACTG	
FutP	CAGTCAGTCAGAATTCAAGAAGGAGATATACATATGGTGATTATCAAAATGATGGGTGGT CTGGGCAACCAGATGTTCCAGTACGCACTGTACAAAGCATTTCGAGCAGAAGCACATCGAT GTGTATGCAGACCTGGCATGGTACAAAAACAAATCCGTGAAATTTGAACTGTACAACCTC GGCATTAAAAATCAACGTAGCATCCGAGAAAGACATCAACCGTCTGAGCGATTGCCAGGCG GACTTTGTTTCCCGCATCCGCCGTAAAAATCTTTGGTAAAAAAAAGAGCTTCGTATCTGAA AAAAATGACTCCTGCTATGAAAACGACATCCTGCGTATGGACAACGTTTATCTGAGCGGT TATTGGCAGACCGAAAAATACTTCTCTAACACGCGTGAGAAGCTGCTGGAGGATTATTCC TTCGCTCTGGTAAACTCTCAGGTGTCCGAATGGGAAGACTCCATTTCGCAACAAAAACAGC GTTAGCATCCATATCCGTCGTGGTGATTATCTACAGGGCGAACTGTATGGTGGTATTGTC ACCTCTCTGTACTACGCCGAAGCAATCGAGTACATTAAAAATGCGTGTTCCGAACGCAAAA TTCTTCGTTTTCTCTGATGACGTTGAATGGGTAAACAGCAAGAAGACTTCAAAGGCTTC GTAATCGTTGATCGCAACGAGTATTCTAGCGCTCTGTCCGATATGTACCTGATGTCCCTG TGCAAGCATAACATTATTGCTAACTCCTCTTTAGCTGGTGGGCGAGCTTGGCTGAACCGT AACGAAGAAAAAATGTAATCGCGCCGCGCGTGGCTGAACGGCAAGTGCACCCAGAT ATCTGGTGTAAAAAATGGATTTCGTATCTAGCTCGAGTGACTGACTG	277
FutQ	CAGTCAGTCAGAATTCAAGAAGGAGATATACATATGGTGATCGTACAGCTGAGCGGCGGT CTGGGCAACCAGATGTTTCGAATACGCGCTGTACCTGAGCCTGAAAGCAAAAGGCAAGAA GTGAAAATTGACGATGTTACGTGTTACGAGGGCCCTGGCACCCGTCCGCGTCAACTGGAT GTTTTTGGTATCACGTACGATCGCGCGTCTCGTGAGGAGCTGACTGAGATGACGGACGCG AGCATGGATGCGCTGTCTCGTGTTCGTGCGAACTGACCGGTCCGCCACTAAAGCGTAC CGCGAACGCGACATCAACTTCGATCCACTGGTTATGAAAAAGACCCGGCACTGCTGGAA GGCTGTTTCCAGTCTGACAAATACTTTCGTGATTGCGAAGGCCGCGTGCGCGAAGCGTAT CGTTTCCGCGGCATTGAATCCGGCGCGTTCGCGTCCGGAAGACTATCTGCGCCTGGAA AAGCAGATCGAAGATTGTCAGTCCGTATCCGTACACATCCGTCGTGGCGACTACCTGGAC GAATCTCATGGTGGTCTGTACACCGGCATTGTAAGTGGAGGCGTACTATAAAGAGGCTTTT GCTCGCATGGAACGTCTGGTTCGCGGCGCACGTTTCTTCTGTCTCTAACGATCCAGAA TGGACTCGTGAGCACTTTGAGAGCAAGAACTGCGTTCTGGTTGAAGGTAGCACCGAAGAC ACGGGTACATGGACCTGTACCTGATGAGCCGCTGCCGCCACAATATTATTGCCAACTCT TCTTTCAGCTGGTGGGCGCTTGGCTGAATGAGAACCCTGAGAAAAAGTCATCGCACCG GCTAAATGGCTGAACGGTCTGTGAGTGCCGTGATATCTATACCGAACGCATGATTCTGTCTG TAGCTCGAGTGACTGACTG	278
FutR	CAGTCAGTCAGAATTCAAGAAGGAGATATACATATGATCATTGTTTCGTCTGAAAGGCGGT CTGGGCAACCAACTGTCTCAGTATGCACTGGGCGGTAAGATCGCGCATCTGCACAATACC GAACTGAAACTGGACACCCTTGGTTCACCACTATCTCCTCCGACACTCCACGTACCTAC CGTCTGAACAATTATAACATCATCGGCATATTGCATCCGCAAAGGAAATCCAGCTGATC GAACGTGGTTCGCGCGCAAGGCCGTGGCTACCTGCTGTCTAAAAATTTCTGATCTGCTGACT CCGATGTACCGTCTGATACCTACGTCCGTGAACGTATGCATACCTTCGATAAAGCTATCCTG ACCGTTCCCGGACAACGTGTACCTGGATGGTTACTGGCAGACCGAAAAAGTACTTCAAAGAC ATCGAAGAAATCCTGCGCCGTGAGGTTACGCTGAAAGATGAACCGGATAGCATCAACCTG GAAATGGCTGAACGATTTAGGCTTGCCACAGCGTTTCCCTGCACGTGCGTCTGGCGAC TACGTTTCCAACCCGACCCTCAACAATTCACGGCTGTTGCTCCATTGACTACTACAAC CGCGCTATCTCTCTGATTGAAGAAAAAGTGGATGACCCGCTTTCTTTATTTTTTCTGAC GATCTGCCGTGGGCTAAAGAAAACCTGGACATCCCTGGCGAAAAAACCTTCGTTGCGCAT AACGGCCCCGAAAAAGAGTATTGCGATCTGTGGCTGATGTCTCTGTGCCAGCACCATATC ATCGCAAACTCTTCTTTCAGCTGGTGGGGTGCCCTGGCTGGGTCAAGACGCCGAAAAGATG GTGATCGCGCCGCTGCTGGGCCCTGTCCGAGAGCTTTGACACTTCTGACATCATTCCG GACTCTTGGATTACTATCTAGCTCGAGTGACTGACTG	279
FutS	CAGTCAGTCAGAATTCAAGAAGGAGATATACATATGGTACGCATTGTGGAAATCATCGGC GGTCTGGGTAACCAGATGTTCCAGTACGCATTCTCCCTGTACCTGAAAAACAAATCTCAC ATCTGGGACCGTCTGTATGTGGACATCGAGGCGATGAAAACTACGATCGTCACTATGGT CTGGAACCTGGAGAAAGTTTTCAATCTGAGCCTGTGTCCAATCTCTAACCGTCTGCACCGC AACCTGCAAAAACGCTCCTTCGCAAAACACTTTGTAAAGAGCCTGTACGAGCACTCTGAA	280

	TGCGAGTTCGACGAACCGGTGTACCGTGGCCTGCGTCCTTATCGCTATTATCGCGGCTAC TGGCAAAACGAAGTTACTTCGTTGATATTGAACCGATGATCCGTGAGGCTTTTCAGTTTC AACGTTAACATCCTGAGCAAAAAGACTAAAGCGATCGCATCCAAAATGCGCCGTGAAGTCG TCCGTATCTATCCATGTTTCGCGGTGGTGATTACGAAAACCTGCGCGAAGCGAAAGCGATG CATGGCGGTATTTGTTCTCTGGACTATTACCAAAAGCGATCGACTTCATCCGCCAGCGT CTGGATAATAACATCTGTTTCTATCTGTTCTCCGACGATATCAATTGGGTAGAAGAAAAC CTGCAACTGGAAAACCGTTGCATCATCGACTGGAACCAGGGCGAAGATAGCTGGCAGGAC ATGTACCTGATGAGCTGCTGCCGCCACCACATTATCGCAAACAGCTCTTTCTCCTGGTGG GCGGCATGGCTGAATCCAAACAAGAACAAAATCGTACTGACCCCGAACAAATGGTTCAAC CATACTGACGCAGTGGGTATCGTCCCAAAGTCCTGGATTAAAATTCTGTGTTTTAGCTC GAGTGACTGACTG	
FutU	CAGTCAGTCAGAATTCAGAAGGAGATATACATATGAAAATCGTTAAAATCCTGGGCGGC CTGGGTAACCAGATGTTTCAGTACGCCCTGTTCTGTCTCTGAAAGAACGCTTCCCGCAT GAACAGGTGATGATTGACACCAGCTGCTTCCGCAATTACCCACTGCACAACGGTTTCGAA GTGGATCGTATCTTCGCCCAGAAAGCACCAGGTTGCCTCTTGGCGTAACATCCTGAAGGTT GCC'TACCCGTACCCGAAC'TACCGTTTCTGGAAAATCGGTAAATACATCCTGCCTAAACGT AAAACCATGTGTGTAGAGCGTAAAACTTCAGCTTTGACGCCGAGTCCTGACCCGTAAA GGCGATTGCTACTATGATGGCTACTGGCAGCATGAGGAATATTTCTGTGATATGAAAGAA ACGATTTGGGAGGCTTTCTCCTTCCCTGAGCCGGTTGATGGTCGTAACAAGGAGATCGGT GCCCTGCTACAGGCATCTGATAGCGCTTCCCTGCACGTTTCGTCGCGGTGACTACGTGAAC CACCCACTGTTTTCGTGGTATTTGTGACCTGGACTATTATAAACGTGCCATCCACTACATG GAAGAACGCGTCAACCCACAGCTGTAAGTTCGCTTTTCAGCAACGATATGGCCTGGTGCAG TCCACCTGCGTGCACTGCTGCCAGGCAAGAAGTAGTTTATGTTGACTGGAACAAGGGT GCGGAATCTTACGTTGATATGCGTCTGATGAGCCTGTGCCGTCAACATCATCGCTAAC TCTTCTTTTACGCTGGTGGGGCGCATGGCTGAACCGTAACCCGAGAAAGTGGTGGTAGCG CCGGAACGTTGGATGAACAGCCCGATTGAAGACCCAGTGAGCGACAAATGGATTAAACTG TAGCTCGAGTGACTGACTG	281
FutV	CAGTCAGTCAGAATTCAGAAGGAGATATACATATGATCATCATCCAGCTGAAAGGTGGC CTGGGCAACCAATGTTCCAGTACGCGCTGTACAAAATCCCTGAAAAACGTGGTAAAGAA GTTAAATTTGATGACAAAAC'TGGCTTCGTGAACGACAAAATGCGTATCCCGGTACTGTCC CGTTGGGGTGTGAGTACGATCGTGCAACCGACGAAGAGATTATTAACCTGACCGACTCC AAAATGGACCTGTTCTCTCGCATCCGCCGTAAACTGACTGGCCGCAAAACGTTCCGTATC GACGAAGAATCCGTAATTTCAACCCGGAATCCTGAAAAAGAGAACGCTTATCTGGTG GGTTATTGGCAGTGCGACAAGTACTTCGACGACAAAGATGTGGTTTCGCGAAATTCGTGAA GCGTTTCGAGAAAAAACCGCAGGAGCTGATGACCGACGCCAGCTCTTGGTCTACTCTACAG CAGATTGAATGCTGCGAGTCCGTATCCCTGACGTACGTCTGATTACGTGGACGAG GAACATATTCATATCCATAACATCTGTACGGAAAAATACTATAAAAAACGCCATTGATCGT GTGCGTAAACAGTACCCGAGCGCAGTGTTCCTTATCTTACCCGATGATAAAGAATGGTGC CGCGACCACTTTAAAGGTCCGAAC'TTCATCGTAGTCGAACTGGAAGAAGGCGACGGTACC GACATCGCTGAAATGACTCTGATGTCCCGCTGTAAACATCACATCATCGCTAATTCTAGC TTTAGCTGGTGGGCGGCGTGGCTGAACGACTCCCGGAAAAAATCGTGATCGCTCCTCAG AAATGGATTAAACAACCGCGACATGGACGATATTTACACCGAGCGTATGACTAAAAATCGCA CTGTAGCTCGAGTGACTGACTG	282
FutW	CAGTCAGTCAGAATTCAGAAGGAGATATACATATGCGCCTGGTTAAAATGATCGGCGGT CTGGGTAATCAGATGTTTCATCTACGCGTTTTACCTACAGATGCGTAAGCGTTTCTCCAAC GTTTCGTATCGACCTGACCGATATGATGCAC'TACAACGTACACTATGGCTACGAACTGCAC AAAGTTTTTCGCTCTGCCGCGCACCGAGTTCTGTATGAACCAGCCTCTGAAAAAGGTTCTG GAGTTCTGTTCCTTCCGTACCATTGTTGAACGTAAACAGCACGGTTCGTATGGAGCCGTAT ACTTGCCAGTATGTTTGGCCGCTGGTTTACTTTAAGGGCTTCTATCAGTCCGAACGTTAC TTCTCCGAAGTTAAGGACGAAGTTTCGTGAGTGTTCACCTTCAATCCGGCACTGGCGAAT CGTTCTTCCCAACAGATGATGGAACAGATCCAGAATGATCCTCAGGCTGTCTCTATCCAC ATCCGTCTGTGGCGACTATCTGAATCCGAAGCACTACGACACTATCGGTTGTATCTGTGAG CTGCCGTATTACAAGCACGCCGTTTTCCGAAATTAAGTACGTTTCTAACCCCTCACTTT TACGTTTTTCTCCGAAGACCTGGATTGGGTCAAAGCAAACCTGCCGCTGGAAAAACGCACAG TACATCGATTGGAACAAAGGCGCAGATAGCTGGCAGGATATGATGCTGATGAGCTGTTGC AAACACCACATTATCTGTAACCTCACCTTTAGCTGGTGGGCGGCGTGGCTGAACCCATCT GTGCAAAAAACCGTGATCATGCCGGAACAGTGGACGTCTCGTCAAGATTCCTGTGGACTTT GTGGCTAGCTGTGGCCGTTGGGTCCGTGTTAAAACGGAGTAGCTCGAGTGACTGACTG	283
FutX	CAGTCAGTCAGAATTCAGAAGGAGATATACATATGCGTCTGATCAAGATGATCGGCGGC CTGGGTAACCAGATGTTTATCTACGCGTTCTACCTGAAAATGAAACACCATTACCCGGAT ACGAACATCGATCTGTCTGACATGGTTTATTATAAAGTTCAACAGGTTATGAGATGAAC	284

	CGTATCTTTGACCTGAGCCAGACTGAATTTTGCATCAACCGTACCCTGAAAAAATCCTG GAGTTCCTGTTCTTCAAAAAATCTACGAACGTCGCCAGGACCCGTCTACTCTGTATCCA TACGAAAAACGTTATTTTGGCCGCTGCTGTACTTTAAAGGTTTCTACCAGTCTGAACGC TTCTTCTTCGATATCAAAGACGACGTTTCGTAAAGCCTTCTCTTTTAACCTGAACATCGCT AACCCGGAAGCCTGGAAGTGTGAAACAGATCGAAGTTGACGACCAAGCTGTTTCTATC CACATCCGCCGTGGTGACTACCTGCTGCCGCGTCACTGGGCAAACACGGGTTCGCTGTGC CAGCTGCCGTATTACAAGAACGCGATCGCGGAAATGGAGAACCGTATTACTGGCCCGAGC TACTACGTGTTCTCTGATGATATCTCTTGGGTAAAGAAAACATCCCGCTGAAGAAAGCG GTCTACGTGACGTGGAACAAGGGCGAAGACAGCTGGCAGGATATGATGCTGATGAGCCAC TGTCGTCACCAATTATCTGTAATTCTACGTTCTCCTGGTGGGGTGCTTGGCTGAACCCA CGTAAAGAGAAAAATCGTCATCGCGCCGTGTGCTGGTTCCAGCATAAAGAAACCCCGGAC ATGTACCCGAAAGAATGGATCAAAGTACCGATTAACTAGCTCGAGTGACTGACTG	
FutZ	CAGTCAGTCAGAATTCAAGAAGGAGATATACATATGTATTCTTGCCTGTCTGGTGGCCTG GGTAACCAAATGTTTCAATACGCAGCAGCGTATATCCTGAAGCAGTATTTTCAGTCTACC ACTCTGGTCTTGGATGATAGCTATTACTATTCCAGCCGAAACGTGATACCGTTTCGTAGC CTGGAACCTGAATCAGTTCAACATCTCTTATGATCGTTTTCAGCTTCGCGGATGAAAAAGAG AAGATCAAACCTGCTGCGCAAATTCAAACGTAACCCGTTCCCTAAACAGATTTCCGAGATC CTGTCTATTGCGCTGTTTCGGCAAATACGCGCTGTCCGACCGTGCATTTTACACCTTCGAA ACTATCAAAAACATCGACAAAGCGTGCTGTCTCCTTTTACCAGGACGCCGATCTGCTG AATAAATATAAGCAGCTGATCCTGCCGCTGTTTGAAGTGCAGGATGACCTGCTGGATATC TGCAAGAACCTGGAAGTGTATTCCTGATCCAACGCAGCAACAATACCACTGCACTGCAT ATCCGCCGTGGCGACTACGTGACCAACCAGCAGCCGCGAAATACCAACGGCGTGCTGGAC ATCAGCTACTATAACCACGCAATGGAATACGTGGAACGTGAACGCGGCAAACAGAACTTC ATTATCTTCAGCGATGATGTACGTTGGGCACAGAAAGCGTTTCTGGAGAACGATAATTGC TACGTGATTAACAACCTCCGACTACGATTTCTCTGCGATCGATATGATCTGATGTCTCTG TGCAAAAACAACATCATCGCAAATTCACCTACTCCTGGTGGGGTGCGTGGCTGAACAAA TACGAGGACAAACTGGTTATCTCTCCGAAACAATGGTTTCTGGGTAACAACGAAACCTCT CTGCGTAACGCGTCTTGGATCACCTGTAGCTCGAGTGACTGACTG	285
FutZA	CAGTCAGTCAGAATTCAAGAAGGAGATATACATATGCGTCTGATCAAGATGACCGGTGGC CTGGGTAACAGATGTTTCATCTACGCGTTTATCTGCGTATGAAAAACGTTATCCGAAA GTTTCGTATTGATCTGTCTGATATGGTTTATTATCACGTTTACCACGGCTATGAAATGCAC CGTGTCTTCAATCTGCCGCACACCGAATTTTGCATCAACCAGCCGCTGAAAAAAGTGATC GAGTTCTCTGTTTTTCAAAAAGATTTACGAACGTAAACAGGACCTAATTCTCTGCGTGCA TTCGAGAAGAAGTATCTGTGGCCGCTGCTGTACTTCAAAGGTTTCTATCAATCTGAGCGC TTCTTTGCTGACATCAAAGACGAGGTTTCGTAAAGCATTCACCTTTGACTCTTCTAAAGTG AACGCTCGCTCTGCCGAAGTGTGCGTGCCTGGATGCCGATGCTAACGCGGTTAGCCTG CACATTCGTGCGCGTGACTATCTACAGCCGCAGCATTGGGCTACCACTGGTTCTGTCTGC CAGCTGCCGTACTACCAGAACGCGATCGCTGAAATGAACCGTCGCGTTGCTGCCCCGAGC TACTACGTTTTTACGCGATGACATCGCGTGGGTGAAGGAAAACATCCCTCTACAGAACGCA GTGTACATCGACTGGAATAAAGGCGAAGAAAGCTGGCAGGATATGATGCTGATGAGCCAC TGCCGCCACCAATTATCTGTAACAGCACCTTCTCTTGGTGGGGCGCGTGGCTGGACCCG CACGAGGACAAAATTGTAATCGTTCCGAATCGTTGGTTCCAGCATTGCGAAACTCCTAAC ATCTATCCGCGAGGCTGGGTGAAAGTTGCGATTAAATTAGCTCGAGTGACTGACTG	286

In any of the methods described herein, the $\alpha(1,2)$ fucosyltransferase genes or gene products may be variants or functional fragments thereof. A variant of any of genes or gene products disclosed herein may have 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% sequence identity to the nucleic acid or amino acid sequences described herein.

Variants as disclosed herein also include homolog, orthologs, or paralogs of the genes or gene products described herein that retain the same biological function as the genes or gene products specified herein. These variants can be used interchangeably with the genes recited in these methods. Such variants may demonstrate a percentage of homology or

identity, for example, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity conserved domains important for biological function, preferably in a functional domain, e.g. catalytic domain.

The term "% identity," in the context of two or more nucleic acid or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same, when compared and aligned for maximum correspondence, as measured using one of the following sequence comparison algorithms or by visual inspection. For example, % identity is relative to the entire length of the coding regions of the sequences being compared, or the length of a particular fragment or functional domain thereof.

For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are input into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. The sequence comparison algorithm then calculates the percent sequence identity for the test sequence(s) relative to the reference sequence, based on the designated program parameters.

Percent identity is determined using search algorithms such as BLAST and PSI-BLAST (Altschul et al., 1990, J Mol Biol 215:3, 403-410; Altschul et al., 1997, Nucleic Acids Res 25:17, 3389-402). For the PSI-BLAST search, the following exemplary parameters are employed: (1) Expect threshold was 10; (2) Gap cost was Existence:11 and Extension:1; (3) The Matrix employed was BLOSUM62; (4) The filter for low complexity regions was "on".

Changes can be introduced by mutation into the nucleic acid sequence or amino acid sequence of any of the genes or gene products described herein, leading to changes in the amino acid sequence of the encoded protein or enzyme, without altering the functional ability of the protein or enzyme. For example, nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues can be made in the sequence of any of sequences expressly disclosed herein. A "non-essential" amino acid residue is a residue at a position in the sequence that can be altered from the wild-type sequence of the polypeptide without altering the biological activity, whereas an "essential" amino acid residue is a residue at a position that is required for biological activity. For example, amino acid residues that are conserved among members of a family of proteins are not likely to be amenable to mutation. Other amino acid residues, however, (e.g., those that are poorly conserved among members of the protein family) may not be as essential for activity and thus are more likely to be

amenable to alteration. Thus, another aspect of the invention pertains to nucleic acid molecules encoding the proteins or enzymes disclosed herein that contain changes in amino acid residues relative to the amino acid sequences disclosed herein that are not essential for activity (*i.e.*, fucosyltransferase activity).

An isolated nucleic acid molecule encoding a protein essentially retaining the functional capability compared to any of the genes described herein can be created by introducing one or more nucleotide substitutions, additions or deletions into the corresponding nucleotide sequence, such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein.

Mutations can be introduced into a nucleic acid sequence by standard techniques such that the encoded amino acid sequence is altered, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. Certain amino acids have side chains with more than one classifiable characteristic. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, tryptophan, cysteine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tyrosine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid residue in a given polypeptide is replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of a given coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for given polypeptide biological activity to identify mutants that retain activity. Conversely, the invention also provides for variants with mutations that enhance or increase the endogenous biological activity. Following mutagenesis of the nucleic acid sequence, the encoded protein can be expressed by any recombinant technology known in the art and the activity of the protein can be determined. An increase, decrease, or elimination of a given biological activity of the variants disclosed herein can be readily measured by the ordinary person skilled in the art, *i.e.*, by measuring the capability for mediating oligosaccharide modification, synthesis, or degradation (via detection of the products).

The present invention also provides for functional fragments of the genes or gene products described herein. A fragment, in the case of these sequences and all others provided herein, is defined as a part of the whole that is less than the whole. Moreover, a fragment ranges in size from a single nucleotide or amino acid within a polynucleotide or polypeptide sequence to one fewer nucleotide or amino acid than the entire polynucleotide or polypeptide sequence. Finally, a fragment is defined as any portion of a complete polynucleotide or polypeptide sequence that is intermediate between the extremes defined above.

For example, fragments of any of the proteins or enzymes disclosed herein or encoded by any of the genes disclosed herein can be 10 to 20 amino acids, 10 to 30 amino acids, 10 to 40 amino acids, 10 to 50 amino acids, 10 to 60 amino acids, 10 to 70 amino acids, 10 to 80 amino acids, 10 to 90 amino acids, 10 to 100 amino acids, 50 to 100 amino acids, 75 to 125 amino acids, 100 to 150 amino acids, 150 to 200 amino acids, 200 to 250 amino acids, 250 to 300 amino acids, 300 to 350 amino acids, 350 to 400 amino acids, 400 to 450 amino acids, or 450 to 500 amino acids. The fragments encompassed in the present invention comprise fragments that retain functional fragments. As such, the fragments preferably retain the catalytic domains that are required or are important for functional activity. Fragments can be determined or generated by using the sequence information herein, and the fragments can be tested for functional activity using standard methods known in the art. For example, the encoded protein can be expressed by any recombinant technology known in the art and the activity of the protein can be determined. The biological function of said fragment can be measured by measuring ability to synthesize or modify a substrate oligosaccharide, or conversely, to catabolize an oligosaccharide substrate.

Within the context of the invention, “functionally equivalent”, as used herein, refers to a gene or the resulting encoded protein variant or fragment thereof capable of exhibiting a substantially similar activity as the wild-type fucosyltransferase. Specifically, the fucosyltransferase activity refers to the ability to transfer a fucose sugar to an acceptor substrate via an α -(1,2)-linkage. As used herein, “substantially similar activity” refers to an activity level within 5%, 10%, 20%, 30%, 40%, or 50% of the wild-type fucosyltransferase.

To test for lactose-utilizing fucosyltransferase activity, the production of fucosylated oligosaccharides (*i.e.*, 2'-FL) is evaluated in a host organism that expresses the candidate enzyme (or syngene) and which contains both cytoplasmic GDP-fucose and lactose pools. The production of fucosylated oligosaccharides indicates that the candidate enzyme-encoding sequence functions as a lactose-utilizing α (1,2)fucosyltransferase.

Engineering of *E. coli* to produce human milk oligosaccharide 2'-FL

Described herein is a gene screening approach, which was used to validate the novel α (1,2) fucosyltransferases (α (1,2) FTs) for the synthesis of fucosyl-linked oligosaccharides in metabolically engineered *E. coli*. Of particular interest are α (1,2) FTs that are capable of the synthesis of the HMOS 2'-fucosyllactose (2'-FL). 2'-FL is the most abundant fucosylated oligosaccharide present in human milk, and this oligosaccharide provides protection to newborn infants against infectious diarrhea caused by bacterial pathogens such as *Campylobacter jejuni* (Ruiz-Palacios, G.M., et al. (2003). *J Biol Chem* 278, 14112-120; Morrow, A.L. et al. (2004). *J Pediatr* 145, 297-303; Newburg, D.S. et al. (2004). *Glycobiology* 14, 253-263). Other α (1,2) FTs of interest are those capable of synthesis of HMOS lactodifucotetraose (LDFT), laco-N-fucopentaose I (LNFI), or lacto-N-difucohexaose I (LDFH I).

The synthetic pathway of fucosyl oligosaccharides of human milk is illustrated in FIG. 1. Structurally, 2'-FL consists of a fucose molecule α 1,2 linked to the galactose portion of lactose (Fuc α 1-2Gal β 1-4Glc). An α (1,2) FT from *H. pylori* strain 26695 termed FutC has been utilized to catalyze the synthesis of 2'-FL in metabolically engineered *E. coli* (Drouillard, S. et al. (2006). *Angew Chem Int Ed Engl* 45, 1778-780).

Candidate α (1,2) FTs (*i.e.*, syngenes) were cloned by standard molecular biological techniques into an expression plasmid. This plasmid utilizes the strong leftwards promoter of bacteriophage λ (termed P_L) to direct expression of the candidate genes (Sanger, F. et al. (1982). *J Mol Biol* 162, 729-773). The promoter is controllable, e.g., a trp-cl construct is stably integrated into the *E. coli* host's genome (at the ampC locus), and control is implemented by adding tryptophan to the growth media. Gradual induction of protein expression is accomplished using a temperature sensitive cl repressor. Another similar control strategy (temperature independent expression system) has been described (Mieschendahl et al., 1986, *Bio/Technology* 4:802-808). The plasmid also carries the *E. coli* *rcsA* gene to up-regulate GDP-fucose synthesis, a critical precursor for the synthesis of fucosyl-linked oligosaccharides. In addition, the plasmid carries a β -lactamase (*bla*) gene for maintaining the plasmid in host strains by ampicillin selection (for convenience in the laboratory) and a native *thyA* (thymidylate synthase) gene as an alternative means of selection in *thyA*⁻ hosts. Alternative selectable markers include the proBA genes to complement proline auxotrophy (Stein et al., (1984), *J Bacteriol* 158:2, 696-700 (1984) or purA to complement adenine auxotrophy (S. A. Wolfe, J. M. Smith, *J Biol Chem* **263**, 19147-53

(1988)). To act as plasmid selectable markers each of these genes are first inactivated in the host cell chromosome, then wild type copies of the genes are provided on the plasmid.

Alternatively a drug resistance gene may be used on the plasmid, e.g. beta-lactamase (this gene is already on the expression plasmid described above, thereby permitting selection with ampicillin). Ampicillin selection is well known in the art and described in standard manuals such as Maniatis et al., (1982) Molecular cloning, a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring, NY.

The nucleic acid sequence of such an expression plasmid, pEC2-(T7)FutX-rcsA-thyA (pG401) is provided below. The underlined sequence represents the FutX syngene, which can be readily replaced with any of the novel $\alpha(1,2)$ FTs described herein using standard recombinant DNA techniques.

TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGT
CTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTGGCGGGTGTGGGG
CTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATATGCGGTGTGAAATACC
GCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCCTCCTCAACCTGTATATTGTAACACCGCC
CAATGGGAGCTGTCTCAGGTTTGTTCCTGATTGGTTACGGCGCGTTTCGCATCATTTGTTGAGTTTTTC
CGCCAGCCCGACGCGCAGTTTACCGGTGCCTGGGTGCAGTACATCAGCATGGGGCAAATTCCTTCCAT
CCCGATGATTGTGCGGGTGTGATCATGATGGTCTGGGCATATCGTCGCAGCCACAGCAACACGTTT
CCTGAGGAACCATGAAACAGTATTTAGAACTGATGCAAAAAGTGCTCGACGAAGGCACACAGAAAAAC
GACCGTACCGGAACCGGAACGCTTTCCATTTTTGGTCATCAGATGCGTTTTAACCTGCAAGATGGATT
CCCGCTGGTGACAACTAAACGTTGCCACCTGCGTTCCATCATCCATGAACTGCTGTGGTTTCTGCAGG
GCGACACTAACATTGCTTATCTACACGAAAAAATGTCAACCATCTGGGACGAATGGGCCGATGAAAAC
GGCGACCTCGGGCCAGTGTATGGTAAACAGTGGCGCGCCTGGCCAACGCCAGATGGTCGTATATTGA
CCAGATCACTACGGTACTGAACCAGCTGAAAAACGACCCGGATTTCGCGCCGCATTATTGTTTCAGCGT
GGAACGTAGGCGAACTGGATAAAATGGCGCTGGCACCGTGCCATGCATTCTTCCAGTTCTATGTGGCA
GACGGCAAACTCTCTTGCCAGCTTTATCAGCGCTCCTGTGACGTCTTCTCGGCCTGCCGTTCAACAT
TGCCAGCTACGCGTTATTGGTGCATATGATGGCGCAGCAGTGCATCTGGAAGTGGGTGATTTTGTCT
GGACCGGTGGCGACACGCATCTGTACAGCAACCATATGGATCAAACCTCATCTGCAATTAAGCCGCGAA
CCGCGTCCGCTGCCGAAGTTGATTATCAAACGTAAACCCGAATCCATCTTCGACTACCGTTTCGAAGA
CTTTGAGATTGAAGGCTACGATCCGCATCCGGGCATTAAAGCGCCGGTGGCTATCTAATTACGAAACA
TCCTGCCAGAGCCGACGCCAGTGTGCGTCGGTTTTTTTACCCTCCGTTAAATTCTTCGAGACGCCTTC
CCGAAGGCGCCATTTCGCCATTAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGC
TATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCC
CAGTCACGACGTTGTAAAACGACGGCCAGTGCCAAGCTTTCTTTAATGAAGCAGGGCATCAGGACGGT
ATCTTTGTGGAGAAAGCAGAGTAATCTTATTACGCTGACTGGTGGGAAACCACCAGTCAGAATGTGT
TAGCGCATGTTGACAAAAATACCATTAGTCACATTATCCGTCAGTCGGACGACATGGTAGATAACCTG
TTTATTATGCGTTTTTGATCTTACGTTTAAATATTACCTTTATGCGATGAAACGGTCTTGGCTTTGATAT
TCATTTGGTCAGAGATTTGAATGGTTCCTGACCTGCCATCCACATTGCAACATACTCGATTGCGTT
CGGCTCAATGATAACGTCGGCATATTTAAAAACGAGGTTATCGTTGTCTCTTTTTTCAGAATATCGCC
AAGGATATCGTCGAGAGATTCCGGTTTAAATCGATTTAGAACTGATCAATAAATTTTTTCTGACCAATA
GATATTCATCAAATGAACATTGGCAATTGCCATAAAAACGATAAATAACGTATTGGGATGTTGATTA
ATGATGAGCTTGATACGCTGACTGTTAGAAGCATCGTGGATGAAACAGTCCTCATTAAATAAACACCAC
TGAAGGGCGCTGTGAATCACAAGCTATGGCAAGGTCAACCGGTTTCAATGTGCTTGATTTCTCTTT
TTTTAACCCCTCTACTCAACAGATAACCGGTTAAACCTAGTCGGGTGTAACCTACATAAATCCATAATA
ATCGTTGACATGGCATACCTCACTCAATGCGTAACGATAAATCCCCTTACCTGAATATTTTCATCATG
ACTAAACGGAACAACATGGGTACCTAATGCGCCACTCTCGCGATTTTTTCAGGCGGACTTACTATCCC
GTAAAGTGTGTATAATTTGCCTGGAATTGTCTTAAAGTAAAGTAAATGTTGCGATATGTGAGTGAGC
TTAAACAAATATTTGCTGCAGGAGTATCCTGGAAGATGTTGCTAGAAGCTTACTGCTCACAAGAAA

AAAGGCACGTCATCTGACGTGCCTTTTTTATTTGTACTACCCTGTACGATTACTGCAGCTCGAGCTAG
TTAATCGGTACTTTGATCCATTCTTTCCGGGTACATGTCCGGGGTTTCTTTATGCTGGAACCAGCGACA
CGGCGCGATGACGATTTTCTCTTTACGTGGGTTCAGCCAAGCACCCACCAGGAGAACGTAGAATTAC
AGATAATGTGGTGACGACAGTGGCTCATCAGCATCATATCCTGCCAGCTGTCTTCGCCCTTGTTCAC
GTCACGTAGACCGCTTTCTTCAGCGGGATGTTTTCTTTAACCCAAGAGATATCATCAGAGAACACGTA
GTAGCTCGGGCCAGTAATACGGTTCTCCATTTCCGCGATCGCGTTCTTGTAATACGGCAGCTGGCACA
CGGAACCCGTGTTTTGCCAGTGACGCGGCAGCAGGTAGTCACCACGGCGGATGTGGATAGAAACAGCT
TGGTCGTCAACTTCGATCTGTTTCAGCAGTTCCAGGCTTTCCGGGTAGCGATGTTTCAGGTTAAAGA
GAAGGCTTTACGAACGTCTCTTTGATATCGAAGAAGAAGCGTTTCAGACTGGTAGAAACCTTTAAAGT
ACAGCAGCGGCCAAAATAACGTTTTTCGTATGGATACAGAGTAGACGGGTCTGGCGACGTTTCGTAG
ATTTTTTTGAAGAACAGGAACCTCCAGGATTTTTTTTCAGGGTACGGTTGATGCAAAATTCAGTCTGGCT
CAGGTCAAAGATACGGTTCATCTCATAACCGTTGTGAACTTTATAATGAACCATGTACAGACAGATCGA
TGTTTCGTATCCGGGTAAATGGTGTTCATTTTCAGGTAGAACGCGTAGATAAACATCTGGTTACCCAGG
CCGCCGATCATCTTGATCAGACGCATATGTATATCTCCTTCTTGAATTCTAAAAATTGATTGAATGTA
TGCAAAATAAATGCATACACCATAGGTGTGGTTTAATTTGATGCCCTTTTTTCAGGGCTGGAATGTGTAA
GAGCGGGGTATTTATGCTGTTGTTTTTTTGTACTCGGGAAGGGCTTTACCTCTTCCGCATAAACGC
TTCCATCAGCGTTTATAGTTAAAAAATCTTTCGGAACGGTTTTTGCGCTTACCCCAACCAACAGGGG
ATTTGCTGCTTTCCATTGAGCCTGTTTCTCTGCGCGACGTTTCGCGGCGGCGTGTGTCATCCATCT
GGATTCTCCTGTGCTAGCTTTGGTGGTGTGTGGCAGTTGTAGTCCTGAACGAAAACCCCCGCGAT
TGGCACATTGGCAGCTAATCCGGAATCGCACTTACGGCCAATGCTTCGTTTCGTATCACACACCCCAA
AGCCTTCTGCTTTGAATGCTGCCCTTCTTCAGGGCTTAATTTTTAAGAGCGTCACCTTCATGGTGGTC
AGTGCGTCCTGCTGATGTGCTCAGTATCACCGCCAGTGGTATTTATGTCAACACCGCCAGAGATAATT
TATCACCGCAGATGGTTATCTGTATGTTTTTTATATGAATTTATTTTTTGCAGGGGGGCATTGTTTGG
TAGGTGAGAGATCAATTCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGG
CGCTCTTCCGCTTCCCTCGCTCACTGACTCGCTGCGCTCGGTCTCGGCTGCGGCGAGCGGTATCAGC
TCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAA
AAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCC
CCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATA
CCAGGCGTTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACC
TGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCG
GTGTAGGTGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTTCAGCCCGACCGCTGCGCCTT
ATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTG
GTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTAC
GGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGA AAAAGAGT
TGGTAGCTCTTGATCCGGCAAACAAACACCGCTGGTAGCGGTGGTTTTTTTTGTTTTGCAAGCAGCAGA
TTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGG
AACGAAAACCTACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTT
AAATTA AAAATGAAGTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAAT
GCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCCTGACTCCCC
GTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGA
CCCACGCTCACC GGCTCCAGATTTATCAGCAATAAACAGCCAGCCGGAAGGGCCGAGCGCAGAAGTG
GTCCTGCACTTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCGGGAAGCTAGAGTAAGTAGTTTCG
CCAGTTAATAGTTTTCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGTCTGTTGG
TATGGCTTCATTACGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAA
AAGCGGTTAGCTCCTTCGGTCTCCGATCGTTGTGAGAAGTAAGTTGGCCGAGTGTTTACTCATG
GTTATGGCAGCACTGCATAATTCTCTTACTGTATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGA
GTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATAC
GGGATAATACCGCGCCACATAGCAGAACTTTAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGA
AACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTTCGATGTAAACCACTCGTGACCCAACTGATC
TTCAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAA
AGGGAATAAGGGCGACACGGAATGTTGAATACTCATACTCTTCTTTTCAATATTATTGAAGCAT
TATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGT
TCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGCTAAGAAACCATTATTATCATGACATTAACCT
ATAAAAATAGGCGTATCACGAGGCCCTTTCGTC (SEQ ID NO: 287)

The expression constructs were transformed into a host strain useful for the production of 2'-FL. Biosynthesis of 2'-FL requires the generation of an enhanced cellular pool of both lactose and GDP-fucose (FIG. 2). The wild-type *Eschericia coli* K12 prototrophic strain W3110 was selected as the parent background to test the ability of the candidates to catalyze 2'-FL production (Bachmann, B.J. (1972). Bacteriol Rev 36, 525-557). The particular W3110 derivative employed was one that previously had been modified by the introduction (at the *ampC* locus) of a tryptophan-inducible P_{trpB} *cI*⁺ repressor cassette, generating an *E.coli* strain known as GI724 (LaVallie, E.R. et al. (2000). Methods Enzymol 326, 322-340). Other features of GI724 include *lacIq* and *lacPL8* promoter mutations. *E.coli* strain GI724 affords economical production of recombinant proteins from the phage λ P_L promoter following induction with low levels of exogenous tryptophan (LaVallie, E.R. et al. (1993). Biotechnology (N Y) 11, 187-193; Mieschendahl, et al. (1986). Bio/Technology 4, 802-08). Additional genetic alterations were made to this strain to promote the biosynthesis of 2'-FL. This was achieved in strain GI724 through several manipulations of the chromosome using λ Red recombineering (Court, D.L. et al. (2002). Annu Rev Genet 36, 361-388) and generalized P1 phage transduction.

First, the ability of the *E. coli* host strain to accumulate intracellular lactose was engineered by simultaneous deletion of the endogenous β -galactosidase gene (*lacZ*) and the lactose operon repressor gene (*lacI*). During construction of this deletion, the *lacIq* promoter was placed immediately upstream of the lactose permease gene, *lacY*. The modified strain maintains its ability to transport lactose from the culture medium (via LacY), but is deleted for the wild-type copy of the *lacZ* (β -galactosidase) gene responsible for lactose catabolism. Therefore, an intracellular lactose pool is created when the modified strain is cultured in the presence of exogenous lactose. A schematic of the P_{lacIq} *lacY*⁺ chromosomal construct is shown in FIG. 12.

Genomic DNA sequence of the P_{lacIq} *lacY*⁺ chromosomal construct is set forth below (SEQ ID NO: 288):

```
CACCATCGAATGGCGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCAAGTGTAGGCTGGAGC
TGCTTCGAAGTTCCTATACCTTCTAGAGAATAGGAACCTTCGGAATAGGAACCTAAGGAGGAT
ATTTCATATGTACTATTTAAAAACACAACTTTTGGATGTTTCGGTTTATTCTTTTCTTTTACTTTTTATCATG
GGAGCCTACTTCCCGTTTTTCCCGATTTGGCTACATGACATCAACCATATCAGCAAAAGTGATACGGGTATTATT
TTTGCCGCTATTTCTCTGTTCTCGCTATTATTTCAACCGCTGTTTGGTCTGCTTTCTGACAAACTCGGGCTGCGC
AAATACCTGCTGTGGATTATTACCGGCATGTTAGTGATGTTTGCGCCGTTCTTTATTTTATCTTCGGGCCACTG
TTACAATACAACATTTTAGTAGGATCGATTGTTGGTGGTATTTATCTAGGCTTTTGTTTTAAACGCCGGTGCGCCA
GCAGTAGAGGCATTTATTGAGAAAGTCAGCCGTCGCAGTAATTTTGAATTTGGTTCGCGCGCGGATGTTTGGCTGT
GTTGGCTGGGCGCTGTGTGCCTCGATTGTTCGGCATCATGTTTACCACATCAATAATCAGTTTGTCTGGCTGGGC
TCTGGCTGTGCACTCATCCTCGCGTTTTTACTCTTTTTCGCCAAAACCGATGCGCCCTCTTCTGCCACGGTTGCC
AATGCGGTAGGTGCCAACCATTTCGGCATTTAGCCTTAAGCTGGCACTGGAACCTGTTTCTAGACAGCCAAAACCTGTGG
```


TTTTTGTCACTGTATGTTATTGGCGTTTCTGACCTACGATGTTTTTGACCAACAGTTTGCTAATTTCTTTACT
 TCGTTCTTTTGCTACCGGTGAACAGGGTACGCGGTATTTGGCTACGTAACGACAATGGGCGAATTACTTAACGCC
 TCGATTATGTTCTTTGCGCCACTGATCATTAAATCGCATCGGTGGGAAAAACGCCCTGCTGCTGGCTGGCACTATT
 ATGTCTGTACGTATTATTGGCTCATCGTTCCGCACCTCAGCGCTGGAAGTGTTATTTCTGAAAAACGCTGCATATG
 TTTGAAGTACCGTTCTGCTGGTGGGCTGCTTTAAATATATTACCAGCCAGTTTGAAGTGCGTTTTTCAGCGACG
 ATTTATCTGGTCTGTTTCTGCTTCTTTAAGCAACTGGCGATGATTTTTATGTCTGTACTGGCGGGCAATATGTAT
 GAAAGCATCGGTTTCCAGGGCGCTTATCTGGTGTGGGTCTGGTGGCGCTGGGCTTCACCTTAATTTCCGTGTTT
 ACGCTTAGCGGCCCCGGCCCCGCTTTCCCTGCTGCGTCTCAGGTGAATGAAGTCGCTTAAGCAATCAATGTCGGA
 TGCGGCGCGAGCGCCTTATCCGACCAACATATCATAACGGAGTGATCGCATTGTAAATTATAAAAAATTGCCTGAT
 ACGTGCCTTATCAGGCCTACAAGTTCAGCGATCTACATTAGCCGCATCCGGCATGAACAAAGCGCAGGAACAA
 GCGTCGCA

Second, the ability of the host *E. coli* strain to synthesize colanic acid, an extracellular capsular polysaccharide, was eliminated by the deletion of the *wcaJ* gene, encoding the UDP-glucose lipid carrier transferase (Stevenson, G. et al. (1996). J Bacteriol 178, 4885-893). In a *wcaJ* null background GDP-fucose accumulates in the *E. coli* cytoplasm (Dumon, C. et al. (2001). Glycoconj J 18, 465-474). A schematic of the chromosomal deletion of *wcaJ* is shown in FIG. 13.

The sequence of the chromosomal region of *E. coli* bearing the $\Delta wcaJ::FRT$ mutation is set forth below (SEQ ID NO: 289):

GTTTCGGTTATATCAATGTCAAAAACCTCACGCCGCTCAAGCTGGTGATCAACTCCGGGAACGGCGCAGCGGGTCC
 GGTGGTGGACGCCATTGAAGCCCGCTTTAAAGCCCTCGGCGCGCCCGTGGAATTAATCAAAGTGCAACACGCC
 GGACGGCAATTTCCCAACGGTATTCTTAACCCACTACTGCCGGAATGCCGCGACGACCCCGCAATGCGGTCTAT
 CAAACACGGCGCGGATATGGGCATTGCTTTTGATGGCGATTTTGACCGCTGTTTCTGTGTTGACGAAAAAGGGCA
 GTTTATTGAGGGCTACTACATTGTGCGCCTGTTGGCAGAAGCATTCTCGAAAAAATCCCGCGCGAAGATCAT
 CCACGATCCACGTCTCTCCTGGAACACCGTTGATGTGGTGACTGCCGAGGTGGCAGCCCGTAATGTGCAAAAC
 CGGACACGCCTTTATTAAAGAACGTATGCGCAAGGAAGACGCCATCTATGGTGGCGAAATGAGCGCCACCATTA
 CTTCCGTGATTTTCGCTTACTGCGACAGCGGCATGATCCCGTGGCTGCTGGTCCCGAACTGGTGTGCCTGAAAGA
 TAAACGCTGGGCGAACTGGTACGCGACCGGATGGCGCGCTTTCCGGCAAGCGGTGAGATCAACAGCAAACTGGC
 GCAACCCGTTGAGGCGATTAACCGCGTGGAACAGCATTTTAGCCGTGAGGCGCTGGCGGTGGATCGCACCGATGG
 CATCAGCATGACCTTTGCCGACTGGCGCTTTAACCTGCGCACCTCCAATACCGAACCAGGTGGTGCCTGAATGT
 GGAATCGCGCGGTGATGTGCCGCTGATGGAAGCGCGAAGCGCAACTCTGCTGACGTTGCTGAACGAGTAATGTGCG
 GATCTTCCCTTACCCCACTGCGGGTAAGGGGCTAATAACAGGAACAACGATGATTCCGGGGATCCGTCGACCTGC
 AGTTCGAAGTTCTATTCTCTAGAAAAGTATAGGAACCTCGAAGCAGCTCCAGCCTACAGTTAACAAAGCGGCATA
 TTGATATGAGCTTACGTGAAAAAACCATCAGCGGCGCGAAGTGGTTCGGCGATTGCCACGTTGATCATCATCGGCC
 TCGGGCTGGTGCAGATGACCGTGCTGGCGCGGATTATCGACAACCACAGTTTCGGCCTGCTTACCGTGTGCTGG
 TGATTATCGCGCTGGCAGATACGCTTTCTGACTTCGGTATCGCTAACTCGATTATTACGCGAAAAAGAAATCAGTC
 ACCTTGAACACACAGTTGTACTGGCTGAACGTGGGCTGGGGATCGTGGTGTGCGTGGCGGTGTTTTTGTGGA
 GTGATCTCATCGCGACGTGCTGAATAACCCGACCTGGCACCGTTGATTAAACATTATCGCTGGCGTTTGTGG
 TAATCCCCACGGGCAACAGTTCCGCGCGTTGATGCAAAAAGAGCTGGAGTTCAACAAAATCGGCATGATCGAAA
 CCAGCGCGGTGCTGGCGGGCTTCACTTGTACGGTGGTTAGCGCCCATTTCTGGCCGCTGGCGATGACCGCGATCC
 TCGGTTATCTGGTCAATAGTGGGTGAGAACGCTGCTGTTTGGCTACTTTGGCCGCAAAATTTATCGCCCCGGTC
 TGCATTTCTCGCTGGCGTCCGTGGCACCGAATTACGCTTTGGTGCCTGGCTGACGGCGGACAGCATCATCAACT
 ATCTCAATACCAACCTTTCAACGCTCGTGCTGGCGCGTATTCTCGGCGCGGCGTGGCAGGGGGATACAACCTGG
 CGTACAACGTGGCCGTTGTGCCACCGATGAAGCTGAACCAATCATACCCGCGTGTGTTTCCGGCATTTCGCCA
 AAATTCAGGACGATACCGAAAAGCTGCGTGTAACTTCTACAAGCTGCTGTGCGTAGTGGGGATTATCAACTTTC
 CGGCGCTGCTCGGGCTAATGGTGGTGTGGAATAACTTTGTACCGCTGGTCTTTGGTGAGAAGTGGAACAGCATTAT
 TTCCGGTGTGCAATTGCTGTGTGTGGTGGTCTGCTGCGCTCCG

Third, the magnitude of the cytoplasmic GDP-fucose pool was enhanced by the introduction of a null mutation into the *lon* gene. Lon is an ATP-dependant intracellular

protease that is responsible for degrading RcsA, which is a positive transcriptional regulator of colanic acid biosynthesis in *E. coli* (Gottesman, S. & Stout, V. Mol Microbiol 5, 1599-1606 (1991)). In a *lon* null background, RcsA is stabilized, RcsA levels increase, the genes responsible for GDP-fucose synthesis in *E. coli* are up-regulated, and intracellular GDP-fucose concentrations are enhanced. The *lon* gene was almost entirely deleted and replaced by an inserted functional, wild-type, but promoter-less *E. coli lacZ⁺* gene ($\Delta lon::(kan, lacZ^+)$). λ Red recombineering was used to perform the construction. A schematic of the *kan, lacZ⁺* insertion into the *lon* locus is shown in FIG. 14.

Genomic DNA sequence surrounding the *lacZ⁺* insertion into the *lon* region in the *E. coli* strain is set forth below (SEQ ID NO: 290):

```
GTGGATGGAAGAGGTGGAAGAAAGTGGTTATGGAGGAGTGGGTAATTGATGGTGAAAGGAAAGGGTTGGTGATTTA
TGGGAAGGGGGAAGGGGAAGAGGGATGTGGTGAATAATTAAGGATTGGGATAGAATTAGTTAAGGAAAAAGGGGG
GATTTTATGTGGGTTTAAATTTTGGTGTATTGTGGGGTTGAATGTGGGGGAAAGATGGGGATATAGTGAGGTA
GATGTTAATAGATGGGGTGAAGGAGAGTGGTGTGATGTGATTAGGTGGGGGAAATTAAGTAAGAGAGAGGTGTA
TGATTGGGGGATGGGTGGAGGTGGAGTTGGAAGTTGTTATTGTGTAGAAAAGTATAGGAAGTTGAGAGGGGTTTT
GAAGGTGAGGGTGGGGGAAGGAGTGAGGGGGGAAGGGGTGCTAAAGGAAGGGGAAGAGGTAGAAAGGGAGTGGGG
AGAAAGGGTGGTGAGGGGGATGAATGTGAGGTAGTGGGTATGTGAGAGAAGGGGAAAGGGAAGGGGAAAGAGAA
AGGAGGTAGGTTGGAGTGGGGTTAGATGGGGATAGGTAGAGTGGGGGTTTTATGGAGAGGAAGGGGAAGGGGAAT
TGGGAGGTGGGGGGGGTGTGGTAAGGTGGGAAGGGGTGGAAGCTAAAGTGGATGGGTTTGTGGGGGGAAGGA
TGTGATGGGGGAGGGGATGAAGATGTGATGAAGAGAGAGGATGAGGATGCTTTGGGATGATTGAAGAAGATGGAT
TGGAGGGAGGTTGTGGGGGGGGTGGGTGGAGAGGGTATTGGGGTATGAGTGGGGAGAAGAGAGAATGGGGTGGT
GTGATGGGGGGGTGTGGGGGTGTGAGGGAGGGGGGGGGGGTGTGTTTTGTGAAGAGGGAGGTGTGGGGTGGGG
TGAATGAAGTGGAGGAGGAGGGAGGGGGGGTATGGTGGGTGGGAGGAGGGGGGTTGGTTGGGGAGGTGTGGTGG
AGGTTGTGAGTGAAGGGGGAAGGGAGTGGGTGGTATTGGGGGAAGTGGGGGGGAGGATGTGGTGTGATGTGAGG
TTGGTGGTGGGAGAAAGTATGGATGATGGGTGATGGAATGGGGGGGGTGGATAGGGTTGATGGGGCTAGGTGGG
GATTGGAGGAGGAAGGGAAAGATGGGATGGAGGGAGGAGGTAGTGGGATGGAAGGGGGTGTGTTGGATGAGGATG
ATGTGGAGGAAGAGGATGAGGGGGTGGGGGGAGGGGAAGTGTGGGGAGGGTGAAGGGGGGATGGGGGAGGGGGA
GGATGTGGTGGTGAGGGATGGGGATGGGTGGTTGGGAATATGATGGTGGAAAATGGGGGGTTTTGTGGATTGAT
GGAGTGTGGGGGGTGGGTGTGGGGAGGGGTATGAGGAGATAGGGTTGGGTAGGGGTGATATTGGTGAAGAGGT
TGGGGGGGAATGGGGTGGGGGTTGGTGGTGGTTTAGGGTATGGGGGGTGGGGATTGGGAGGGGATGGGGTTGTA
TGGGGTTGTTGAGGAGTTGTTGTAATAAGGGGATGTTGAAGTTGGTATTGGGAAGTTGGTATTGTGTAGAAAAGTA
TAGGAAGTTGGAAGGAGGTGGAGGGTAGATAAAGGGGGGGGTTATTTTTGAGAGGAGAGGAAGTGGTAATGGTAG
GGAGGGGGGTTGGAGTGAATTTGGGGGATAGTGAGGGGGTGGAGGAGTGGTGGGGAGGAATGGGGATATGGAAG
GGGTGGATATTGAGGGATGTGGGTGTGGGGGTGGAGGAGATGGGGATGGGTGGTTTGGATGAGTTGGTGTGTTGA
GTGTAGGGGTGATGTTGAAGTGAAGTGGGGGGGGAGTGGTGTGGGGGATAATTGAATTTGGGGGTGGGGGAG
GGGAGAGGGTTTTGGGTGGGGAAGAGGTAGGGGGTATAGATGTTGAGAATGGGAGATGGGAGGGGTGAAAAGAGG
GGGGAGTAAGGGGTGGGGATAGTTTGTGGGGGGTAAATGGGAGGGAGTTTAGGGGGTGTGGTAGGTGGGGGA
GGTGGGAGTTGAGGGGAATGGGGGGGGATGGGGTGTATGGGTGGGGAGTTGAAGATGAAGGGTAATGGGGATTT
GAGGAGTAGGATGAATGGGGTAGGTTTTGGGGGTGATAAATAAGGTTTTGGGGTGTGGTGGGAGGGGTGAGGGG
TGGTAATGAGGAGGGGATGAGGAAGTGTATGTGGGGTGGAGTGAAGAAGGGTGGTTGGGGGTGGTAATGGGGGG
GGGGGTGGAGGGTTGGAGGGAGGGGTAGGGTGAATGGGGTGGGTGAGTTAGGGGAATGTGGTTATGGAGGG
GTGGAGGGGTGAAGTGTATGGGGAGGGGGGTGAGGAGTTGTTTTTATGGGGAATGGAGATGTGTGAAAGAAAGG
GTGAGTGGGGTTAAATTGGGAAGGGTTATTAGGGAGGTGGATGGAAGAAATGGATTTGGGTGGTGGTGGATGGG
GGATGGGGTGGGAGGGGGGGGGAGGGTGAGAGTGGGTTTGGGGAGAGGGGAGTGGTGGGAGGGGGTGTGTT
GGGGGGGTGTGAGGATGGGGTGGGGTTGGGTGGAGTAGGGGTAGTGTGAGGGAGAGTTGGGGGGGGGTGTGGG
GGTGGGGTAGTTGAGGGAGTTGAATGAAGTGTATTAGGTTGTGGAGGGAGATGGAGAGGGAGTTGAGGGGTGGGA
GGGGGTTAGGATGGAGGGGGAGGATGGAGTGGAGGAGGTGGTTATGGGTATGAGGGAAGAGGTATTGGGTGGTGA
GTTGGATGGTTTGGGGGATAAAGGGAAGTGGAAAAAGTGGTGGTGGTGTGTTTGGTTGGGTGAGGGGTGGATGGG
GGGTGGGGTGGGGAAGAGGAGAGGGTTGATAGAGAACTGGGGATGGTTGGGGGTATGGGGAAAATGAGGGGGGT
AAGGGGAGGAGGGGTGGGGTTTTGATGATATTTAATGAGGAGTGTGAGGGAGTGGGAGAGGAAGGGGGGGT
GTAAAGGGGGATAGTGAGGAAGGGGTGGGAGTATTTAGGGAAGGGGGAAGAGTGTAGGGATGGGGTGGGGGT
ATTGGGAAGGATCAGGGGGGGGTGTGTGGAGGTAGGGAAGGGATTTTTGATGGAGGATTTGGGGAGAGGGG
GGAAGGGGTGGTGTGATGGAGGGGGGGTAGATGGGGGAAATAATATGGGTGGGGGTGGTGTGGGGTGGGGGGG
```

GTTGATAGTGGAGGGGGGGGAAGGATGGAGAGATTTGATGGAGGGATAGAGGGGGTGGTGATTAGGGGGGTGGG
 GTGATTGATTGGGGAGGGAGGAGATGATGAGAGTGGGGTGATTAGGATGGGGGTGGAGGATTGGGGTTAGGGGTT
 GGGTGATGGGGGGTAGGGAGGGGGGATGATGGGTGAGAGGATTGATTGGGAGGATGGGGTGGGTTTGAATATTGG
 GTTGATGGAGGAGATAGAGGGGGTAGGGGTGGGAGAGGGTGTAGGAGAGGGGATGGTTGGGATTAATGGGAAGAGG
 GGAGGGGGTTAAAGTTGTTGTGGTTGATGAGGAGGATATGGTGGAGGATGGTGTGGTGATGGATGAGGTGAGGAT
 GGAGAGGATGATGGTGGTGAGGGTTAAGGGGTGGAATGAGGAAGGGGTGGGGTTGAGGAGGAGGAGGATTTT
 GAATGGGGAGGTGGGGGAAAGGGAGATGGGAGGGTTGTGGTTGAATGAGGGTGGGGTGGGGGTGTGGAGTTGAA
 GGAGGGGAGGATAGAGATTGGGGATTGGGGGGTGGAGAGTTTGGGGTTTTGGAGGTTGAGAGGTAGTGTGAGGG
 GATGGGGATAAGGAGGAGGGTGATGGATAATTTGAGGGGGGAAAGGGGGGTGGGGTGGGGAGGTGGGTTTGAG
 GGTGGGATAAAGAAAGTGTTAGGGGTAGGTAGTGAGGGAAGTGGGGGAGATGTGAAGTTGAGGGTGGAGTAGAG
 GGGGGGTGAAATGATGATTAAAGGGAGTGGGAAGATGGAAATGGGTGATTTGTGTAGTGGGTTTATGGAGGAAGG
 AGAGGTGAGGGAATAAGGGGTGATGGGGGAGATATGGTGATGTTGGAGATAAGTGGGGTGAAGTGGAGGGGAGGA
 GGATGAGGGGGAGGGGTTTTGTGGGGGGGTAAAAATGGGGTGAGGTGAAATTGAGAGGGGAAAGGAGTGTGGT
 GGGGGTAAGGGAGGGAGGGGGGTGGAGGAGAGATGAAAGGGGGAGTTAAGGGGATGAAAAATAATTGGGGTGT
 GGGGTTGGTGTAGGGAGGTTTGATGAAGATTAAATGTGAGGGAGTAAGAAGGGGTGGGATTGTGGGTGGGAAGAA
 AGGGGGGATTGAGGGTAATGGGATAGGTGAGGTGGTGTAGATGGGGGGATGGTAAGGGTGGATGTGGGAGTTTG
 AGGGGAGGAGGAGAGTATGGGGGTGAGGAAGATGGGAGGGAGGGAGGTTTGGGGGAGGGGTGTGGTGGGGGAAA
 GGAGGGAAAGGGGGATTGGGGATTGAGGGTGGGGAAGTGTTGGGAAGGGGGATGGGTGGGGGGGTGTTGGGTATT
 AGGGGAGGTGGGGAAAGGGGGATGTGGTGAAGGGGATTAAAGTTGGGTAAAGGGGAGGGTTTTGGGAGTGAGGAGG
 TTGTAAAAGGAGGGGGAGTGAATGGGTAATGATGGTGATAGTAGGTTGGTGAGGTTGTGAGTGGAATAAGTGTGA
 GGTGGGGGAAAATGGAGTAATAAAAAGAGGGGTGGGAGGGTAATTGGGGGTGGGAGGGTTTTTTTGTGTGGGTA
 AGTTAGATGGGGGATGGGGGTGGGGTTATTAAGGGGTGTTGTAAGGGGATGGGTGGGGTGATATAAGTGGTGGG
 GGTTGGTAGGTTGAAGGATTGAAGTGGGATATAAATATAAAGAGGAAGAGAAGAGTGAATAAATGTGAATTGAT
 GGAGAAGATTGGTGGAGGGGTGATATGTGTAAGGTGGGGGTGGGGGTGGGTAGATGGTATTATTGGTTGGGT
 AAGTGAATGTGTGAAAGAAGG

Fourth, a *thyA* (thymidylate synthase) mutation was introduced into the strain by P1 transduction. In the absence of exogenous thymidine, *thyA* strains are unable to make DNA and die. The defect can be complemented in *trans* by supplying a wild-type *thyA* gene on a multicopy plasmid (Belfort, M., Maley, G.F., and Maley, F. (1983). Proc Natl Acad Sci U S A 80, 1858-861). This complementation was used here as a means of plasmid maintenance.

An additional modification that is useful for increasing the cytoplasmic pool of free lactose (and hence the final yield of 2'-FL) is the incorporation of a *lacA* mutation. LacA is a lactose acetyltransferase that is only active when high levels of lactose accumulate in the *E. coli* cytoplasm. High intracellular osmolarity (e.g., caused by a high intracellular lactose pool) can inhibit bacterial growth, and *E. coli* has evolved a mechanism for protecting itself from high intra cellular osmolarity caused by lactose by "tagging" excess intracellular lactose with an acetyl group using LacA, and then actively expelling the acetyl-lactose from the cell (Danchin, A. Bioessays 31, 769-773 (2009)). Production of acetyl-lactose in *E. coli* engineered to produce 2'-FL or other human milk oligosaccharides is therefore undesirable: it reduces overall yield. Moreover, acetyl-lactose is a side product that complicates oligosaccharide purification schemes. The incorporation of a *lacA* mutation resolves these problems. Sub-optimal production of fucosylated oligosaccharides occurs in strains lacking either or both of the mutations in the colanic acid pathway and the lon protease. Diversion of lactose into a side product (acetyl-lactose) occurs in strains that do not contain the *lacA*

mutation. A schematic of the *lacA* deletion and corresponding genomic sequence is provided above (SEQ ID NO: 288).

The strain used to test the different $\alpha(1,2)$ FT candidates incorporates all the above genetic modifications and has the following genotype:

$\Delta amp^C::P_{trp}^B cI$, $\Delta(lacI-lacZ)::FRT$, $P_{lacI} lacY^+$, $\Delta wcaJ::FRT$, $thyA::Tn10$, $\Delta lon:(npt3, lacZ^+)$, $\Delta lacA$

The *E. coli* strains harboring the different $\alpha(1,2)$ FT candidate expression plasmids were analyzed. Strains were grown in selective media (lacking thymidine) to early exponential phase. Lactose was then added to a final concentration of 0.5%, and tryptophan (200 μ M) was added to induce expression of each candidate $\alpha(1,2)$ FT from the P_L promoter. At the end of the induction period (~24 h) equivalent OD 600 units of each strain and the culture supernatant was harvested. Lysates were prepared and analyzed for the presence of 2'-FL by thin layer chromatography (TLC).

A map of plasmid pG217 is shown in FIG. 11, which carries the *B. vulgatus* FutN.

The sequence of plasmid pG217 is set forth below (SEQ ID NO: 291):

TCTAGAATTCTAAAAATTGATTGAATGTATGCAAATAAATGCATACACCATAGGTGTGGTTTAATTTGATGCCCT
TTTTTCAGGGCTGGAATGTGTAAGAGCGGGTTATTTATGCTGTTGTTTTTTTGTACTCGGGAAGGGCTTTACCT
CTTCCGCATAAACGCTTCCATCAGCGTTTATAGTTAAAAAATCTTTCGGAACCTGGTTTTGCGCTTACCCCAACC
AACAGGGGATTTGCTGCTTTCCATTGAGCCTGTTTCTCTGCGCGACGTTTCGCGGCGCGCTGTTTGTGCATCCATC
TGGATTCTCCTGTCTAGTTAGCTTTGGTGGTGTGTGGCAGTTGTAGTCCTGAACGAAAAACCCCGCGATTGGCAC
ATTGGCAGCTAATCCGGAATCGCACTTACGGCCAATGCTTCGTTTCGTATCACACACCCCAAGCCTTCTGCTTT
GAATGCTGCCCTTCTTCAGGGCTTAATTTTTTAAGAGCGTCACCTTCATGGTGGTCAGTGCCTCTGCTGATGTGC
TCAGTATCACCGCCAGTGGTATTTATGTCAACACCGCCAGAGATAATTTATCACCGCAGATGGTTATCTGTATGT
TTTTTATATGAATTTATTTTTTTCAGGGGCGATTGTTTGGTAGGTGAGAGATCAATTCTGCATTAATGAATCGG
CCAACGCGCGGGGAGAGGCGGTTTTCGCTATTGGGCGCTCTTCCGCTTCTCGCTCACTGACTCGCTGCGCTCGGT
CGTTTCGGCTGCGGCGAGCGGTATCAGCTCAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGC
AGGAAAGAACATGTGAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCA
TAGGCTCCGCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGAAAACCCGACAGGACTATA
AAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCTGCGCTTACCGGATACCT
GTCCGCTTTCTCCCTTCGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCCGTGTAGGT
CGTTTCGCTCCAAGCTGGGCTGTGTGCAGAACCCCCGTTTCAGCCGACCGCTGCGCTTATCCGGTAACATCG
TCTTGAGTCCAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG
GTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACCTACGGCTACACTAGAAGGACAGTATTTGGTAT
CTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCGCTGG
TAGCGGTGGTTTTTTTGTTCGAAAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTT
TTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAAGGAT
CTTCACCTAGATCCTTTTAAATTAATAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAACTTGGTCTGA
CAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCTGTTTCATCCATAGTTGCCTGACT
CCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC
ACGCTCACCGGCTCCAGATTTATCAGCAATAAACAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCTGCAAC
TTTATCCGCTCCATCCAGTCTATTAATTGTTGCCGGAAGCTAGAGTAAGTAGTTCCGCGAGTTAATAGTTTGCG
CAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGTCTGTTGGTATGGCTTCATTACGCTCCGGTTC
CCAACGATCAAGGCGAGTTACATGATCCCCATGTTGTGCAAAAAAGCGTTAGCTCCTTCGGTCTCCGATCGT
TGTCAGAAGTAAGTTGGCCGAGTGTATCACTCATGTTATGGCAGCACTGCATAATTCTCTTACTGTCTATGCC
ATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAG
TTGCTCTTGCCCGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAA
ACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACC
CAACTGATCTTCAGCATCTTTTACTTTACCAGCGTTTCTGGGTGAGCAAAAAAGGAAAGGCAAAATGCCGCAAA
AAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCTTTTTCAATATTAATTGAAGCATTTATCA
GGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATT

TCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAAATAGGCGTATCAC
 GAGGCCCTTTTCGTCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCAC
 AGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTGGCGGGTGTCTGGGG
 CTGGCTTAACATATGCGGCATCAGAGCAGATTGTACTGAGAGTGACCATATATGCGGTGTGAAATACCGCACAGA
 TCGCTAAGGAGAAAAATACCGCATCAGGCGCCTCCTCAACCTGTATATTCTGTAACACAGCCCAATGGGAGCTGTC
 TCAGGTTTGTTCCTGATTGGTTACGGCGCGTTTCGCATCATTGTTGAGTTTTTCCGCCAGCCGACGCGCAGTTT
 ACCGGTGCCTGGGTGCAGTACATCAGCATGGGGCAAATTCTTTCCATCCCGATGATTGTGCGGGGTGTGATCATG
 ATGGTCTGGGCATATCGTCGCAGCCACAGCAACACGTTTCTTGAGGAACCATGAAACAGTATTTAGAAGTATG
 CAAAAAGTGCTCGACGAAGGCACACAGAAAAACGACCGTACCGGAACCGGAACGCTTTCCATTTTTGGTTCATCAG
 ATGCGTTTTTAACCTGCAAGATGGATTCCCGCTGGTGACAATAACGTTGCCACCTGCGTTCCATCATCCATGAA
 CTGCTGTGGTTTTCTGCAGGGCGACACTAACATTGCTTATCTACACGAAAAACAATGTACCATCTGGGACGAATGG
 GCGGATGAAAACGGCGACCTCGGGCCAGTGATGGTAACAGTGGCGCGCCTGGCCAACGCCAGATGGTCTCAT
 ATTGACCAGATCACTACGGTACTGAACCAGCTGAAAAACGACCCGATTTCGCGCCGATTATTGTTTTAGCGTGG
 AACGTAGGCGAATGGATAAAATGGCGCTGGCACCCTGCCATGCATTCTTCCAGTTCTATGTGGCAGACGGCAAA
 CTCTCTTGCCAGCTTTATCAGCGCTCCTGTGACGTCTTCTCGGCCCTGCGTTCAACATTGGCAGCTACGCGTTA
 TTGGTGCATATGATGGCGCAGCAGTGCGATCTGGAAGTGGGTGATTTTGTCTGGACCGGTGGCGACACGCATCTG
 TACAGCAACCATATGGATCAAACCTCATCTGCAATTAAAGCCGCGAACCAGCGTCCGCTGCCGAAGTTGATTATCAA
 CGTAAACCCGAATCCATCTTCGACTACCGTTTCGAAGACTTTGAGATTGAAGGCTACGATCCGCATCCGGGCATT
 AAAGCGCCGGTGGCTATCTAATTACGAAACATCTTGCCAGAGCCGACGCCAGTGTGCGTCAACTGTTGGGAAGGGCGATC
 CGTTAAATTCTTCGAGACGCCCTTCCCGAAGGCGCCATTTCGCCATTAGGCTGCGCAAGGCGATTAAAGTTGGGTAAACGCC
 GGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAAGTTGGGTAAACGCC
 AGGGTTTTTCCAGTCACGACGTTGTAAACGACGGCCAGTGCCAAGCTTTCTTTAATGAAGCAGGGCATCAGGAC
 GGTATCTTTGTGGAGAAAGCAGAGTAATCTTATTCAGCCTGACTGGTGGGAAACCACCAGTCAGAATGTGTTAGC
 GCATGTTGACAAAAATACCATTAGTCACATTATCCGTGAGTCGACGACATGGTAGATAACCTGTTTATTATGCG
 TTTTGATCTTACGTTTAAATATTACCTTTATGCGATGAAACGGTCTTGGCTTTGATATTCAATTTGGTCAGAGATTT
 GAATGGTTCCCTGACCTGCCATCCACATTCGCAACATACTCGATTTCGGTTCGGCTCAATGATAACGTCGGCATAT
 TTAACAAACGAGGTTATCGTTGTCTCTTTTTTTCAGAAATATCGCCAAGGATATCGTCGAGAGATTCCGGTTTAAATCG
 ATTTAGAAGTATCAATAAATTTTTTCTGACCAATAGATATTATCAAAATGAACATTGGCAATTGCCATAAAAA
 CGATAAATAACGTATTGGGATGTTGATTAAATGATGAGCTTGATACGCTGACTGTTAGAAGCATCGTGGATGAAAC
 AGTCCTCATTAATAAACACCACTGAAGGGCGCTGTGAATCACAAGCTATGGCAAGGTCATCAACGGTTTCAATGT
 CGTTGATTTCTCTTTTTTTAAACCCCTCTACTCAACAGATACCCGGTTAAACCTAGTCGGGTGTAACATATAAAT
 CCATAATAATCGTTGACATGGCATACCCCTCACTCAATGCGTAACGATAATTCCCTTACCTGAATATTTTCATCAT
 GACTAAACGGAACAACATGGGTACCTAATGCGCCACTCTCGGATTTTTTCAGGCGGACTTACTATCCCGTAAAG
 TGTTGTATAATTTGCCTGGAATTGTCTTAAAGTAAAGTAAATGTTGCGATATGTGAGTGAGCTTAAAAACAAATAT
 TTCGCTGCAGGAGTATCCTGGAAGATGTTTCGTAGAAGCTTACTGCTCACAAGAAAAAGGCACGTATCTGACGT
 GCCTTTTTTTATTTGTACTACCCTGTACGATTACTGCAGCTCGAGTTAGGATACCGGCACCTTTGATCCAACAGTC
 GGGTAGATATCCGGTGCTTCGGAGTGCTGGAACCAACGGCTCGGCACAATAACAGTCTTATCCATATTAGGGTTC
 AGCCAGGCACCCCAAGAAAACGTGCTGTTACAAATGATGTGATGTTTGCAATGAGACATCAGCATCATATCC
 TGCCAGGAGTCTTCATCAGTGTTCCAGTCAATATAAACCGCATCTGTCAGTGGCAGATTTTCTTTAAACCCACGCG
 ATATCGTCGGAGAAGATATAGTAAGATGGGCTAGCAACACGACGGGACATTTCCGCGATAGCATTCTGGTAATAC
 GGCAGCTGGCACACGGAACCGGTAGTAGCCAGTGTTTCGGCTGCAGATAGTCACCACGACGAATGTGACGGGAA
 ACCGCGTTTTTCATCTTTGTCCAGGATTTCCAGCATGTTTCAGGCTGCGGGAATTTGCTTTTCTTATCAAAGGTG
 AAGGATTCACGCACTTCGTCTTTGATATCAGCGAAGAAACGCTCGCTCTGATAGAAACCTTTAAAGTACAGCAGC
 GGCCAGAAATACTTCTTCTCGAACGCACGCAGAGAGTTTCGGCGCCTGCTTGCCTTCGTAGATTTTTTTAAAAAAC
 AGGAATTCGATAACTTTTTTCAGCGGTTGGTTGATGCAGAATTTCGGTGTGCGGACGGTTGAACACGCGGTGCATT
 TCGTAACCGTAATGGACTTTGTAATGCATCATGTGCTCAGGTCGATACGGACCTTCGGGTAATACTTTTTTCATA
 CGCAGATAGAAAGCATAGATAAACATCTGGTTGCCAGACCGCCAGTCACTTTGATCAGACGCATTATATCTCCT
 TCTTG

Fucosylated oligosaccharides produced by metabolically engineered *E. coli* cells are purified from culture broth post-fermentation. An exemplary procedure comprises five steps. (1) Clarification: Fermentation broth is harvested and cells removed by sedimentation in a preparative centrifuge at 6000 x g for 30 min. Each bioreactor run yields about 5-7 L of partially clarified supernatant. (2) Product capture on coarse carbon : A column packed with coarse carbon (Calgon 12x40 TR) of ~1000 ml volume (dimension 5 cm diameter x 60 cm

length) is equilibrated with 1 column volume (CV) of water and loaded with clarified culture supernatant at a flow rate of 40 ml/min. This column has a total capacity of about 120 g of sugar. Following loading and sugar capture, the column is washed with 1.5 CV of water, then eluted with 2.5 CV of 50% ethanol or 25% isopropanol (lower concentrations of ethanol at this step (25-30%) may be sufficient for product elution.) This solvent elution step releases about 95% of the total bound sugars on the column and a small portion of the color bodies. In this first step capture of the maximal amount of sugar is the primary objective. Resolution of contaminants is not an objective. (3) Evaporation: A volume of 2.5 L of ethanol or isopropanol eluate from the capture column is rotary-evaporated at 56 C° and a sugar syrup in water is generated. Alternative methods that could be used for this step include lyophilization or spray-drying. (4) Flash chromatography on fine carbon and ion exchange media: A column (GE Healthcare HiScale50/40, 5x40cm, max pressure 20 bar) connected to a Biotage Isolera One FLASH Chromatography System is packed with 750 ml of a Darco Activated Carbon G60 (100-mesh): Celite 535 (coarse) 1:1 mixture (both column packings were obtained from Sigma). The column is equilibrated with 5 CV of water and loaded with sugar from step 3 (10-50 g, depending on the ratio of 2'-FL to contaminating lactose), using either a celite loading cartridge or direct injection. The column is connected to an evaporative light scattering (ELSD) detector to detect peaks of eluting sugars during the chromatography. A four-step gradient of isopropanol, ethanol or methanol is run in order to separate 2'-FL from monosaccharides (if present), lactose and color bodies. Fractions corresponding to sugar peaks are collected automatically in 120-ml bottles, pooled and directed to step 5. In certain purification runs from longer-than-normal fermentations, passage of the 2'-FL-containing fraction through anion-exchange and cation exchange columns can remove excess protein/DNA/caramel body contaminants. Resins tested successfully for this purpose are Dowex 22.

The gene screening approach described herein was successfully utilized to identify new $\alpha(1,2)$ FTs for the efficient biosynthesis of 2'-FL in metabolically engineered *E. coli* host strains. The results of the screen are summarized in Table 1.

Production Host Strains

E. coli K-12 is a well-studied bacterium which has been the subject of extensive research in microbial physiology and genetics and commercially exploited for a variety of industrial uses. The natural habitat of the parent species, *E. coli*, is the large bowel of mammals. *E. coli* K-12 has a history of safe use, and its derivatives are used in a large

number of industrial applications, including the production of chemicals and drugs for human administration and consumption. *E. coli* K-12 was originally isolated from a convalescent diphtheria patient in 1922. Because it lacks virulence characteristics, grows readily on common laboratory media, and has been used extensively for microbial physiology and genetics research, it has become the standard bacteriological strain used in microbiological research, teaching, and production of products for industry and medicine. *E. coli* K-12 is now considered an enfeebled organism as a result of being maintained in the laboratory environment for over 70 years. As a result, K-12 strains are unable to colonize the intestines of humans and other animals under normal conditions. Additional information on this well known strain is available at <http://epa.gov/oppt/biotech/pubs/fra/fra004.htm>. In addition to *E. coli* K12, other bacterial strains are used as production host strains, e.g., a variety of bacterial species may be used in the oligosaccharide biosynthesis methods, e.g., *Erwinia herbicola* (*Pantoea agglomerans*), *Citrobacter freundii*, *Pantoea citrea*, *Pectobacterium carotovorum*, or *Xanthomonas campestris*. Bacteria of the genus *Bacillus* may also be used, including *Bacillus subtilis*, *Bacillus licheniformis*, *Bacillus coagulans*, *Bacillus thermophilus*, *Bacillus laterosporus*, *Bacillus megaterium*, *Bacillus mycoides*, *Bacillus pumilus*, *Bacillus lentus*, *Bacillus cereus*, and *Bacillus circulans*. Similarly, bacteria of the genera *Lactobacillus* and *Lactococcus* may be modified using the methods of this invention, including but not limited to *Lactobacillus acidophilus*, *Lactobacillus salivarius*, *Lactobacillus plantarum*, *Lactobacillus helveticus*, *Lactobacillus delbrueckii*, *Lactobacillus rhamnosus*, *Lactobacillus bulgaricus*, *Lactobacillus crispatus*, *Lactobacillus gasseri*, *Lactobacillus casei*, *Lactobacillus reuteri*, *Lactobacillus jensenii*, and *Lactococcus lactis*. *Streptococcus thermophiles* and *Propionibacterium freudenreichii* are also suitable bacterial species for the invention described herein. Also included as part of this invention are strains, modified as described here, from the genera *Enterococcus* (e.g., *Enterococcus faecium* and *Enterococcus thermophiles*), *Bifidobacterium* (e.g., *Bifidobacterium longum*, *Bifidobacterium infantis*, and *Bifidobacterium bifidum*), *Sporolactobacillus* spp., *Micromonospora* spp., *Micrococcus* spp., *Rhodococcus* spp., and *Pseudomonas* (e.g., *Pseudomonas fluorescens* and *Pseudomonas aeruginosa*).

Suitable host strains are amenable to genetic manipulation, e.g., they maintain expression constructs, accumulate precursors of the desired end product, e.g., they maintain pools of lactose and GDP-fucose, and accumulate endproduct, e.g., 2'-FL. Such strains grow well on defined minimal media that contains simple salts and generally a single carbon source. The strains engineered as described above to produce the desired fucosylated

oligosaccharide(s) are grown in a minimal media. An exemplary minimal medium used in a bioreactor, minimal "FERM" medium, is detailed below.

Ferm (10 liters): Minimal medium comprising:

40g $(\text{NH}_4)_2\text{HPO}_4$

100g KH_2PO_4

10g $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$

40g NaOH

1X Trace elements:

1.3g NTA (nitrilotriacetic acid)

0.5g $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$

0.09g $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$

0.09g $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$

0.01g $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$

0.01g $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$

0.02g H_3BO_3

0.01g $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$ (pH 6.8)

Water to 10 liters

DF204 antifoam (0.1ml/L)

150 g glycerol (initial batch growth), followed by fed batch mode with a 90% glycerol-1% MgSO_4 -1X trace elements feed, at various rates for various times.

A suitable production host strain is one that is not the same bacterial strain as the source bacterial strain from which the fucosyltransferase-encoding nucleic acid sequence was identified.

Bacteria comprising the characteristics described herein are cultured in the presence of lactose, and a fucosylated oligosaccharide is retrieved, either from the bacterium itself or from a culture supernatant of the bacterium. The fucosylated oligosaccharide is purified for use in therapeutic or nutritional products, or the bacteria are used directly in such products.

EXAMPLES

Example 1: Identification of novel $\alpha(1,2)$ fucosyltransferases

To identify additional novel $\alpha(1,2)$ fucosyltransferases, a multiple sequence alignment query was generated using the alignment algorithm of the CLCbio Main Workbench package, version 6.9 (CLCbio, 10 Rogers Street #101, Cambridge, Massachusetts 02142, USA) using four previously identified lactose-utilizing $\alpha(1,2)$ fucosyltransferase protein sequences: *H. pylori* futC (SEQ ID NO: 1), *H. mustelae* FutL (SEQ ID NO: 2), *Bacteroides vulgatus* futN (SEQ ID NO: 3), and *E. coli* 0126 wbgL (SEQ ID NO: 4). This sequence alignment and percentages of sequence identity between the four previously identified lactose-utilizing $\alpha(1,2)$ fucosyltransferase protein sequences is shown in FIG. 3. An iterative PSI-BLAST was performed, using the FASTA-formatted multiple sequence alignment as the query, and the NCBI PSI-BLAST program run on a local copy of NCBI BLAST+ version 2.2.29. An initial position-specific scoring matrix file (.pssm) was generated by PSI-BLAST, which was then used to adjust the score of iterative homology search runs. The process is iterated to generate an even larger group of candidates, and the results of each run were used to further refine the matrix.

A portion of the initial position-specific scoring matrix file used is shown below:

Last position-specific scoring matrix computed																				
	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V
1 M	-1	-1	-2	-3	-2	0	-2	-3	-2	1	2	-1	6	0	-3	-2	-1	-2	-1	1
2 A	2	-2	0	4	-2	-1	1	-1	-1	-2	-3	-1	-2	-3	-1	1	-1	-3	-3	-1
3 F	-2	-3	-3	-4	-3	-3	-3	-3	-1	0	0	-3	0	7	-4	-3	-2	1	3	-1
4 K	0	3	0	-1	-2	1	0	-1	-1	-3	-3	3	-2	-3	-1	2	0	-3	-2	-2
5 V	-1	-3	-3	-4	-1	-3	-3	-4	-3	4	2	-3	1	0	-3	-2	-1	-3	-1	3
6 V	-1	-3	-3	-3	-1	-3	-3	-4	-3	4	1	-3	1	-1	-3	-2	0	-3	-1	3
7 Q	-1	4	0	-1	-3	4	1	-2	0	-3	-2	3	-1	-3	-2	0	-1	-3	-2	-3
8 I	-1	-3	-3	-4	-1	-2	-3	-4	-3	3	2	-3	1	0	-3	-2	-1	-3	-1	3
9 C	-1	-1	0	-1	5	3	0	-2	4	-2	-2	0	-1	-2	-2	0	2	-2	-1	-1
10 G	0	-3	-1	-1	-3	-2	-2	6	-2	-4	-4	-2	-3	-3	-2	0	-2	-3	-3	-3
11 G	0	-3	-1	-1	-3	-2	-2	6	-2	-4	-4	-2	-3	-3	-2	0	-2	-3	-3	-3
12 L	-2	-2	-4	-4	-1	-2	-3	-4	-3	2	4	-3	2	0	-3	-3	-1	-2	-1	1
13 G	0	-3	-1	-1	-3	-2	-2	6	-2	-4	-4	-2	-3	-3	-2	0	-2	-3	-3	-3
14 N	-2	-1	6	1	-3	0	0	-1	1	-4	-4	0	-2	-3	-2	1	0	-4	-2	-3
15 Q	-1	1	0	0	-3	6	2	-2	0	-3	-2	1	-1	-3	-1	0	-1	-2	-2	-2
16 M	-1	-2	-3	-4	-2	-1	-2	-3	-2	1	3	-2	5	0	-3	-2	-1	-2	-1	1
17 F	-2	-3	-3	-4	-3	-3	-4	-3	-1	0	0	-3	0	7	-4	-3	-2	1	3	-1
18 Q	-1	0	-1	-1	-3	5	1	-2	0	1	-1	1	0	-2	-2	-1	-1	-2	-2	0
19 Y	-2	-2	-3	-3	-3	-2	-3	-3	1	-1	-1	-2	-1	5	-3	-2	-2	2	6	-1
20 A	4	-1	-1	-1	-1	-1	-1	0	-2	-2	-2	-1	-1	-2	-1	2	0	-3	-2	-1
21 F	-2	-3	-3	-4	-3	-3	-4	-3	-1	0	0	-3	0	7	-4	-3	-2	1	3	-1
22 A	3	-2	-1	-2	-1	-1	-1	4	-2	-2	-2	-1	-2	-3	-1	1	-1	-3	-2	-1
23 K	-1	0	-1	-2	-3	0	-1	-3	1	-2	-2	3	-1	2	-2	-1	-1	1	6	-2
24 S	2	-1	-1	-2	-1	-1	-1	-1	-2	-1	1	-1	0	-1	-1	3	0	-3	-2	0
25 L	-2	3	-2	-3	-2	-1	-2	-3	-2	1	3	0	1	0	-3	-2	-1	-2	-1	0
26 Q	0	0	0	-1	-2	4	1	-2	-1	-1	0	0	3	-2	-2	2	0	-2	-2	-1
27 K	-1	2	0	-1	-3	1	0	-2	-1	-2	-2	4	-1	-3	-1	0	2	-3	-2	-2
28 H	-1	0	0	-2	-3	0	0	-2	6	1	-1	2	-1	-1	-2	-1	-1	-3	0	0
29 S	-1	-1	3	-1	-2	-1	-1	-2	0	-1	1	-1	0	1	-2	1	0	0	4	-1
30 N	-1	-1	4	0	-3	-1	-1	3	0	-3	-3	-1	-2	0	-3	0	-1	-1	4	-3
31 T	-1	-2	-1	-2	-2	-1	-2	-2	-2	1	-1	-1	-1	-2	5	0	3	-3	-2	0
32 P	-1	0	-2	-1	-3	0	-1	-2	-2	-3	-3	2	-2	-4	7	-1	-1	-4	-3	-3
33 V	-1	-3	-3	-4	-1	-2	-3	-4	-3	2	2	-3	1	-1	-3	-2	0	-3	-1	4
34 L	-2	3	-2	-3	-2	-1	-2	-3	0	0	2	0	1	1	-3	-2	-1	0	4	-1
35 L	-2	-3	-4	-4	-2	-3	-3	-4	-3	3	3	-3	1	3	-3	-3	-1	-1	1	1
36 D	-2	-2	1	6	-4	0	1	-2	-1	-4	-4	-1	-3	-4	-2	0	-1	-5	-3	-4

The command line of PSI-BLAST that was used is as follows:

```
psiblast-db<LOCAL NR database name> -max_target_seqs 2500-in_msa<MSA file in FAST
format> -out <results output file> -outfmt "7sskingdoms sscinames scomnames sseqid stitle
evaluate length pident" -out_pssm<PSSM file output> -out_ascii_pssm<PSSM (ascii) output>
-num_iterations 6 -num_threads 8
```

This PSI-BLAST search resulted in an initial 2515 hits. There were 787 hits with greater than 22% sequence identity to FutC. 396 hits were of greater than 275 amino acids in length. Additional analysis of the hits was performed, including sorting by percentage identity to FutC, comparing the sequences by BLAST to an existing $\alpha(1,2)$ fucosyltransferase inventory (of known $\alpha(1,2)$ fucosyltransferases, to eliminate known lactose-utilizing enzymes and duplicate hits), and manual annotation of hits to identify those originating from bacteria that naturally exist in the gastrointestinal tract. An annotated list of the novel $\alpha(1,2)$ fucosyltransferases identified by this screen are listed in Table 1. Table 1 provides the bacterial species from which the enzyme is found, the GenBank Accession Number, GI Identification Number, amino acid sequence, and % sequence identity to FutC.

Multiple sequence alignment of the 4 known $\alpha(1,2)$ FTs used for the PSI-BLAST query and 12 newly identified $\alpha(1,2)$ FTs is shown in FIG. 4.

Example 2: Validation of novel $\alpha(1,2)$ FTs

To test for lactose-utilizing fucosyltransferase activity, the production of fucosylated oligosaccharides (*i.e.*, 2'-FL) is evaluated in a host organism that expresses the candidate enzyme (*i.e.*, syngene) and which contains both cytoplasmic GDP-fucose and lactose pools. The production of fucosylated oligosaccharides indicates that the candidate enzyme-encoding sequence functions as a lactose-utilizing $\alpha(1,2)$ fucosyltransferase. Of the identified hits, 12 novel $\alpha(1,2)$ fucosyltransferases were further analyzed for their functional capacity to produce 2'-fucosyllactose: *Prevotella melaninogenica* FutO, *Clostridium bolteae* FutP, *Clostridium bolteae* +13 FutP, *Lachnospiraceae* sp. FutQ, *Methanosphaerula palustries* FutR, *Tannerella* sp. FutS, *Bacteroides caccae* FutU, *Butyrivibrio* FutV, *Prevotella* sp. FutW, *Parabacteroides johnsonii* FutX, *Akkermansia muciniphilia* FutY, *Salmonella enterica* FutZ, and *Bacteroides* sp. FutZA.

Syngenes were constructed comprising the 12 novel $\alpha(1,2)$ FTs in the configuration as follows: EcoRI – T7g10 RBS – syngene – XhoI. FIG. 5A and FIG. 5B show the syngene fragments after PCR assembly and gel-purification.

The candidate $\alpha(1,2)$ FTs (*i.e.*, syngenes) were cloned by standard molecular biological techniques into an exemplary expression plasmid pEC2-(T7)-Fut syngene-*rcsA*-*thyA*. This plasmid utilizes the strong leftwards promoter of bacteriophage λ (termed P_L) to direct expression of the candidate genes (Sanger, F. et al. (1982). J Mol Biol 162, 729-773). The promoter is controllable, e.g., a *trp*-cI construct is stably integrated into the *E. coli* host's genome (at the *ampC* locus), and control is implemented by adding tryptophan to the growth media. Gradual induction of protein expression is accomplished using a temperature sensitive cI repressor. Another similar control strategy (temperature independent expression system) has been described (Mieschendahl et al., 1986, Bio/Technology 4:802-808). The plasmid also carries the *E. coli rcsA* gene to up-regulate GDP-fucose synthesis, a critical precursor for the synthesis of fucosyl-linked oligosaccharides. In addition, the plasmid carries a β -lactamase (*bla*) gene for maintaining the plasmid in host strains by ampicillin selection (for convenience in the laboratory) and a native *thyA* (thymidylate synthase) gene as an alternative means of selection in *thyA*⁻ hosts.

The expression constructs were transformed into a host strain useful for the production of 2'-FL. The host strain used to test the different $\alpha(1,2)$ FT candidates incorporates all the above genetic modifications described above and has the following genotype:

$\Delta ampC::P_{trp}^B cI$, $\Delta(lacI-lacZ)::FRT$, $P_{lacI} lacY^+$, $\Delta wcaJ::FRT$, *thyA*::Tn10, $\Delta lon::(npt3, lacZ^+)$, $\Delta lacA$

The *E. coli* strains harboring the different $\alpha(1,2)$ FT candidate expression plasmids were analyzed. Strains were grown in selective media (lacking thymidine) to early exponential phase. Lactose was then added to a final concentration of 0.5%, and tryptophan (200 μ M) was added to induce expression of each candidate $\alpha(1,2)$ FT from the P_L promoter. At the end of the induction period (~24 h) the culture supernatants and cells were harvested. Heat extracts were prepared from whole cells and the equivalent of 0.2 OD₆₀₀ units of each strain analyzed for the presence of 2'-FL by thin layer chromatography (TLC), along with 2 μ l of the corresponding clarified culture supernatant for each strain.

FIG. 6 shows the oligosaccharides produced by the $\alpha(1,2)$ FT-expressing bacteria, as determined by TLC analysis of the culture supernatant and extracts from the bacterial cells. 2'FL was produced by exogenous expression of WbgL (used as control), FutO, FutP, FutQ, FutR, FutS, FutU, FutW, FutX, FutZ, and FutZA.

Table 4 summarizes the fucosyltransferase activity for each candidate syngene as determined by the 2'FL synthesis screen described above. 11 of the 12 candidate $\alpha(1,2)$ FTs were found to have lactose-utilizing fucosyltransferase activity.

Table 4. 2'FL synthesis screen results

	syngene	24h OD (induced)	2'-FL culture medium	2'-FL cell extract		
<i>Escherichia coli</i>	WbgL	9.58	5	5	pG204 pEC2-WbgL-rcsA-thyA	E640
<i>Prevotella melaninogenica</i>	FutO	12.2	3	2	pG393 pEC2-(T7)FutO-rcsA-thyA	E985
<i>Clostridium bolteae</i>	FutP	10.4	1	2	pG394 pEC2-(T7)FutP-rcsA-thyA	E986
<i>Lachnospiraceae sp.</i>	FutQ	10.6	3	4	pG395 pEC2-(T7)FutQ-rcsA-thyA	E987
<i>Methanosphaerula palustris</i>	FutR	11.9	0	1	pG396 pEC2-(T7)FutR-rcsA-thyA	E988
<i>Tannerella sp.</i>	FutS	11.3	2	3	pG397 pEC2-(T7)FutS-rcsA-thyA	E989
<i>Bacteroides caccae</i>	FutU	12.1	0	2	pG398 pEC2-(T7)FutU-rcsA-thyA	E990
<i>Butyrivibrio</i>	FutV	11.3	0	1	pG399 pEC2-(T7)FutV-rcsA-thyA	E991
<i>Prevotella sp.</i>	FutW	10.5	3	3	pG400 pEC2-(T7)FutW-rcsA-thyA	E992
<i>Parabacteroides johnsonii</i>	FutX	10.7	3	5	pG401 pEC2-(T7)FutX-rcsA-thyA	E993
<i>Akkermansia muciniphila</i>	FutY	9.1	0	0	pG402 pEC2-(T7)FutY-rcsA-thyA	E994
<i>Salmonella enterica</i>	FutZ	11.0	0	3	pG403 pEC2-(T7)FutZ-rcsA-thyA	E995
<i>Bacteroides sp.</i>	FutZA	9.9	3	3	pG404 pEC2-(T7)FutZA-rcsA-thyA	E996

Example 3: Characterization of cultures expressing novel $\alpha(1,2)$ FTs

Further characterization of the bacterium expressing novel $\alpha(1,2)$ FTs FutO, FutQ, and FutX was performed. Specifically, proliferation rate and exogenous $\alpha(1,2)$ FT expression was examined.

Expression plasmids containing fucosyltransferases WbgL (plasmid pG204), FutN (plasmid pG217), and novel $\alpha(1,2)$ FTs FutO (plasmid pG393), FutQ (plasmid pG395), and FutX (pG401) were introduced into host bacterial strains. For example, the host strains utilized has the following genotype: $\Delta ampC::P_{trp}^B cI$, $\Delta(lacI-lacZ)::FRT$, $P_{lacIq} lacY^+$, $\Delta wcaJ::FRT$, $thyA::Tn10$, $\Delta lon::(npt3, lacZ^+)$, $\Delta lacA$

Bacterial cultures expressing each exogenous fucosyltransferase were induced by addition of tryptophan (to induce expression of the exogenous fucosyltransferases) in the presence of lactose. Growth of the cultures was monitored by spectrophotometric readings at A600 at the following timepoints: 4 hours and 1 hour before induction, at the time of induction (time 0), and 3 hours, 7 hours, and 24 hours after induction. The results are shown in FIG. 7, and indicate that expression of the exogenous fucosyltransferase did not prevent cell proliferation. Furthermore, the growth curve for the bacterial cultures expressing the novel $\alpha(1,2)$ fucosyltransferases FutO, FutQ, and FutX is similar to those expressing the known $\alpha(1,2)$ FT enzymes WbgL and FutN.

Protein expression was also assessed for the bacterial cultures expressing each fucosyltransferase after induction. Cultures were induced as described previously, and

protein lysates were prepared from the bacterial cultures at the time of induction (0 hours), 3 hours, 7 hours, and 24 hours after induction. The protein lysates were run on an SDS-PAGE gel and stained to examine the distribution of proteins at each time point. As shown in FIG. 8, induction at 7 hours and 24 hours showed increases in a protein band at around 20-28 kDa for bacterial cultures expressing exogenous FutN, FutO, and FutX. These results indicate that induction results in significant expression of the exogenous fucosyltransferases.

Finally, additional TLC analysis to assess the efficiency and yield of 2'FL production in bacterial cultures expressing novel $\alpha(1,2)$ FTs FutO, FutQ, and FutX compared to known fucosyltransferases WbgL and FutN. Cultures were induced at 7 hours and 24 hours, and run out on TLC. FIG. 9A shows the level of 2'FL in the cell supernatant. The level of 2'FL found in the bacterial cells were also examined. As shown in FIG. 9B, 2'FL was produced in cell lysates from bacteria expressing the novel $\alpha(1,2)$ FTs FutO, FutQ, and FutX at 7 hours and 24 hours after induction.

Example 4: FutN exhibits increased efficiency for production of 2'FL

Fucosylated oligosaccharides produced by metabolically engineered *E. coli* cells to express *B. vulgatus* FutN was purified from culture broth post-fermentation.

Fermentation broth was harvested and cells were removed by sedimentation in a preparative centrifuge at 6000 x g for 30 min. Each bioreactor run yields about 5-7 L of partially clarified supernatant. A column packed with coarse carbon (Calgon 12x40 TR) of ~1000 ml volume (dimension 5 cm diameter x 60 cm length) was equilibrated with 1 column volume (CV) of water and loaded with clarified culture supernatant at a flow rate of 40 ml/min. This column had a total capacity of about 120 g of sugar. Following loading and sugar capture, the column is washed with 1.5 CV of water, then was eluted with 2.5 CV of 50% ethanol or 25% isopropanol (lower concentrations of ethanol at this step (25-30%) may be sufficient for product elution.) This solvent elution step released about 95% of the total bound sugars on the column and a small portion of color bodies (caramelized sugars). A volume of 2.5 L of ethanol or isopropanol eluate from the capture column was rotary-evaporated at 56 C° and a sugar syrup in water was generated. A column (GE Healthcare HiScale50/40, 5x40cm, max pressure 20 bar) connected to a Biotage Isolera One FLASH Chromatography System was packed with 750 ml of a Darco Activated Carbon G60 (100-mesh): Celite 535 (coarse) 1:1 mixture (both column packings were obtained from Sigma). The column was equilibrated with 5 CV of water and loaded with sugar from step 3 (10-50 g, depending on the ratio of 2'-FL to contaminating lactose), using either a celite loading

cartridge or direct injection. The column was connected to an evaporative light scattering (ELSD) detector to detect peaks of eluting sugars during the chromatography. A four-step gradient of isopropanol, ethanol or methanol was run in order to separate 2'-FL from monosaccharides (if present), lactose and color bodies. Fractions corresponding to sugar peaks were collected automatically in 120-ml bottles, pooled.

The results from two fermentation runs are shown in FIG. 10A and FIG. 10B. The cultures were grown for 136 (run 36B) or 112 hours (run 37A), and the levels of 2'-FL produced was analyzed by TLC analysis. As shown in both FIG. 10A and FIG. 10B, the 2'-fucosyllactose was produced at 40 hours of culture, and production continued to increase until the end point of the fermentation process. The yield of 2'-FL produced from run 36B was 33 grams per liter. The yield of 2'-FL produced from run 37A was 36.3 grams per liter. These results indicate that expression of exogenous FutN is suitable for high yield of 2'-fucosyllactose product.

Table 1. Hits from PSI-BLAST multiple sequence alignment query for novel $\alpha(1,2)$ fucosyltransferases

Bacterium names	Accession No.	GI No.	Gene name [bacterium]	% identity FutC	Alias	SEQUENCE	WO 2015/175801
Helicobacter pylori	AAD29869.1	4808599	alpha-1,2-fucosyltransferase [Helicobacter pylori]	98	FutC	MAFKVVQICGGLGNQMFQYAFAKSLQKHSNTPVLDDTSFDWSDRKMQLLELPINLPYASAKEIAIAKMQHLPKLVRDALCKMGEDRVSQEIVFEYEPPELLKPSRLTYFYGYFQDPRYFDAISPLIKQTFTLPPPPENNNKKEEYHRKLSLILAAKNSVFHRRGDYVGIGCQLGIDYQKKALEYMAKRVPNMELFVCEDEFTQNLDLGYPFMDMTTRNKEEEAYWDMILLMQSCQHGIANSTYSWWAAAYLIENPEKIIIGPKHWFHGHENILCKEWWKIESHFEVKSQKYNA	1
Helicobacter mustelae; Helicobacter mustelae 12198	YP_003517.185.1	291277413	alpha-1,2-fucosyltransferase [Helicobacter mustelae 12198]	70.85	FutL	MDFKIVQVHGGGLGNQMFQYAFAKSLQTHLNIPVLDDTTWFDYGNRELGLHLFPIDLQCASAAQIAAAHMQNLPRLVRGALRRMGLGRVSKYFVPELFEPSRIAYFHGYFQDPRYFEDISPLIKQTFTLPHPTHEAEQYSRKLSQLAAKNSVFHRRGDYMRGWLQDISYQLRAIAYMAKRQVQNLFLFCEDLEFVQNLDLGYPFVDMTTRDGAHAWDMMLMQSCKHGIITNSTYSWWAAAYLIKPEKIIIGPSHWIYGNENILCKDWVKIESQFETKS	2
Bacteroides; Bacteroides vulgatus ATCC 8482; Bacteroides sp. 4_3_47FAA; Bacteroides sp. 3_1_40A; Bacteroides vulgatus PC510; Bacteroides vulgatus CL09T03C04; Bacteroides vulgatus dh1.KV7; Bacteroides vulgatus CAG:6	YP_001300.461.1	150005717	glycosyl transferase family protein [Bacteroides vulgatus ATCC 8482]	24.83	FutN	MRLIKVTGGGLGNQMFYAFYLRMKKYYPKVRIDLSDDMMHYKVHYGYEMHRVFNLPHTFCINQPLKKVIEFLFFKKIYERKQAPNSLRAFEKKYFWPLLYFKGFYQSERFFADIKDEVRESFTFDKNKANSRLNMLEILDKDENA VSLHRRGDYLPQKHAWATTGSVCQLPYYQNAIAEMSRRAVSPSYIIFSDDIAWVVKENLPLQNAVVIDWNTD EDSWQDMMLMSHCKHHIICNSTFSWWGAWLNPNNMDKTVIVPSRWVQHSAPDIYPTGWIKVPVS	3
Escherichia coli; Escherichia coli UME4 3065-1	WP_02155.4465.1	545259828	protein [Escherichia coli]	23.13	WbgL	MSIIRLOGGGLGNQLQFQSFYALS KINGTPLYFDISHYAENDDDHGGYRLNNLQIPEEYLYQYTPKINNYKLLVRGSRLYPDIFLFLGCNEFHAYGYDFEYIAQKWKKYGYWQSEHFFHKLHLDLKEFFIPKNVSEQANLLAAKILESQSSLSIHRRGDYIKNKTATLTHGVCSLEYKALNKIRDLAMIRDVFISDDIFWCKENIETLLSKKYNYYSEDLSQEEDLWLSLANHHIIANSFSWWGAYLGSSAQVIYPTPWYDITPKNTYIPVNHWINVDKHS	PCT/US2015/030823

Clostridium bolteae; Clostridium bolteae 90A9; Clostridium bolteae 90B3; Clostridium bolteae 90B8	WP_00257 0768.1	488634090	protein [Clostridium bolteae]	29.86	FutP	MIFQYALYKAFEQKHIDVADLAWYKNKSVKFEYNFGIKINVASEKDINRLSDCQADFVSRIRKIFGKKKSFV SEKNDSCYENDILRMDNVDVLSGYWQTEKYFSNTRKLEDDYSFALVNSQVSEWDSIRNKNKSVSHIRRGDYL QGELYGGICTSLYAEAEIYKMRVPAKFFVSDDDVEWVKQQEDFKGFVIVDRNEYSSALSDMYLMSLCKH NIANSFSFWAAWLNRENEKIVAPRRWLNGKCTPDIWCKKIRI	WO 2015/175801
Lachnospiraceae bacterium 3_1_57FAA_CT1	WP_00925 1343.1	496545268	protein [Lachnospiraceae bacterium 3_1_57FAA_CT1]	29.25	FutQ	MVIVQLSGGLGNQMFEYALYLSLKAKGKEVKIDDVTCEGPGTRPRQLDVFGITYDRASREELTEMTDASML DALSRVRRKLTGRRTKAYRERDINFDP LVMKEDPALLEGCFQSDKYFRDCEGRVREAYRFRGIESGAPLPED YLRLEKQIEDCQSVSHIRRGDYLDESHGGLYTGTICTEAYKEAFARMERLVPGARFFLFSNDPEWTRHFES KNCVLVEGSTEDTGYMDLYLMSRCRHHIANSFSFWGAWLNENPEKKVIAPAKWLNNGRECRDIYTERMI RL	12
Methanospiraea palustris; Methanosp iraea palustris E1- 9c	YP_002467 213.1	219852781	glycosyl transferase family protein [Methanospiraea palustris E1- 9c]	28.52	FutR	MIIVRLKGGGLGNQLSQYALGRKIAHLNTELDLDTTFTTSSDTPRTYRLNNYNIIGTIAAKEIQLIERGRAQ GRGYLLSKISDILLTPMYRRTYVRRMHTFDKAILTVPDNVLDGYWQTEKYFKDIEILRREVTLKDEPD SINLE MAERIQACHSVSLHVRRGDYVSNPTTQQFHGCCSIDYVRAISLIEEKVD DDPFFISDDLPAWAKENLDIPGE KTFVAHNGPEKEYCDLWMLSLCQHIIANSFSFWGAWLGGQDAEKVMVAPRRWALSESFDTSDIIPDSWI TI	13
Tannerella sp. CAG:118	WP_02192 9367.1	547187521	glycosyl transferase family 11 [Tannerella sp. CAG:118]	28.38	FutS	MVRIVEIIGGLGNQMFOYAFSLYLNKSHIWDRLVYDIEAMKTYDRHYGLELEKVFENLSLCPISNRLHRNLQK RSFAKHFVKSLEYHSECEFEDEPVYRGLRPVYRYRGYVWQNEGYFVDPIMIREAFQFNVNILSKTKAIASKMIR RELSVSIHVRRGDYENLPEAKAMHGGICSLDYHKAIDFIRQLDNNICFYLFSDDINWVEENLQLENRCIID WNQGEDSWQDMYLMSCCRHHIANSFSFWAAWLNPNKNKIVLTPNKNWFNHTDAVGIVPKSWIKIPVF	14
Bacteroides cacciae; Bacteroides cacciae ATCC 43185	WP_00567 5707.1	491925845	protein [Bacteroides cacciae]	28.09	FutU	MIKIVKILGGLGNQMFOYALFSLKERPFHEQVMIDTSCFRNYPHLNGFEVDRIFAQKAPVASWRNLIKVAYP YPNRYFWKIGKYLPRKRTMVCVERKNFSFDAVALTRKGDYDGYWQHEEYFCDMKETIWEAFSPPEVDG RNKEIGALLQASDSASLHVRRGDYVNHPLFRGICDLDYKRAIHVMEERVNPQLYCVFSNDMAWCESHRA LLPGKEVVVVDWNGGAESYVDMRLMSLCRHHIANSFSFWGAWLNRNPQKVVVAPERWVNSPIEDPV SDKWIKL	15
Butyrivibrio sp. AE2015	WP_02277 2718.1	551028636	protein [Butyrivibrio sp. AE2015]	27.8	FutV	MIIILKGGGLGNQMFOYALYKSLKRGKEVKIDDKTGFVNDKLIPVLSRWGVEYDRATDEEINLTD SKMDL FSRIRRLTGRKTRIDEESGKFNPEILEKENAYLVGVWQCDKYFDDKD VVREIREAFKPKQELMTDASSWS TLQQIECCESVSLHYVRTDYVDEEHHIHNICTEKYKNAIDRVKQYPSAVFFITDDKEWCRDHFKGPNFIV VELEEGDGTDAEMTLMSCRKHIIANSFSFWAAWLNDSPEKVIAPQKWINNRMDDDIYTERMTKIAL	6

Previotella sp. CAG:891	WP_02248 1266.1	548264264	uncharacterized protein [Prevotella sp. CAG:891]	27.4	FutW	MRLVKMIGGLGNQMFIYAFYLMRKRFNSVRIDLTDMMHYNVHYGYELHKVFLGRLPTEFCMNPQLKKVL EFLFRTIVERKQHGHRMEPYTCQYVWPLVYFKGYQSERFYSEVKDEVRECFTFNPALANRSSQMMEQIQ NDPOAVSIHRRGDYLNPKHYDTIGICQLPYKHAVSEIKKVSNPVHFVSELDWVVKANLPLENAQYIDW NKGADSWQDMMLMSCCKHHIICNSTFSWAAWLNPSVEKTVIMPEQWTSRQSDVDFVASCGRWVRV KTE	WO 2015/175801
Parabacteroides johnsonii;Parabacter oides johnsonii CLO2T12C29	WP_00815 5883.1	495431188	glycosyl transferase [Parabacteroides johnsonii]	26.69	FutX	MRLKIMIGGLGNQMFIYAFYLMKHHYPDTNIDLSDMVHYKVHNGYEMNRIFDLSQTEFCINRTLKKILEFL FFKIIYERRQDPSTLYPYEKRYFWPLLYFKGYQSERFFEDIKDDVRKAFSFLNIANPESLELLKQIEVDQAV SIHRRGDYLLPRHWANTGSVCQLPYKNAIAEMENRITGPSYVFSDDISWVKENIPLKKAIVYVWTKGED SWQDMMLMSHCRHHIICNSTFSWAAWLNPRKEKVIAPCRWFQHKETPDMPYKEWIKVPIN	18
Akkermansia muciniphila;Akkermansia muciniphila ATCC BAA-835	YP_001877 555.1	187735443	glycosyl transferase family protein [Akkermansia muciniphila ATCC BAA-835]	25.67	FutY	MRLFGGLGNQLFOYAFALSRQGGKARLETSSYEHDDKRVCELHFRVSLPIEGGPPPWAFKRSRIPACLR LFAAPKYPHFREKRHGFDPGLAAPPRRHTYFKGYQTEQYFLHCREQLCREFLKTPLTPENARILEDIRSCS ISLHRRDYLNSPYPSPPLLEYLRSMAEMEGRLRAAGAPQESLRYFIISDIEWARQNLRLPALPHVHVDIND GGTGYDLELMRNCRHIIICNSTFSWAAWLNPRKEKVIAPCRWFQHKETPDMPYKEWIKVPIN	19
Salmonella enterica;Salmonella enterica subsp. enterica serovar Poona str. ATCC BAA- 1673	WP_02321 4330.1	555221695	fucosyltransferase [Salmonella enterica]	25.99	FutZ	MYSCLSGGLGNQMFIYAFYLMKRYPKVRIDLSDMVHYVHHGYEMHVRVFNLPHTFCINQPLKKVIEF RKFRNPFPKQISEILSIALFGKALSDRAFYTFETIKNIDKACLFQYQDADLLNKYQILPLFELRDDLLDICKN LELYSLIORSNNTTALHRRGDYVTNQHAQYHGVLDISYNNHAMEYERERERKQNFIFSDVVRWAQKAF ENDNCYVINNSDYDFSADIMYLSLCKNNIIANSTYSWWGAWLNKYEDKLVSPKQWFLGNNETSLRNAS WITL	20
Bacteroides sp. CAG:633	WP_02216 1880.1	547748823	glycosyltransferase family 11 [Bacteroides sp. CAG:633]	26.01	FutZ A	MRLKIMIGGLGNQMFIYAFYLMKRYPKVRIDLSDMVHYVHHGYEMHVRVFNLPHTFCINQPLKKVIEF LFFKKIYERKQDPNLSRAFEKKYVWPLLYFKGYQSERFFADIKDEVKRAFTDSSKVNARSALLRLDADAN AVSLHRRGDYLPQHWATTGSVCQLPYQNAIAEMNRRAVAPSYVFSDDIAWVKENIPLQNAVVIDWN KGEESWQDMMLMSHCRHHIICNSTFSWAAWLNPRKEKVIAPCRWFQHKETPDMPYKEWIKVPIN	21
Clostridium sp. CAG:306	WP_02224 7142.1	547839506	alpha-1 2- fucosyltransferase [Clostridium sp. CAG:306]	34.28		MEKIKVKLQGGMGNMFOYAFGKGLSKFGCKVLFDKINYDELQKTIINNTGKNAEGICVRKYELGIFNLNI DFATAEQICIGELKNACALPGFIRKIFNLKNTVSNRFEKKYGEYDEELKDYSLAYDGYQNPKYFEDI SDKIKKEFTLPEIKNHDIYNNKKLEKIQFENSFHVRRDDYLNINCEIDLDDYQKAVKYILKHIEPNPKFFVCAE DPDYKHNHFDIGYDFELVGENNKTDQTYENMRLMMACKHAIANSSYSWWAAWLSDDYDNKIVIAPTPWL PGISNEIICKNWIIQKRGISNE	2

PCT/US2015/030823

Prevotella sp. oral taxon 306;Prevotella sp. oral taxon 306 str. F0472	WP_00943 4595.1	497004957	protein [Prevotella sp. oral taxon 306]	32.11		MKIVKILGGLGNQMFQYALYSLQESFPKERVALLSCFNGYHLHNGELERIFSLTAQKASATIMRIAYYP NYLLWRIGKRLLPRRTKMCLESSTFRYDESVLTREGNRYFDGYWQDERYFVACREKVLKFTFPFAKRTENLS LLRLDKNSVAIHVRRGDYIGNQLYQGICDLDYRAAIDKISTYVTPSVFCISNDIAWCQTHLOPYLKAPVY VTWNTGTESYRDMQLMSSCAHNIAANSFSWWGAWLNNQNEKVVIAPKRWLNMDDCCQPLPASWVKI	WO 2015/175801
Brachyspira sp. CAG:484	WP_02191 7109.1	547139308	glycosyl transferase family 11 [Brachyspira sp. CAG:484]	30.14		MLVVKLMGGLGNQMFQYAFKALGDKNILFYGDYKKHSLRKVELNRFKCKAVYIPRELKYLKFVFTKFDKIE YMRSGYVPEYLNDRGNHIYIGFWQTEKYFKQIRPRLKDLFTPRKLLDREAGIISKMQQINSVSVHRRDYY DESHYGDNTLDYYKRAIEYISSKIENPEFFFSDDMAYVKEKFAGLKPHSFIDINSNNYSKDLILMKNCNHN IANSTFSWWGAWLNNENEKIVIAPAKWFTVGTENDKDVPDEWIKL	24
Thalassospira profundimaris;Thalas sospira profundimaris WP0211	WP_00888 9330.1	496164823	glycosyl transferase family 11 [Thalassospira profundimaris]	30		MVVKLLGGLGNQMFQYATGRAVASRLDVELLDDVSAFAHYDLRRYELDDWNITARLATSEELARSGVTAA PPSFFDRIARLRLDLPVNCFREASTYDPRILEVSSPVYLDGYWQSERVFLDIEKKLRQEFQLKASIDANNHSFK KKIDGLGKQAVSLHVRGDDYVTPQTASYHGVCSDYYRAAVDYIAEHVSDPCFFVSDDEWVQTNLNIK QPIVLVDANGPDNGAADMALMMAACRHHIIANSFSWWGAWLNNPLNDKIIVAPKKWFGGRANHDTTDLVP DSWVRL	25
Acetobacter sp. CAG:267	WP_02207 8656.1	547459369	alpha-1 2- fucosyltransfera se [Acetobacter sp. CAG:267]	29.9		MAVSPQESKYSAHVSPDKLIRVRLGGGLGNQMFQYAFGLAAGDVLWDNTSFLTNHYRSFDLGLNYSGDF ASNEQIKKCKNEIRFKNILPISRIRKKNLGFIVLKTNRVCEQINRYEPILLSKGDVYDGVQTEKYFKPLRE RLLDHFTLTPLDAANLMLAKIRAADAVAVHIRRGDYLNPRSPFTYLDKDYFLNAMDYIGKRVDPKPHFFIS SDTDWVRTNIQTAYPQTIVEINDEKHGYFDLELMRNCRHHIIANSSTFSWWGAWLNTNPDKIVVAPKQWFR PDAAEYSGDIVPNDWIKL	26
Dysgonomonas mossii;Dysgonomona s mossii DSM 22836	WP_00684 2165.1	493896281	protein [Dysgonomonas mossii]	29.9		MVTVLLSGGLGNQMFQYAAKSLAIRLNTALSVLDLYTFSKKTQATVRPYELGIFNIEDVVETSSLKAKAVIKAR PFIQRHRSFFQRFQVFTDITYAILYQPTFEALTGGVIMSGYFQNESYFKNISELLRKDFSFKYPLIGENKDVAGQI SENQSVAVHIRRGDYLNNKNSQSNFAILEKDYEKAINYISAHVKNPEFYVSEDFDWIKDNLNFKFEPVTFID WNGKGDYIDMQLMSLCKHNIIANSFSWWGAWLNNSEERKIVAPERWVFWDEQKNELLDCFYPQGWIKI	27
Clostridium sp. KLE 1755	WP_02163 6924.1	545396671	glycosyltransfer ase, family 11 [Clostridium sp. KLE 1755]	29.83		>gi 545396671 ref WP_021636924.1 glycosyltransferase, family 11 [Clostridium sp. KLE 1755]MIIIEISGGLGNQMFQYALGQKFMISMGKEVKYDLSFYNDRVQTLRQFELDFDLDCPVASNSELSRFGK GNSLKSRLKQKLGWDKEKIYEENLDLGYQPRIFELDDIYLSGYWQSELYFKDIQILRLTYTFPIQLDYMMGVFL RKIENSNSVSIHRRGDYLNENNLKIYGNICTLNNYNNKALQIAKKTNPFIIFVTNDIEWVVRKELEIPNMVIVDC NSGKLSYWDMYLMSKCKANIVANSFSWWGAWLNNKNNRIIISPKRWLNNHEQSTLTCDNWIRCGDD	8

Gillisia limnaea; Gillisia limnaea DSM 15749	WP_00698 8068.1	494045950	alpha-1,2- fucosyltransferase [Gillisia limnaea]	29.28	MFISKNTVIKLVGGGLGNQMFQFAIAKIAEKEKSEVLVDITFYTELTKKPRHFSLGIFNSSFAIASKKEIDYF TKLSNFNFKKKLGLNYPITIFHESSENFKAQVLELKAPIYLNQYFQSFYFLGKEYVIRKIFKFPDEALDKDNDNI KRKIIGKTSVSLHIRRGDYVNNKTKQFHGNCTIDYYQSAIAYLSSKLTDFNLIFSDDIHWVRQKFNISNQKI YVSGNLNHSWKDMYLM5LSDHNIANSFSSWWGAWLNKNPEKIIAPKRWFADTEQDKNSIDLIPSEWY RI	WO 2015/175801 30
Methylobacterium mobilis; Methylobacterium mobilis JMW8	YP_003048 467.1	253996403	glycosyl transferase family protein [Methylobacterium mobilis JMW8]	29.19	MLVSRIIGGLGNQMFYAAAARASLRISVQLKLDLSGFETYDLHAYGLNFNINVEDVAKKDDYFAGAPESLLKK IKKYLRLGLIQLSFRESDFSQVLELNDNTYLDGYWQCERYFIDFKQIRQDSFKFAPDALNQRYLELDSV NAVSVHRRGDYVNSNTTNEIHGVCDDYYQRAAEFRARIGPENLHFFVSDDDTWVKNISFGSDTTTFS HNDAAKNYEDMRLMSACKHHIIANSFSSWWAAWLNPSKQVVIAPRQWFKSTLLNSDDIVPASWVRL	31
Runella slithyiformis; Runella slithyiformis DSM 19594	YP_004658 567.1	338214504	glycosyl transferase family protein [Runella slithyiformis DSM 19594]	29.14	MIIVKSGGLGNQLFQYAFGRHLATVNOQKELKLTSAITKTSWNTNRSYALDAFNIRAQEAATPEEIKALAGKP NRLLQVRGRKVGITPIQYFQEPHFHFYSSALSISKSHYLEGYVQSEKYFEAITPIIRFEAFITSPSTHAQTIKEKI SNGTSVSIHLRRGDYVNTSKANRYLRPLTMDYYQKADYINQVRKNPNFFLFSDDIKWAKSQVTFPTTHFST GTSAHEDLWLMTHCRHHIIANSFSSWWGAWLNQPDQKIVIAQKWFESTERFDTKDLLPEPWQL	
Pseudoalteromonas haloplanktis; Pseudoalteromonas haloplanktis ANT/505	WP_00295 8454.1	489048235	alpha-1,2- fucosyltransferase [Pseudoalteromonas haloplanktis]	29.1	MIKVAIGGLGNQLFQYATARAIAEKRGDGVVDMDSFSSVKTHTPCLNKRCKATYESKPKLNKLLSNEKIR NLLQKLGFIKKYFETQLPNEEDVLLNNSINYLTYFQSEKYLSIRECLDELTLIEDLNIAETA VSKAIKNAKNSI SIHRRGDYVNSNEGANKTHGVCDSDYFKKALNYFSEKLLDEHTLFIFSDDIWCRNLSFDYKMFNFDGSS ERPEVDMVLM5QCKHQVISNTFSSWWGAWLNKNDEKVVAPKEWFKSTDLSTDIVPNQWIKL	32
uncultured bacterium	EKE06679.1	406985989	glycosyl transferase family 11 [uncultured bacterium]	28.67	MLTLKGGGLGNQMFQYAAASHNLAKNKKTKINFDLSSFFSDIEVRDIKRDYLLDKFNISADISFDQKNSISGRK FLVKVISKFFGEVYRLLKFLSSKYLDGYFQSEKYFNVEEDIRKDTFLKDEMGVEAKKIEQQIVNSKNSVSLHIR RGDYVDDLKTNIYHGVCLDYKRSIKYLKENFGEINIFVSDDIWVKNLAFENLQFVSRPDIKDYEELML MSKCEHNIANSFSSWWGAWLNENKNIIPKEWFKFQKFNINEKHIVPKSWIRL	33
Clostridium sp. KLE 1755	WP_02163 6949.1	545396696	glycosyltransferase, family 11 [Clostridium sp. KLE 1755]	28.57	MVIVQLSGGLGNQMFYALYLSLKAKGVKIDITCYEGGTRPKQLDVFVGSYERATKQELTEMTDSSLD PVSRIIRKLTGRKTKAYREKIDINFPQVMERDPALEGGCFQSEKYFQDQREQVREAYRFRGIESGAYPLPEAY RRLKEIADCKSVSHIRRGDYLEESHGGLYTGICTEQYQEAFAARMEKEVPGAKFLFSDPDWTRHEFKGE NRILVEGSTEDTGYLDLYLMSCKKHNIANSFSSWWGAWLNNDNPKEKVTAPARWLNRECRDIYTERMIRI	1
Francisella philomiragia; Francisella philomiragia subsp. philomiragia ATCC 25015	WP_00428 7502.1	490414974	alpha-1,2- fucosyltransferase [Francisella philomiragia]	28.57	MKIKIQGGGLGNQMFQYAFYKSLKNNCICDVYDKNIYDYLKHYGFELNRIKFNIDLSFARKYHKKEVLGKLFESI IPSKFIVKFNKNYILQKNFAEDKAYFEIDNCYLDGYWQSEKYFKITKDIYDAFTFEPIDSNFELKNIQDYNLV SIHVRGDYVNHPLHGGICDLEYYNKAISFIRSKVANHIFLFSNDILWCKDNKLDRVTYIDHNRMDSYK DMHLM5LCKKHNIANSFSSWWGAWLNQNDKIVAPSKWFWFNDDKINQKDKICPNSWVRI	5

Pseudomonas fluorescens;Pseudomonas fluorescens NCIMB 11764	WP_01733 7316.1	515906733	protein [Pseudomonas fluorescens]	28.52	MVIAHLIGLGNQMFQYAAARALSSAKKEPLLLDTSSFESYTLHQGFELSKLFAGEMCIARDKNDINHVLWSQ AFPRIRNFLHRPKLAFLRKASLIIEPSFHYWNGIQKAPADCYLMGYWQSERVFQDAEEIRKDFTEKLNIMSPQ NIATADQLINTNAISLHVRRGDYVNNSVYAACTVEYQAAIQLLSKRVDAPTFFVFSDDIIDWVKNNLNIIGFPH CYVHNKSGESYNDMLMSMCQHNIANSFSWWGAWLNSNADKIVVAPKQWFNNNTNVNDLFPAPW VTL	WO 2015/175801
Herbaspirillum sp. YR522	WP_00811 7381.1	495392680	glycosyl transferase family 11 [Herbaspirillum sp. YR522]	28.48	MIATRLIGLGNQMFQYAAAGRALRVGSPLLDVSGFANYELRRYELDGRIDATAASAQQLARLGVNATP GTSLLARVLRKVWPQPADRLREASTYDARIEQASAPVLDGYWQSERVFARIRQHLLDEFTLKGDWGSDN AAMAAQIATAGAGAVSLHVRRGDYVNSAHTAQYHGVCSLDYRDVAHIGGRVEAPHFVFSDDHEWVR ENLQJGHPATFVQINSADHGIYDMMMLMKSCRHHIANSFSWWGAWLNPAAEDKIVVAPQRWFKDATNDT RDLPAAWVRL	37
Prevotella histicola;Prevotella histicola F0411	WP_00882 2166.1	496097659	rotein [Prevotella histicola]	28.43	MKIVKILGGLGNQMFQYALYLSUKETFPQENVTVDLSCFHGYHLHNGFEIARIFSLHPDKATVMEILRIAYYP NYFFWQIGKRVLPQRKTMCTESTKLLFDKSVLQREGDRYFDGYWQDERYFIDCRITLNTKFPPTDDNNL ALLKMDTNSVSIHVRRGDYVGNKLYQGICDLNYYREAIMKISSYSPSMFCVFSNDIEWCRDNLSEFIKAPIY YVDWNSGTESYRDMQLMSCCGHNIANSFSWWGAWLQNSSKIVIAPKRWINLKNCGFMLPSRWVKI	38
Flavobacterium sp. WG21	WP_01749 4954.1	516064371	protein [Flavobacterium sp. WG21]	28.42	MIVVQLIGLGNQLFQYAAAKALALQTKQKFSLDVVSQFESYKLHNVALNHFNVISKNYKPNRYLRKIKSFYQ KNVYKEVDFGYNPDLIHLKGGIFLEGYFQSEKYFIKEKEIREDFELRPLKKTAAIAKIESVNSVSIHRRGD YINNPLHNTSKEEYNNKALEIVENKINNPVYFVFSDDMEWVKANFSTKQETIFIDFNDASTNFEDLKLMTSK HNIIANSFSWWGGWLNKNPDKIVIAPKRWFWNDDSDINTNDIPTNWVKI	39
Polaribacter franzmannii	WP_01894 4517.1	517774309	protein [Polaribacter franzmannii]	28.42	MIIVRVGGLGNQMFQYAKALQKQGYQVKIDITKFKYNLHGGYQLDQFKIDLETSSPIANVLCRIGLRRS VKEKSLFDEKFLIPQREYIKGYFQTEKYFESSITPILRKQFIVQKELCNTTLRYLKEITIQKNACSLHIRRGDYISDE KANSVHGTCDLPYKKSIRIQDEYKDAHFFIFSDDISWAKNLTNKNITTFIEHIVMPHEDMHLMSLCKHNIT ANSSFSWWGAWLQNHENKTVIAPKNWFVFNRENEVACANWQL	40
Polaribacter sp. MED152	YP_007670 847.1	472321325	glycosyl transferase family 11 [Polaribacter sp. MED152]	28.42	MVVVRILGGLGNQMFQYAKSLAEKGYEVQIDISKFSYKYLHGGYHLDKFRIDLETANSSSAFLSKIGLKTIK EPNLLFHKDLKLVNNNAFIKGYFAEQYFSDIREILNQFIKKELAKSTLAIKNQJELLKTTCSLHVRRGDYISDK KANKVHGTCDLDYSSAIEHISKQNSNVHFFVFSDDIWVKDNLNITNATYIDHNVPHEDMYLMTLNHNH TANSSFSWWGAWLQNPDKIVIAPKNWFVDKENEVACKSWITL	1

PCT/US2015/030823

Methanococcus maripaludis;Methanococcus maripaludis C7	YP_001329 558.1	150402264	glycosyl transferase family protein [Methanococcus maripaludis C7]	28.19	MKIQLKGGGLGNQMFQYALYKSLKRGQEVLLDISWYLKNAHNGYELEWVFGLSPEYASIRQCCKLGDPI NLIYVVRKVPKKTTHFEKSNFYDNNVFEVTNGYFEGYQWQENYFKNFRSEILNDFSKNIDKRNAEFSEY LKINSVSVHVRRGDYVTNQKALNVHGNICNLEYNKAJNANNLNKPKFVIFSDDITWCKSNLGDIDPVPV DWNTGPYSYQDMYLMNSCKNIIANSFSWWGAWLNQNTKFKVFSPPKKVWVDRNNVNIIVPNGWIKI	WO 2015/175801
Gallionella sp. SCGC AAA018-N21	WP_01829 3379.1	517104561	protein [Gallionella sp. SCGC AAA018-N21]	28.15	MIIAIIIGGLGNQMFQYAGRALSLARGVPFKLDISGEGYDLHQGFELQRFVFNCAAGIASAEVRDSL GW QFSSPIRRIVARPSLAVLRBSTFVVEPHFYWAGIKQVDPNCLAGYWQSEQYFQSHAAVIRTDFAKPPLSG QNSKLAMQIAQQGNVSLHRRGDYANNPKTTATHGLCSLDYYRAAIIHIAERQSPHFIFSDDIAWVKSNI AINFPHQYVDHNQGTESYNDMRLMSLCQHNIIANSFSWWGAWLNNTNAHKIVIAPKQWFANTTHVADLI PSSWERL	43
Azospira oryzae;Dechlorosoma suillum PS	YP_005026 324.1	372486759	Glycosyl transferase family 11 [Dechlorosoma suillum PS]	28.04	MQSPACIAGARAWVVGYGMAEAMQPVVVGSLGGLGNQMFQYAAAGRALAHLRGLHPLSLDSWFQGRG DRHFAIAPHIAASLERAWPRLPPAMQAQLSRLSRWAPRIMGAPVFREPHEHYVPAFAALAAPVFLEGY WQSERYFRELREPLLQDFSLRQPLPASCQPIAAIGNSDAICVHVRRGDYLSPVAAKHGVCPVDYQQGV AELSASLARPHCFVSDDPWEVVRGSLAFPCMTVVDVNGPAEAHFDLALMAACQHFVIANSSLSWWGAW LGQAAGKRVIAPSRWFLTSDKDARDLLPPSWERR	44
Prevotella paludivivens	WP_01846 3017.1	517274199	protein [Prevotella paludivivens]	28	MKIVKIIGGLGNQMFQYALAMALNKNFTDEEVKLDIHCNFGYTKHQGFEDRVFNGEFELASYRDRVAKVAY PYNFKLWRIGSRIFDRRHIMISEDTSFKIMPEVITSHINYKYDGYWQHEEYKNIHDEILDFAKFPKQDER NKALAEERLSDSNSISIHRRGDYLNDELFGTGCIEYKKAIEEINERTVPTLFCVFSNDIHWCKENIEPLNGKE TIWVDWNTGSDNYRDMQLMTKCKHNIIANSFSWWGAWLNNTKDKIVIAPIRWYNTKEKVSVPVANSWIKL	45
Gramella forsetii;Gramella forsetii KT0803	YP_860609. 1	120434923	alpha-1,2- fucosyltransferase [Gramella forsetii KT0803]	27.96	MSNKNPVIVEIMGGLGNQMFQYAVAKLAEKNSSVLLVDTNFYKEISQNLKDFPRYFSLGIFDISYKMGTEN GMVNFKNLSFKNRVSRKGLGNYPKIFKEKSYRFDADLFNKKTPYLGKYFQSYKYFVIGESKIRQWFEFPYENL GVGNEEIKSKILEKTSVSVHRRGDYVENKKTKEFHGNCSELYYKNAITYFLDIVKEFNIVFFSDDISWVRDEFK DLPNEKVFVTGNLHENSWKDMYLMSLCDHNIIANSFSWWAAWLNNNSEKNVIAPKKWFADIDQEQLSL DILPPSWIRM	46
Mariprofundus ferrooxydans;Mariprofundus ferrooxydans PV-1	WP_00984 9029.1	497534831	alpha-1,2- fucosyltransferase [Mariprofundus ferrooxydans]	27.92	MIIQFTGGLGNQMFQYALGRRLSLHLDVELKFDLSFYQHDIIRDLMDFRQVNGQVATEKEIEAYTNTPIF ALDRPLDLRLVRWGLYRGIVSVSEPPGKQALMVVNSRVLAQRNTYVQGYWQSEKYFMPIRQKLLDDFSL VDKADQANGAMLEKIRQCHSVSLHVRGDYVSNPLTNHSHGTGCLYYEKAIALIGSKVDDPHFFVSDPPE WTRDHLKCRFPMYVTVCNSADSCEDWDMELMRHCRDHIANSFSWWGAWLNMMNPDKVWVVAAPAAWFN NFSADTSDILIPDSWVRI	7

Bacillus cereus; Bacillus cereus VD107	WP_00217 4293.1	488102896	protein [Bacillus cereus]	27.91	MKIQVSSGLGNQMFQYALYKKISLNDNDVFLDSSTSYMMYKNQHNQNGYELERIFHIKPRHAGKEIDNLSLDL SELISRRIRKLFGAKSMVVELKEFEYDPIFEKKEITYFKGYWQNNYNYKDIQELRKDFVTEKLDKRNEKLANE IRNKNSVSIHRRGDYLLNKVYEEKFGNANLEYLLKAINLVKKKIEDPKFYIFSDIDWAKNNINLTDVAVYISH NOGNESYKDMQLMSLCKHNIANSTFSWWGAFNNNDKIVVAPKKWINKGLEKVELFPENWITY	WO 2015/175801
Firmicutes bacterium CAG:534	WP_02235 2106.1	547951299	protein [Firmicutes bacterium CAG:534]	27.81	MIIIRMTGGLGNQMFQYALYKLRAMGKEVKMDDFTIEYEGREARPLSLWAFGIEYDRASREELCRMTDGF DPVSRIRRKLFGRKSLMEKDCNFDPEILNRDPAYLTGYFQSEKYPADIEEEVROAQRFSERIWEIPSQLLER IRSYEQIKTTMAVSIVHRRGDYLLQNEEAYGGICTERYKTAIEYVKKQQDASFFVFTNDPDYAGEWILKNF GQEKERFVLIETGTEENGYL DLYLMSLCRHHILANSSFSWWGAYLNP SREKMVIVPHKWFNGNQECRDIYME NMIRIAKEQS	49
Sideroxydans lithotrophicus; Sideroxydans lithotrophicus ES-1	YP_003525 501.1	291615344	glycosyl transferase family 11 [Sideroxydans lithotrophicus ES-1]	27.81	MVISNIIGGLGNQMFQYAAARALSLEKLEVPKLDISGFTNYALHQGFELDRIFGCKIEIASEADVHEILGWQSA SGIRRVVSRPGMSIFRRKGFVPEHFSYWNIGIRKITGDCYLAGYWQSEKYFLDAAVEIRKDFSEKLPDLSHNA ELAEKIDQENAVSLHRRGDYANNPLTAATHGLCSLDYRKSIIKHIAGQVRNPFVFSDDIAWVKDNLEIEFP SQVVDYNHGSMSFNDMRLMSLCKHHIANSFSWWGAWLNPKEKVVIAPERWFANRTDQVQDLLPPGW VKL	50
zeta proteobacterium SCGC AB-137-C09	WP_01828 1578.1	517092760	protein [zeta proteobacterium SCGC AB-137-C09]	27.81	MIVSQIIGGLGNQMFQYATGRALSHRLHDTFFLDLDFGSGYQLHQGFELSNVFOCEVNVATRSQMQALLG WRFSVRRLLMKRSLKWARHGRVMIEPHFYWSRFAINEGCYLSGYWQSERYFKPIENIRQDFKFNHLL KGVNLDLAQQMTEVNSVSLHVRGDYASDANTNHTHGLCLPDYRDAILYAQNTPVAPSFIFSDDIEWCRE HLKLSFPATYIDHNKGSNSYCDMQLMSLCHHHIANSFSWWGAWLNTRLDKIVIAPKQWFWANGNRDLDLI PAEWLVM	51
Pedobacter heparinus; Pedobacter heparinus DSM 2366	YP_003090 434.1	255530062	glycosyl transferase family protein [Pedobacter heparinus DSM 2366]	27.8	MKIIRFLGGLGNQMFQYAFYKSLQHRPHVKADLQGYQEYTLHNGFELEHIFNIKVNSVSSFTSDIFYNKKW LYRKLRLNLRNTYIEEKKLFSFDPSSLNPKSAYWGYWQNFQYFEHIADDLRKDFQFRAPLSAQNQEVLD QTKLSNISLHRRGDYKIDPLLGGLCGPEYQTAINTVTSKVNAARFFISDDIDWCANLKLQDCSFSWNGG TSSYIDMQLMSCKHHIVANSFSWWAAWLNPNPDKIVIAPEKWTNDKINVRMSFPQGWISL	52
Methylophilus methylotrophus	WP_01898 5060.1	517814852	protein [Methylophilus methylotrophus]	27.78	MFQYAMGLSLAENNQTPLKLDLSQFTDYKLNHNGFELSKEVNCSEATASVTQIETLLGICKYFIRIRILKNTYLN LRPAQVWVEFFGYWDGVNFGDNVYLEGYWQSQYKFIIDYESTIRHTFTKNIISGENKLSDRKGSNSVSL HIRRGDYVTNNKNAFIGTCSLIYYQNAIEYFSTKIADPIFFISDDITWAKSNLRLANEHYFVGHNGQEDSHFD MQLMSLCKHHIANSFSWWGAWLNP SKDKIIIAPKKWFA SGLNDQDLVPKDWLRI	3

PCT/US2015/030823

Rhodobacterales bacterium HTCC2255	WP_00803 3953.1	495309205	alpha-1,2- fucosyltransferase [Rhodobacterales bacterium HTCC2255]	27.7	MIYTRIRGGGLGNQLFQYSAARSLADYLNVSGLDTREFDENSPYKMSLNHFNIRADLNPPDLIKHKDGIAYI IDHIKGNQKVKYKEPFLSFDKNLFSNVDTGYLGYWQSEKYFLNRNRKNILSDINLIKKTDFNTINLKEIKKSTSI SLHIRRGDYLSNESYNETHGICSLSYTTDAVEYIKNRLGENIKVFAFSDDDPDWVLENLKLSDVDIKIINNNTSANS FEDLRMLNCDHNIIANSFSWGWAWLNNQPEKIVISPKWYNKKQLQONADIVPSSWLKY	WO 2015/175801
Spirulina subsalsa	WP_01730 2658.1	515872075	protein [Spirulina subsalsa]	27.69	MAKIIARIRGGIGNQFIYAAARRLELINNAELVLDVSQGVHDLQYRQHYQLDHFHIPCRKATPAERFEFPSR VRRYLKRLNQLRPFQRRYVIEQSIDEDPRLIEFKPRGTVHLEGYQWSEYFKDIEATIRQDLQIQPTDPTN LAIVQHIHQHTSVAVHIRFFDQPNADTMNNAPSDYYHRAVEAMETFPVGAHYVLFSDQPEAAKSRIPLPDE RVTLVNHNRRGNKLAYADLWLMTCQCHFIANSTESWWGAWLAENKKQKQVIAPGFKEKREGVSWWGFKGL LPKQWIKL	55
Vibrio cyclitrophicus	WP_01043 3911.1	498119755	glycosyl transferase family 11 [Vibrio cyclitrophicus]	27.67	MVIVKITGGGLGNQLFQYATGSALANKLSCSELVLDLSFYPTQTLRKYLAKFNINARVATDREIFLAGGGNDFFS KALKKGLTSIIFPEYIKEQESIKVVGKIDLCCKSGAYLDGYWQNPVLYFSQNKIELTREFLPRAQLSPSALAWKDH SOASNSVSLHVRRGDYVENAHTNNIHGTCSLEYQYHAEIKIRSEVHNPFVFFSDDIEWCKLNLSSLAEEFV DNNTSAIDDLMLMRQCKHSIIANSTESWWGAWLKLDDGLVIAPRNWFFSSASRNKGIYPKEWHIL	56
Lachnospiraceae bacterium NK4A179	WP_02278 3177.1	551039510	protein [Lachnospiraceae bacterium NK4A179]	27.65	MRSVVVDIKGGYGNQLFCYSFGYAVSKETGSELIIDTSMILDMNNVDRNYQLGVLTITYDSHISYKYGKDFLSR KTGLNRLRKKSAIGFTVVFKEQYVYDPSVFEIKRDTYFDGFWQSSRYFEKYSDDLRLKMLPKKISNAAEKL AEDARDCLSVSHIRRGDYVSLGWTLKDDYIKALDIKERYGSEPFVFSDNKKYADDFFSAAGLYRLMD YETDDAVRDDMFLMSRCSHNIMANSSYWWGAFNLNDNKDKTVICPETGVWGGDFYPEGWMKVTASSG K	57
uncultured bacterium	EKE02186.1	406980610	glycosyl transferase family protein [uncultured bacterium]	27.57	MIIVLYGGGLGNQMFQYALGRHLAEKNNTLKLDISAFESYKLRYELGNLNIIEKALPEEISRLSTLPTGKIER FIRKTLRKPVKKPESYIKENITGGFNPKILDQLQNNIYLEGYWQSEKYFIEIDIRKFSFKPATGKNKEILENINI NSVSLHIRRGDYVTNPEVNQVHGVCSLDYKSCVDIEKKLESFYIFSDDIEWVYKNNLQIQSQVYVVDHNT VDNAIEDMRLMIFSCKHNLANSFSSWWGAWLNSNPDKMVITPRKWFNNTTYDSNDLIPERWIKL	58

PCT/US2015/030823

WO 2015/175801	59	protein [Bacteroides fragilis]	492366053	WP_00582 2375.1	Bacteroides fragilis; Bacteroides fragilis HMW 616	MKIGIIVTGPYIKFWNEFYSSQLYFCVEAEKNYEVFTDSELASQRLPNVHMHLIEDKGVWVNVSSKSFIC EIRNQLTSYDIIFYLNGNFKFSPICYDEILPOAEHNYLTALSFHYLTHPDHPYDRNKNKNCAFIYPYGGQKYYF QGGFYGGRTQEVLSLSEWCRDAIEADFNKVIARFHDSEYNRYLLTQHPPKVLNDKYAFQDIWPVEGEYKAIV LNKEEVPEDNNLOEMKQNYIDPSLFLNDELKFIPIVQLYGGNGMQMFGYAFYLRHISTQERKLLIDPAP CKRYGNHNGVELPSFSKICQDIHISDETKNNIRKLKRGTSLSIEEVRASMQSFKEKQPIFYSGCWQCCTYV ETVKDEIKKDFIFDESKLNPSAQMLRIIRSNVSVHRRNDYLGNNFEYGGICTKSYEKAISQMYTLLKDE PIFIYTTDDPEWVRNFDKSYLVDWNNKNDNWQDMYMSACRHHIANSSFSWWAAWLGFGPEKKVI APSTWLNQMTPDILPTEWIKIPITDKKILDRICNHLHLSSYMKQLGNSGKMGVWIFFHYARYTONPLVE NYAGDLFDELYEIEHKGISFSLDGLGIAWAVEYL VHEQFIEGNTDDSLAIDFKVMQIDPRFTDYSFETGL EGIACVLSRLSPRVCSSTLSDSVYKDLTEACRKYVPVDKANYTRLFLNIESKEVSGYFKDVLQMVLNHSEK AFGSDGLTWQTGLTMIMR	27.46	3
WO 2015/175801	60	protein [Butyrivibrio sp. AE3009]	551034739	WP_02277 8576.1	Butyrivibrio sp. AE3009	MHHQLKGGNGNQMFQYALYKELRSRGKEVKIDDVTGFVDDLRTPVLQRFGEYDRATREEVVKLTDSKMDI FSIIRRLKLTGRKTCRIDEESGTFNPDILELDEAYLVGYWQSDKYFRNEDVIAQLRQEQKRPQEIIMTDSASWA TLQJECQCSVSLHRRTDYIDEHNHHLNCTEKYKGAIDRISQYPSAVFFITDDKEWCRNHFRRGPNFFV VELAEKENTDIAEMLLMSSCKHHICANSSFSWWSAWLNDSPEKMWVVPKWINNRDMDDIYTRDMTKM AI	27.46	3
WO 2015/175801	61	protein [Bacteroides ovatus]	490447027	WP_00431 7929.1	Bacteroides ovatus; Bacteroides ovatus CL02T12C04	MKQTIILSGGLGNQMFQYAFILSMKAKGKSCSLDTTLFQTNKMHNGFELKSVFDIPDSPNQASALHSLIKM LRRYKPKSILTIDEPYTFPCDALESKKSFLMGDWLSPKYFESIKDVVNVNAYRHNIGNKNVDANEMHGNNS VSIHRRGDYLLKPYCVCNENYRQAEIQKDRVDNPIFYVFSNEPSWCDSEFMKEFRVNFKIVNWNQKDS YQDMYLMITQCKHNIIANSTFSWWGAWLNNNTDKIIVAPSKWFKNSEHNINCKEWLLIDTSK	27.43	3
WO 2015/175801	62	protein [Desulfospira joergensenii]	550911345	WP_02266 4368.1	Desulfospira joergensenii	MGKYYETVNVGGLGNQIFQFSAGFALSRLNLDVLNISTFSDSCQKRNFEITYTPKIKNSFACIKDDDPGVFS RLRIPFLNFKKIKQFHESHFFDPAPAFDIREPVRIEGYFQSYKFEKYSQDLKDILLDIPLTSRLKTVLKVISSKES VSVHRRGDYISDQGINEVHGTLINEAYYLSIKLMKEMKFPESFFLFTDDPHYVEENFKFLEDTSCISDNDCLP YEDMYLMAHCHNIIANSSFSWWGAWLNNQNPKEIVAPRKWFSRKLIMEKPVMDLLPDDWILL	27.42	3
PCT/US2015/030823	3	protein C819_03052 [Lachnospiraceae bacterium 10-1]	507817890	EOS74299.1	Lachnospiraceae bacterium 10-1	MNIIRMTGGGLGNQMFQYALFLRLKAQGEVDFDRTEYKGEAEARPIILLWAFGIDYPAAAGEEVEVNETDGV MKFSHRLRRKLFGGRKSEYREKSCNFDQQLKEPAYFTGYFQSEREYFEEVKEQVRKAFQSGKIWGSVSKEL EERIREYQTKIENKQMPVSVHRRGDYLENDEAYGGICTDAYYRKAIEEMEEKFNTVFYFNSNDTGWAKQ WIDHFYKEKSRFVIEGTTEDTGTYLDLFLMSKRAHIIANSSFSWWGAWLDPDQEKIIVAPSKWVNNQDMK DIYTREMIKSPKGEVR	27.39	3

Bacteroides dorei; Bacteroides dorei DSM 17855	WP_00783 2461.1	495107639	protein [Bacteroides dorei]	27.33	MVVVVGAGLANRMFQYAFALSLREKGLDVFIEDSFIPRDFERTKLSVFNVNQRCDKNSFPLVLRER RFYKLLRISEYMSDNRVIERWNLDYLPYIHKASTNCIFIGWISYKYFQSSQEDAVRKAFTKPLDSIRNVELAT KLVTENSVAHFHRKNIDYLNKLPNTCPSPYYEAINYKVVPNPKFYFSSDNWDVWRENIRGVFEFTAVDWN PSSGIHSHCDMQLMSLCKHNIA NSTYSWWSAYLNENNNKIVVCPKDWYGGMVKKLDTIIPESWIING	WO 2015/175801
	WP_02191 6201.1	547127421	protein [Firmicutes bacterium CAG:24]	27.33	MVIVKMSGGLGNQMFOYALYRKIQQTGKDVKLDLFSFQDKNAFRFRSLDFIPYEQTANLEECRKLGECSYR PVDKIRKMFGLKESYYQEDLDKGQPEILEMNPVYLDGYWQCERYQDIREKILEDTYFPKKISIESRLQERI KNTESVSIHRRGDYLDAAANYKIYGNICTIEYYQSAISRMRKLCCKPNFYLSNDPEWAKEIFGDTEDITVEEDK ERPVEDMFLMSRCKHNIIANSFSWWSAAWLNQENKRVIAVKWFNNHSTVDVICDDWIRIDGDHKGGA	65
Clostridium hathewayi CAG:224	WP_02203 1822.1	547299420	epsH [Clostridium hathewayi CAG:224]	27.3	MIYVNIIRGLGNQLFIYAFARALQKSTNQJITLNYTSFRKHYNNTAMDLEQFNIPEDIMFENSKELPWFAFNT DGKVIIRLRHYFPLKILRSILQKMNVLMLWLGDEYEVKYNKRDRDIVDGFWSQSRFYKSVYKELNLIKPMEM SKEIKTMGDNLINQKESVSVRRGDYVTVKKNRDVYICDEKYLNLSIMRMVELVNPVTVWFISDDADWVK DNIVFPGEVYQPPRVTPLTLYLMKACKHFIISNSSFWSWGQYLSNNDNKIVGPAKWYVDGRKTDIEEE WIKIEV	66
	YP_462663. 1	85860461	alpha-1,2- fucosyltransferase [Syntrophus aciditrophicus SB]	27.3	MVIVRLTGGIGNQMFOYAAARRVSLVNNAPLFLDLGWFEQETGSWTPRKVYELDAFRIAGESASVGDIDKDFKS RRQNAFFRRLLPLFLKKRIHTRQTHIEKSYNFDPEILNLQGNVYLDGYWQSEKVSFSDVSEIRREFSFQTDPAE RNRKILERIASCESVSIHRRGDYVTLPDANAFHGLCTPAYRLAVEQISRKVVPEVFFVFSDDIAWARGNLKL GFETCFMDQNGPDGRGDEDLRLMIACRHHIIANSFSWWSGAWLCSNPEKIVYAPRKWFNNGLDTPDNIPAS WIRI	67
Bacteroides cacciae; Bacteroides cacciae ATCC 43185	WP_00567 8148.1	491931393	protein [Bacteroides cacciae]	27.27	MKIVKIIGGLGNQMFOYALYLSKKKYPKEKIKIDISMFETYGLHNGFELKRIFDIDA EYASREEIRELSFYKIKYL QRIFRKIFPVKTECEVEKYDFKFMSEVW/SNCDRYEGYQWQWVEYEAQTEVRSTFTFKELVGRNAKVIREI QYAKMPVSLHRRGDYLLHKLFGGLCDLNYKKAIDYV/LNNYDTPQFYLSNDIEWCKTYILPLVQGYPFILVD WNSGVESYIDMLMSCCRINIANSFSWWSAAWLNDSSEKIVIAPIKLWAHSPYKGEIQLKSWLLF	68
	WP_02275 6304.1	551011888	protein [Butyrivibrio fibrisolvens]	27.24	MIIEMSGGLGNQMFOYALYKSMHLKGLDVTIDKSIYRDVDHKEQVLDLDRFPNVSIEADRKLSSTLRGYGY NDSIIDKIRNKLNSKRNLYHEDLDKGQYQPEIFEDNVDNLYNGYWCERYFKDKIKNEIKKDFIFPCTQSGDDKIK ALTIEMESCNVSLHVRRGDYLLKPGLEIYGNICTEYKYKSIYKERVDPVYIFSNDMMAWVRDNFKSDDFR YVNEDGAFDGMITDMY/LMTRCRHNIVANSFSWWSGAW/LNKHDDNIVICPNRWVWNTHVTVDICEDWIRI DV	69

PCT/US2015/030823

Parabacteroides distasonis; Parabacteroides distasonis CL03T12C09	WP_005857874.1	492476819	protein [Parabacteroides distasonis]	27.24	MIVGGNDYCKVNVNIIGLGNQMFQYAFALSLKEHFKEEIRIDISHFNVLVNVKVGAAHLHNGYELDKIFFNIELKANAWQLMKLTWFIPIYLIARILKIPVRNSEYIQNSDDCFYDPMVYNKQSGCYEGYWAIGYYE SMRDKLKIFQHPSPGKNGKQYIENMESSVGIHRRGDYLLSDNFRGICEVDYKRAIDKILQDGEKHVYFL FSNQKWCCEEYILPLLGNYEIIFVTGNIGRDSWDMFLMTHCKDLIIANSFSWVGAFNLKRGGRVVTPKR WMNRNIRYDLWMPWEWIRI	WO 2015/175801
Geobacter uraniireducens; Geobacter uraniireducens Rf4	YP_001230447.1	148263741	glycosyl transferase family protein [Geobacter uraniireducens Rf4]	27.21	MIARLQGGGLGNQMFQYAVGLHLALTHNVELKIDITMFSYKWHYSLRPFNIRESIATEEIKALTVDKMDR PYKKIDNFLCRLLKRSQKISATHVKEKHFIHYDPDILKLPDNVYLDGWQSEKYKFEIENIRQTFIKNPQLGRDK ELACKILSTESVCLHRRGNVYTDKTTNSVLGPDCLSYNSCIKSLAGNNKDPHFFVFSNDHEWVSKNLKLDYP TIYVDHNNEKDVEDLRLMSQCKHHIIANSTFSWWSAWLCSNPDKVIYAPQKWFRVDEYNTKDLLPSNWLI L	71
Lachnospiraceae bacterium A4	WP_016280341.1	511026085	protein [Lachnospiraceae bacterium A4]	27.21	MIIVIEGLGNQMFQYAFARSIQVNGKKVFLDTSGYTDQLPLCRTSTRRRYQLNCFNIRIKEVEKKNIEKYSFL IQEDMFGKLISKLAKLHLWMYKVTIQNAQEYKESYLNTRGNVYKGFQWQNPKYFSSIRRLLLKEITPKYKIRI PAELRELLQEDNIVAVHCRRGDYQYIRNCLPVNYKKAMAYMEKLGVPYLFSSDDLSWVKRQFGNKN NYYIEDYKGFEDYQELMIMSRNFIANSTFSWWAAWLCSYENKVVIMPRVWTVYVGGQGVEMSDPPAD WIRI	72
Colwellia psychrerythraea; Colwellia psychrerythraea 34H	YP_270849.1	71282201	alpha-1,2-fucosyltransferase [Colwellia psychrerythraea 34H]	27.15	MKVVRVCGGFGNQLFQYAFYLAVKHKFNETTKLDIHDMAVELHNGYELERIFNLNENYCSAEKLAQSTK NIFTKLLKEIKKYTPFIPRTYIKEKHLHFYSQEVDLGTDTSIYRGSWQNPQYFNSIASEIREKLTTFEFTPKSL ALHQEISEHETVAVHRRGDYLLKHALGGICDLPYQNAIKEGLVEKPLFVIFSDDITWCRANINVEKRVFVD WNSGEQSFQDMHLMSLCTHNIANSFSWWGAWLNPANPNKIVISPNKWIHYTDSMGIVPSEWIKVETSI	73
Roseobacter sp. MED193	WP_009810150.1	497495952	alpha-1,2-fucosyltransferase [Roseobacter sp. MED193]	26.96	MITSLHGRGNQMFQYAAARALAHRLGCGVALDGRGAELRGEGVLTRVFDLPLSAAKPLPLKQHAPLRY GLWRGLGLAPFRFRERERGLGYNTAFETWEDGCVLHGYWQSERYFEEISDLIRADFTFPDFSNRQNAEMAARI MEDNAISLHVRRGDYVALSAHVLCDQAYEAAALTRLEGLSQDAPTYYVFSDDPDWAKANLPLPCKKVVVD FNGPETDFEDMRLMSLCKHNIIGNSFSWWAAWLNANPNQKRVAGPANWFGDPKLSNPDLPSQWLKVA P	74
Cesiribacter andamanensis; Cesiribacter andamanensis AMV16	WP_009197396.1	496488826	Glycosyl transferase family 11 [Cesiribacter andamanensis]	26.89	MMIVRLCGGLGNQLFQYAVGKQLSVKNNIPLKIDDSWLRLPDARKYRLQFFQIEEPASQEQVEFVGPYES QSLYARLYRKVQNNMLPRHRRRYFOESGFWAYEPFELMRIRSQVFLGFWQHAYFTRLHPQVLEALQLREY ROEPYAVLDQIREDAASVSLHRRGDYVSDPNLQFFGVMPLSYQQAVAYMQEQLHAPTIFYFSDLDLWA RAHLKQAPMVFDVIEGGRKEYELEAMRLCRHNILANSFSWWGAYLNTNPNHVRVIAPIRQWVADPELKD KVQIQMPDWILL	75

Rhodopirella sallentina; Rhodopirella sallentina SM41	WP_00867 9055.1	495954476	glycosyl transferase family 11 [Rhodopirella sallentina]	26.89	MIATRLIGLGNMFQYAYGFSLARRRRLVLDVSASFESYDLHALAIDQDISAARMTQAEFARIPGRYRG KSRWAERVANFAGGLQSCDKRPLRLRREKPFGEAKYLAEGSDLYLDGYWQSERYPFGLQALKEEFQLKRG LSDESSRVLDIEQSSMSVAMHVRGDDYVTTAETLRIYRLDAEYRKCLNDLRQFSLNLFVFSNDIQWCCQ DHLDDVGLKQRPVTHNDATTAIEDMELMSQCDHSIANSFSSWWAAFLGRSDAQRRVYYPDPWFNPGLTN GDSLGCANVWSSSISVSRPSRAA	WO 2015/175801
	WP_02276 2282.1	551018054	protein [Butyrivibrio sp. AD3002]	26.85	MIIRMMGGGLGNMFQYALYLQKALGKEVKIDDVYFRDDPQRDPVLEKMYGITYTKASDAEVVDITDSH LDIFSRIIRKLFGRKSHEVIEETGLFDPKVFEEETAYINGYFQSDKYFPDKEVLAQLRREFVIKPDDVFTSADSW ELYRQIRETESVSIHVRRGDYLLPGTVETFGGICDNDYYKRAIDRMVSEHPDAIFVFTSDKEWCEQNVSGKK FRIVDTKEENDDAADLLMSLCKHHILANSSYSWWSAWMNDSPEKTVIVPSKWLNTKPMDDIYTSRMTKI	77
Segetibacter koreensis	WP_01861 1017.1	517440157	protein [Segetibacter koreensis]	26.78	MVVVKLIGGMGNMFQYAGRHLAIKNCPLYFDHIELENKNTANTPRNYELDIENVQYQKNPLOSNERFV AKVYHKLFSVQRIKPEPDTFHPHILNVQGNHILNGYWNQENYFKEIEEIRQDFTFKTPANEKIESILQQAATN SVSLHVRRGDYITLLEANQFHGVCSDTYYQKAKIAKEIPAIPHLFVFSDDIHWVQKNMPFTEEHFTVDGNT GKNSFEDLRMAACRHNILANSSFSWWAGWLNKNPEKMWIAPEKWFRAVHTDIVPPSWIKM	78
Amphritea japonica	WP_01962 1022.1	518450815	protein [Amphritea japonica]	26.76	MVIVRLIGLGNQLFOYAYALSLEQGYDVKLDASAFESYTLHGFGGLGEYAERLEVATTEEDVMVSRVGRIS TLRLKLGKRRVVIKESNFYSYDEKMLTPEDSHYLVGYFQSELYFNKIRGELLALDLKHLKSPYTEASYLAIDA SVSVSMHIRRGDYVSDKAAHNTHGVCSDLYYAAVTFEEERYPDVDFYISDDIEWVKENLNVQRAHYISSEE KRFAGEDYILMSQCDHNIVANSSFSWWGAWLNANEDKIVVAPRQWYADSNMQRLSKTLVDPDTWIRL	79
Desulfovibrio vulgaris; Desulfovibrio vulgaris str. 'Miyazaki F'	YP_002437 106.1	218887785	glycosyl transferase family protein [Desulfovibrio vulgaris str. 'Miyazaki F']	26.76	MRPVVVDIFGGGLGNMFQYAAKSLAERLGVRLDLVSMFSGDPLRAFLSGEFAITDHVRGKRSLLVRFA RSLGFGSSKCVPEFFHYWEGINEIEAPVHMHYWQSEKFKAYEDLIRRTFSACEGVASSGKYAGVSSP MSVSVHLRRGDYKEQKNVVGILGREYYDAAYSIKQGPCSACFFVTDAINEAVDFSHWVNDVLFVDGN NQYQDMYILMSQCRHHIANSYSWWGAWLGAESDGMTVAPKMWFAVDVLKEKSIKDLFPEDWIVL	80
Spirosoma spitsbergense	WP_02060 6886.1	522095677	protein [Spirosoma spitsbergense]	26.76	MIISRTSGLGNQLFOYAVARHLSLKNKTSLYVDLSYLYQYHDDTSRNFKLGNFSPYHTLQQSPVEYVSKAT KLLPNRSLRPFFLQKERQFHFEDEQILQSRAGCVILEGFWQSEAYFRDNADTIRDLQLSGTSPFENQYRELI RETPMSVSIHVRRSDYVNVHPEFSQTGFVGIDYYKRAIELARKELANRFFVFSDDKEWSKTNLPLGEDSVFV QNTGLNGDVADLVLMSSHCHQHIIANSFSSWWGAWLNPNAGKLVITPKNWKYKNKPAWNTKDLLPTTWLS I	PCT/US2015/030823

Lachnospiraceae bacterium 28-4	WP_01629 2012.1	511037988	protein [Lachnospiraceae bacterium 28-4]	26.73		WO 2015/175801 83	MNIIRMSGIGNQMFQYALYKLVSLGKEVKFDDVTEYELDNARPIMLSVFGIDYPKASREELVELTDSMD FLSRVRRKIFGRKSGEYHEASADYDETULEKEHAYLCGCFQSERYFKDIEYEVREAYFRNVVVEIRGGIETY ERQIGESLSVHIRRGDYLDAADVYGGICTDAYNQAIRYMIKKYENPSFFVTNDTFWAEKWCEVRERETG KRFTVIKGTDEETGYIDLMLMSRCKAHIANSSFSWWGAWLDASPDKCVVAPVKWINTRECRDIYTEDMVR IGSNGKISFNSCSSL
Lachnospiraceae bacterium COE1	WP_01630 2211.1	511048325	protein [Lachnospiraceae bacterium COE1]	26.71			MVVRIWEGLGNLQFQYARALSLRTKDRVYLDISEYEMSPKVRKYELCHFIIKQPVINCGRIFPFVNKDS FYTNNQYLRYPAGLIKEDCYKRFDFCELGLLYLKGWFQSEKVFKEFESHIREIYPRNKIKITRGLRKLNSD NTVSVHIRRGDFGKDHNLPIEYENS KRVLIRVDNYPYFIIFSDDLWV'KENMNFGLNCFYMDKEYSKDYEE LMIMSRCKHNIANSTFSWWGAWLNPSKDKIVIAPKKWFLYNP'KKDFDIV'PNDWIRV
Parabacteroides;Para bacteroides sp. 20_3;Parabacteroides distasonis CL09T03C24	WP_00586 7692.1	492502331	alpha-1,2- fucosyltransferase [Parabacteroides sp.]	26.69	FutZ B	84	MKIVNIIGGLGNQMFQYAFVALKAKYPNEEVFIDTQHYKNAFIKYYHGNFYHNGYEIDKVFPNATLEPAR PKDLMKVSFYIPNQV'LARAVR'RIIPKRKTEFVTDQQP'VFIPEALSVIDDCYFDGYWMTPLVFDKYDRDLKEF TFRPFDTKENLEPLELLKQDNSVTVHIRRGDYYGSSFGIGICTLDYYRNAIREAYNLITSPEFFIFSNDQKWC/M ENMRNEFGDAKVHFIHNRGADSYRDMQLLSIARCNILANSSFSWWGAYLNQRKNCFIICPHKWHNTLEY SDLYLPTWIKI
Bacteroides sp. HPS0048	WP_00256 1428.1	488624717	protein [Bacteroides sp. HPS0048]	26.62		85	MFVIRLIGVGNGQLFQYTFGQFLRHKFGVEVCYDIVAFDVT/DKGRNLELQLLDLDESPLFETSNNFFSKYKSWK KRFLYGFLKKNNKYTKYAPEEISLFTKGLSYFDGW'WQY'PALLRDTINNMEDEFIPKQIPVQIQKYNEIL LNNFAVALHVRRGDYFTSKYAKTYAVCNVEYTS'AVNLMCEKLRSCFKYVFSDDLDWV'KSNLILPSNTVYVK NYDINSY'WYVYLM'SLCRHHIISNFSFW'WGATLNRNFH'KIVIAPKY'WSTTKKNNTL'CDNSWIKI
Bacteroides thetaiotaomicron;Bac teroides thetaiotaomicron dnLKV9	WP_01626 7863.1	511013468	protein [Bacteroides thetaiotaomicron n]	26.58		86	MKIINILGGLGNQMFQYAMYLALKNASHSEEEILCSTRSFCGYGLHNGYELGRIFGIQV'KEASLLQLTKLAYPFF NYKSWQV'MRH'WLPVRKTM'TRGAINIPFDYSQV'MREDSVY'YDGYWQNEKFNHIREEILTATYTFPKFDDK NQELADIIVKSNV'SCHIRRGDYLKEINMCVCTSSY'AHAI'SYMMNEEINPNLYCVFSDDIW'CRNNICELMGE DKKIIFIDWNKGEKSRDMQML'SLCKHNIANSFSFW'GAWLN'NRNDKKIVV'APTRWIASEVKNDPLCDSW KRIE
Desulfovibrio alaskensis;Desulfovib rio alaskensis G20	YP_389367. 1	78357918	glycosyl transferase [Desulfovibrio alaskensis G20]	26.56			MKFVG'VWILGGLGNQMFQFAAAYALAKRMGGELRLDLSGFKKYPLRSVSLD'LTVDTP'LWHGLPMSQRRF RIPMDAWTRGSR'LPVSP'PFVMAKEKNFAFSPVYELQO'SCYLYGYWQSYRFQDVEDDRTLFSLSRFATL ELAPV'VAQLNEVES'AVHLRRGDYITDAASN'AVHGVCGIDY'YQRSMSLV'RRSTTKPIFYIFSDEPEVAKKLFAT EDDVV'MPSRRQEEEDLLMSRCKHHIANSFSFW'AAWILGRASGLCIAPRYW'FARPKLESTYLFDLIPDE WLL

PCT/US2015/030823

Prevotella oralis CC98A	ETD21592. 1	564721540	protein HMPREF1199_0 0667 [Prevotella oralis CC98A]	26.56		MDIVVIFNGLNQMSQAFYLAKRKSGSRCHCIFHNVSTGFHNGSELDKVFGIKYEKGIFSLLSKYIDFDGIP KLRKKLNSLGIHIREPRNYDYTASLLPRVSRWGLNYFVGWVHSEKYYTEILQEIKNTFSFKIDDEIKDIDFYEYVS LIHNDINSVSLHIRRGDYVVGANEYSYFQGGVATLEYHKAIDEIYQRIENPTFYVFSDDIGWCKTTFLKNNFIF VDCNCGEKSWRDMFUSQCKHHIIANSTFSWWGAWLSIFHNSITICPKFIKGVIRDVYPTDWIKLSS	8 WO 2015/175801
Comamonadaceae bacterium CR	YP_008680 725.1	550990115	glycosyltransferase [Comamonadaceae bacterium CR]	26.54		MASKISKIIPRIFGGLGNQLFIYAARRLALVNGAELALDV5GVFRDHEVNRHYQLDHFHNIPCRKATAAERLE PFARVRRYLKRKWNQRLPFEQRYLVQESVDFDERLLTFKPRGTVYLEGYVQSEDFKDIQIRADLRHPP TDTVNQQMAERIRATNAVAVHVRFFDAPAOQSALGVGGNNAPGDYVYQRAIKVMQEQAPDAQYVIFSDQp QAAARARIPLRDDHVTLVNHNQCDAYADLWLISQCQHFIIANSTFSWWGAWLGKTPESIVAPIGFEKREG AMFWGFRGLLPDRWVKL	89
Vibrio nigripulchritudo; Vibrio nigripulchritudo AM115; Vibrio nigripulchritudo FTn2; Vibrio nigripulchritudo Pon4; Vibrio nigripulchritudo SO65	WP_02259 6860.1	550250577	WbIA protein [Vibrio nigripulchritudo]	26.51		MKDSRIVKLNGLGNQMFQAFALFALKKLNVAVKFDTTELLDNRTEFKLSLERFGLIVDKLTTEKFKYKGLE SCKYRKICNWIISNFTTNIHKGYYKEKRGVYDRGIFDSNVKYIDGYWQNQEYFNDFRSELLNKFNLNGKVS HAIQYLKEITSVQNSVSIHVRRGDYLLDVYRNLTLDYSEAIKLVITNPDSKFFISNDINWCKSNFKSVDNAI FVDSTVDEFDDMFLLMSKCKTNIIANSTFSWWAAWLNNSGKIVYCPKKWRNDTTEVHKGLPEGWNIIDK	90
Sulfurospirillum deleyianum; Sulfurosp irillum deleyianum DSM 6946	YP_003304 837.1	268680406	glycosyl transferase family protein [Sulfurospirillum deleyianum DSM 6946]	26.48		MIIKIMGGTTSQMHKALGRVLSLKYNVPLKDLTWFDNPKSDTPWEYQLDYFNINATVSEIKKLKGN LFNRIARKIEKFFSIRIYKSYINKSFISDFHKLKSDIYLDGEWNGFKYFEDYQDTIKNELTLKRGSSINQNTIKE LKSSDNSVFLHIRRGDYLSNKNAAAFAHAKCSLDYKAIQIVKEIDNPIFYIFSDDILWVKKNFVINESCRFME KNQNFEDILLMSYCKHGTANSFGSLMAGWLNQNKDKMIIVPQTWVNVNDDRININILNSLEQDNFTIIR	91

PCT/US2015/030823

64

Bacteroides salyersiae; Bacteroides salyersiae WAL 10018 = DSM 18765 = JCM 12988	WP_005923045.1	492689153	protein [Bacteroides salyersiae]	26.41		MKQTIIMSGGLGNQMFQYALYCSMRKIRVKIDISLYEFNRMHNGYMLDYAFGLNISHNKINKYSVLWTR LIRSNRAPPELLFREDESRFCDDVFTTYKPYIDGCWIDERYFFENIKKISQSFHNIDQNLMMVANNMMKVCNS VSLHRRGDYLSQSMYCNESYYKSAIEYISRVESDKFFISDDPEWCYKFMKFNVDYIEIQHNFSGKDSYKD MYLMTQCKHNIIANSTFSWWGAWLNNAGKNVVCPSVWINGRDFNPLEEWYHI	WO 2015/175801
Bacteroides fragilis; Bacteroides fragilis CL03T00C08; Bacteroides fragilis CL03T12C07	WP_005786334.1	492241663	protein [Bacteroides fragilis]	26.38		MDIILLHNLGNQMSQYAFYLSKKKNGIHTSYICLSNDHNGIELDKVFGVEQMGCKKIFLLFILRLMSNRT GFILRVNLLFSKIKIKLITENLDYSHPSFLSASPYCLAFWVGWGHHPQYSEISSQKEAFTFRSLDERNICI EKRMREPNSVCLHRRGDYLTGINYELFGKVCNEQYQKAIIDYIEGKLSIDICYVFSNDMEWAKKILGKNV FVDWNRGEESWKDMYLMKSCNLIIPNSTFSWWAAWLCEHPVNI/CPKLFVYGVGEQSDIYLDNWHKIE	96
Bacteroides nordii; Bacteroides nordii CL02T12C05	WP_007486843.1	494751435	protein [Bacteroides nordii]	26.37		MMGIEKTNMIVRLWGGIGNQLFQYSFGFEFLREKYQVDVYDIASFSGSKDKLRKLELSVVVPGIPVTTDISFSK YVGTNRLLRFIYGLKNSFIEEKYFSDEQLFKYLSKRGDYLQGYWQKTIYAETLRRKGSFFLSQEEPIVLTHTKA KIQEAEGAIALHVRRGDYFSKKHINTFGVCDAHYIEKAVDIMRGVSNAMIFVSDDLDDWVRRYVNLPTNVI YVPNYDIPQYWIYLYLMSLCRHNISNSSFSSWWGAFNLNMTNTKIVVSPSKWTLNSDKTIALDEWFKI	97
Butyrivibrio proteoclasticus; Butyrivibrio proteoclasticus B316	YP_003829743.1	302669783	glycosyl transferase 11 [Butyrivibrio proteoclasticus B316]	26.37		MECSMIIKFCGALGNLFQYALYKMRILGKDKADISAFGDGNEKRFYLDDELGIEFNIAADEIAEYLNKRT IRFVPGFLQHRHYFEKPYVYNNKILSYDDCYLEGYWQNYRYFDDIKDELLKHKMFPCLPLEQKLAEMEN ENSVAHVVRMGDYLNLQDLYGGICDADYDRAFSYIEGNISNPVYVYFSDDDVDKASALLAKHKINWIDYNSE KGAIYDLILMSKCKNNIIANSSFSWWGAYLEYNNGKVYVSPNRWMNCFENSNIAYWGWISL	98
Prevotella ruminicola; Prevotella ruminicola 23	YP_003574648.1	294674032	family 11 glycosyl transferase [Prevotella ruminicola 23]	26.33		MRIVKVLGGGLGNQMFQYALYKALQKQYPEERVLLDLHCFNGYHKHGRFEIDSFGVTVYEKATLKEVASLAYP YPNYQCWRIGSRILPVRKTMKEEPNFTLEPSALSPLDSTYDGYWQHEEYFMHIREELSTYAFPAFDDERN KTTAQLAASTNSCIIHRRGDYLTDLPLRKGTTNGVYIAIKEMQKEVPEKWLVSDDIAW/CQQHLASTLD ATNTIYIDWNTGANSIHDMHLMALCRHHIIANSFSWWGAWLSQQDGTITAPSNWMNLKDVCSVPDN WIKI	99
Prevotella salivae; Prevotella salivae DSM 15606	WP_007135533.1	494223898	protein [Prevotella salivae]	26.33		MKIKIIGLGNQMFQYALYALQKQYKDEEIRLDLNCFRGYNKHQGYLLDEIFGRRFRRAASLQEVARLAWPY PHYQLWRVGRVLPVRRQTMVCEPADGSFSPDLTLEGNRYDGYWQDERYFKAYRKEIEAFKFSFVGDG NRHVENMLRNERFASLHVRGDYLDNLYQNTCGIDYQYRAISQMNAMANPSCYFISDDIAWCKTHIEPL CEGHRPYIDWNKGKAYRDMQLMALCKYHIIANSFSWWGAWLNDADGEDGTITAPQQWYSHGNKPSPAS ESWIKV	PCT/US2015/030823

Bacteroides ovatus;Bacteroides ovatus 3_8_47FAA	WP_00430 3999.1	490431888	protein [Bacteroides ovatus]	26.25	MVVVYAAGLANKMFQAFSRGLMSHGLDVLDTQTSFQPEWSEFJALEEVPFNIEIKNAPNNMFSLAYKK DLISRIYRRMSAFFPNNRYLMERPHYDELIYKKATNNCFGLWQTELNFNCERDVRNRFVFTPFQDDQNI KLAEMKNENSVAIHIRKADYLRNIWDTCSVEYNNQAINYLKEHVSNPVFLFTDNPWEVENLKNIDY KLVDWNPVSGKQSYRDMQMLMSCAKHNIIANSTYSWWGAWLNNNPQKIVVAPKIWFNPKIEKAPYIPDR WIRL	WO 2015/175801
	WP_01995 5906.1	518799952	protein [Loktanella vestfoldensis]	26.23	MIITKLIGGLGNQMFQYAAGRSLAMRHGVPLLLDITELRSPKQGYQFEDVFAGRFEIAGLIRVLGRKAR KVPKTVAVVSPKPPMGDHDVWVRQRTDHYDAAFESIGADCYLSGFQSEKYFATIAPIQRESFRKEALTG ANAAIASRMKEAPSAAIHRRGDYVTDKGAHAFHGLCAWDYDAAIDHISRHEPDARFFVSDDVVAAQER FANRQRAEVVAVNSGRHSYRDMMLMAQCKHQIANSTFSWWAAWLNQNPDKIVVAPGTWFSGNDGQI KDIYCKDWIV	108
Flavobacterium sp. ACAM 123	WP_01699 1189.1	515558304	protein [Flavobacterium sp. ACAM 123]	26.14	MDVVIIFNGLGNQMSQAFYSQKKINNSTYFVPFKDHNGLLETVFSLNTKETLIQKSLYLIRLLTDRLKIV SDPLKWLILNLFKCKIVKESFNYNYPYLKPSKGITFYGGWHAKEYFAKENQQKSVFEETGDLGKINKEHV DIASTNAVSLHVRGDFMNEANIGLFGGVSTKAYFEAGAKLIATKVDHPHFFVFSNDMDWVKENLSMDT YVTCNSGKDSWKDMCLMSLCOHNIIPNSTFSWWGAWLNKNPKIVKIVVCPSPRFLNNDTYTIDYVDSWVKISD Y	109
Bacteroides fragilis;Bacteroides fragilis 3_1_12	WP_00577 9407.1	492219620	glycosyl transferase family 11 [Bacteroides fragilis]	26.1	MMKLVGMTGGLGNQMFYFIQIMKTIFPELRIDMSEMKKYKLNHNGYELEDVFSIRPQTISAHKWLKRIV YAFFSIIREKSEELSIIHKYTHKRWPLVYVYKGFQSEELFFKESDITRDFSFNTENANFRTKEWAKIKEQSSV SIHRRGDYTSAKNKIYGNICTEYQKASIIILKEPKAFFHIFSDDVETWTKAHLKIHHLPHQYISWNLKGPDS WQDMMLMSLCRHNIIANSFSWWGAWLNAYKDKTVIAPSRWSNVKKTTPHILPESWISDI	110
Spirosoma panaciterrae	WP_02059 8002.1	522086793	protein [Spirosoma panaciterrae]	26.09	MIISRVTSGLGNQLFQYAAARSLSLRNKTAFYVDLSYYLYEYPPDDTSRSFKLGFFSVYRILQESPVEYLSKSTKL FPNRSRPFELFLEKQHFDPITLQAHAGCVIMEGFWQSECVFRDHAEIRRELQLSKSPSEFEGYHQIQQA TPVPVSVHVRRGDVVNHPEFSKTFGFIGLDYKTAIRHLTKIKNPHFVYFSDDKKEWARANLPLPTDSVFVTN TGPSGDVADLVLMSTCHHHIIANSFSWWGAWLNPNPDKLVITPKLWYKNQPTWNTKDLLPPTWVSL	111
uncultured bacterium	EKE06672.1	406985982	glycosyl transferase family 11 [uncultured bacterium]	26.09	MIITKLTTGGLGNQLFQYAGRNLIIYNGSDLKLDVSEYDVSNKGNFRHYALDKFNTIQNFASKKETNNFKFGVF KKWLYKSGIVKNKNYFLEKKFNEDKEILKIDNAFLQGYWQSEKYFIGIRDLQKESLKENIELKFGEILKEINES NSVSHVRRGDYVKNPKNLFSHGVCSPKYSESTSKIASLIEKPVFFVSDDEWVWKENLITFPVAVVLSGINKIK SYEELVLMSCCKHNIIANSFSWWGAWLNTNQKIKVIAPKRWFNDVKLDTTDLIPENWIRI	2

PCT/US2015/030823

Thermosynechococcus elongatus; Thermosynechococcus elongatus BP-1	NP_681784.1	22298537	alpha-1,2-fucosyltransferase [Thermosynechococcus elongatus BP-1]	26.07	MIIVHLCGGLGNQMFQYAAAGLAAAHRIIGSEVKFDTHWFDATCLHQGLELRRVFGLELPEPSSKDRLRKVLGA CVHPAVRRLLAGHFLHGLRPKSLVIQPHFYWTGFEHLDPDNVYLEGYWQSERYFSNIAIIRQQFRFVEPLDP HNAALMDEMOSQSVSLHIRRGDYFNNQMRRVHGVLDSEYYPAAVATMIEKTNAERFYVFSDDPQWVL EHLKLPVSYTVVDHNRGAASYRDMQLMSACRHHIIANSTFSWWGAWLNPRDPKVVIAPRHWENVDVFD TRDLYCPGWIVL	WO 2015/175801
Colwellia piezophila	WP_01902 8421.1	517858213	protein [Colwellia piezophila]	26.03	MKIVKIAGFGNQLFQYAFYLADKKYAEQVCLDSLDMAKYRLHNGYELEGIFKLDARYCTEEQRIVRKDNIN IFTKLLSSKKKLGNNKNYILEPKQEHFTFHEKSFQQAQNTPTYYKGWQDVKYLENIIEELKSSLVFPEFELGKNI ELANFISSNSSVSLHVRRGDYVQHKAFGGICDLSYYQRAVEQINTLVKDPFIVFSDDIQWCKDNLNLEKAKFV DWNIGENSFRDMQLMTLCKHNIANSFSWWGAWLNANDDKNVICPKWVHYTSATGVLPSSEWIKIKAS V	114
Prevotella maculosa Salmonella enterica; Salmonella enterica subsp. enterica serovar Worthington str. ATCC 9607; Salmonella enterica subsp. enterica serovar Cubana str. CFSA N001083; Salmo nella enterica subsp. enterica serovar Cubana str. CFSA N002050; Salmo nella enterica subsp. enterica serovar Cubana str. CVM42234	WP_01996 6794.1	518810840	protein [Prevotella maculosa]	26	MKIVKIIGGLGNQMFQYALAILOERWKDEEIKLDLHGFNGYKHQGYQLDMLFGHRFEAATLTDVAQLA WPPYPHYQLWRVGSRLPKRRSMLCEPSKGLLPDSVLKQKGSYYDGYWQDERYFRAIRPQIMAAFKFPDFT DRRNLETEKRLKASEAVSIHVRRGDYLDVLFQGTCTNIAYYQRAIARLCQLKTPVFCIFSNDMAWCKVHIEPL LHGKEILYVDWNRGKESYRDLQMLTLCRHHIIANSFSWWGAWLSKAEDGTTIAPRHWYAHDAKPSPAE RWIKV	115
	YP_008261 369.1	525860034	fucosyltransferase [Salmonella enterica subsp. enterica serovar Cubana str. CFSA N002050]	25.99	MYSCLSGGLGNQMFQYAAAYILKQYFQSTTLVLD DSYYSQPKRDTVRSLELNQFNISYDRFSFADEKEIKILL RKFKNRPFPKQSEILSIALFGKYALS DRAFYTFETIKNIDKACLFYQDADLLNKHQLILPLFELRDDLLDICKN LELYSIJQRSNNTTALHIRRGDYVTNQHAAYHGVLDISYNNHAMEYVERERGGKQNFIFSDDVVRWAQKAFLE ENDNCYVINNSDYDFSADIMYLMSLCKNNIIANSTFSWWGAWLNKYEDKLVISPKQWFLGNNNETSIRNAS WITL	PCT/US2015/030823

Bacteroides sp. 3_2_5	WP_00865 9600.1	495935021	protein [Bacteroides sp. 3_2_5]	25.94		MKKVIFSGGLGNQMFQYAFYFLKKKGKAVIDNSLYSEKMHNGFELIKVFDIKESIYRTYFLKVHLFIKLLMK IPPVRKLSCKDDVIPIGDHEFDPPYARFYLYGWQSKIVNYVIEELRAQFIERNIPQMTEIKGDFLSSINSVSIHIR RGDYMGIPIYQGGICNEIYYERAISFMKEHFLNPRFYVFSNDSIWAFLLEKFDIDMEIIVTPPIYSYWDMLMS RCRNHIIANSTFSWWAAVLNINKDKIVISPTIFKKDECIDIIFDDWVVKISNI	7 WO 2015/175801
Clostridium sp. CAG:510	WP_02212 4550.1	547662453	protein [Clostridium sp. CAG:510]	25.86		MIMLQMTGGMGNQMFYALYRSRQKQKEVCIEDFTHYDTPEKNCLQTVFHLDIRKADREVYQRLTDSEP DFLHKVRRKLTGRKEIKYQEKDAIPEFEVFTQDDVYMGYFQSGRYFEKAVFDLRKDFTFEAWNTFPEKAKKLR EQMQAESSVSLHRRGDYMGNGKFIASIGNICTDAYEAAARRYMKEHFGDCRFYLTDDAEWGRQSESDT VYVDASEGAGAYVDMALMSCRRHHIIANSFSFWWGAWLDENPDKTVIAPAKWLNISEGKDIYAGLCNCLI DANGSVQGE	118
Rhodopirellula europaea SH398	WP_00866 5459.1	495940880	glycosyl transferase family protein [Rhodopirellula europaea]	25.86		MIVTRLIGGLGNLQFQYAFGHSLARSTYQTLIIDDSAFIDYRLHPLAIDHFTISASRLSDADRSRVPKFLRTPV GRAIDKVSRRVPGYQGVLPVRREKPFGRFRESILARESDDLGYWQSEKFFPGLRGLSIREEFQLRQSETRR RLSAQMKSENSVAIHVRRGDYVTSKAKAQYRTLDADYRRCLLDAAHETDLKLYFSNDVVPWCESNLDVGI PFTPQHTDGTATAHEDLHLIAQCRHVIVIANSTFSWWGAYLQGLHPTRRVYVPEPWFHPGTLDGSAMGCD DWISEASLEEQSSILKSSRRAA	119
uncultured bacterium	EKD23702. 1	406873590	glycosyl transferase family protein [uncultured bacterium]	25.82		MIIVKLKGGMGNGMFQYAGRNALATKLTQLDLTLLDRSPRKDFVRDYDLDFALDVAFAGPTDLKPF QFRISHLTKIYNIFPRLLGRPVYSEPHFSEAILKSSDNVLDGYWQSEKYEKEIENSIRDDEFKFRQPLEGRAA EMAAQJKNEDRAVCLNVRADFVTSKKAQEFHGFGLDYYQKAVDLLVSKVGPLHLFISSDDVDWCAANLK FNYPTTFVTKDYSGKKYEAYLQMLTLCRHYIIPNSTFAWWGAWLNSDPNKIIVAPKQWFKFKEASIDTTDIIPT WIRL	120
Bacillus cereus; Bacillus cereus AH1271	WP_00058 7678.1	446510160	protein [Bacillus cereus]	25.74		MIIVKLKGGGLGNQMFQYALGKSLALYYDKPLKIDADYIKNNEGYPVPRDFSLSKFNIELDLYQEADKERVGFILK NNFLAKKLRNYFLKKGKYKGYIENPDNLGLFKKELFENHNSMYIDGYWQSYLYFNIRECLIKFNLKPEYT KEMTEIMQRINETNSAVVHIRRGDYVKGWTLDTTYYKKAIAEIVKNDNPKFVVFSDDTDWVRSNLQELD NAVFIGECNLFQELWLMSTCKHNIISNSTFSWWGAWLNQNDHQVQVVSFAWINGMSVETTSIPDSW KRV	121
Firmicutes bacterium CAG:95	WP_02249 9937.1	548309386	protein [Firmicutes bacterium CAG:95]	25.74		MDIIRMEGGLGNLQFQYALYRQLQFMGRVTKMDVTTEYGREHQRQMLWAFDHYEEATQEEINRLTD GFMDLPSRIRKLTGRRTKKYAEADSNFDPQVLLKTPVYLTGYFQSEKYEKDFKDEGILHTELGFSDYDGEV ADQIRNYQKQIRETSVSLHVRRGDYLEHPEIYGMSCMTEYQAGVRYRERHDPDAEIFVTNDPVEKWL QENFLGDTLIQGTSEETGYLDLMLMSQCKHQIMANSFSSWWGAWLNPKNKDKIVVAPPEWFGDRNFHDI YTEEMIRISPRGEVKKHG	2 PCT/US2015/030823

Prevotella oris;Prevotella oris F0302	WP_00437 4901.1	490508875	alpha-1,2- fucosyltransfera se [Prevotella oris]	25.74	MIAATLFGGLGNQMFYATVKSLSHYQVPMAFNLNHGFANDYKYHRKLELCKFNCQLPTAKWITFDYRGE LNIKRISRRIGRNLCPNYQFVIEEPPHYEKRLEFETNKNIFLEGYWQSPCYFENYSKEIRADFLQKYPLSKEML EEIYALKATGKTLVMLGIRRYQVEGRDICTYKLCDEYYIKAITIYQERIPNALFVFTQDKEWATTHLPKGAE FYFVKDKQDEYATVADMFLMTQCTHAIISNSTFYWWGAWLQCTTKNHIVIAIPDSFINSDCVCKEWIILKRNS LC	1
Escherichia coli	AA037719. 1	37528734	fucosyltransfera se [Escherichia coli]	25.73	MYSLSGGLGNQMFQYAAAYILQRKLRSLVLDSDSYFLDCSNRDTRRRFEELNQFNCYDRLLTTSKEKEISII RHVNR/YRLPLFVTNSIFGVLLKKNYLPEAKFYEEFLNNCKLQVKNGYCLFSYFQDATLIDSHRDMILPLFQINEDL LHLCNDLHIYKKVICENANTTSLHIRRGDYITNPHASKFHGVLPMDDYKEAIRYIEDVQGEQVIIVFSDDVKWA ENTFANQPNYYVYVNNSECEYSAIDMFLMSKCKNNIIANSTYSWWGAWLNTFEDKIVVSPRKWFAGNNKSK LTMDSWINL	124

PCT/US2015/030823

Leeia oryzae	WP_01815 0480.1	516890767	protein [Leeia oryzae]	25.71	MIIVKIIGLGNQMMQYAFAHACAKRLGVPFKLDITAFESYKLPYPYGLHNFEITAPIASLEEIEHAKSMGVITE TSRFFDDSLVSADKGMVLDGYWADRYSESVMGELKPVFTLMDPLTPEQQAAMNLSAPNAVALHVRR GDVVTNPNCELLPQQYRDAIKLVLDQPDADVFCFSDDPDWVEAHLDPAPKWWVRGGIDNGFVDMIL MSKARHRIVANSTFSIWAASRLADQDGLTIVPSQFFRKDDPWLLQYGEVLQPCYPPQWVRVVDVTGDGKKE AENTSTALLQIAGGDVRGRKLRIGVWGFYEEFYQNNYFLNKNAPIGHELLKPFNQLYQYQGAHNLEFVTLDL VADLSTDAVLFFDAPNMRSPVSSVMQLDIKKYLCLECELIKRPDNWQQSLHELFRIFTWHDGLVDNHRYY KNNYVTDLMPWIESAQSALTAPFEETARKGYLQKKLCNIGNKLVSHPFELYSKRIEVRWFESHHPHFEDLYG MGWSASDYPYKGIKIDDKLEVLKGYRFSLCYENAKELPGYTEKIIDCFKAGWVPVSGAPNIADWIPDNCFIG SGKFPD DALTYLISMTEEVHADYLENIROFFLGGKAYPESADAFINTITRTIVQDCFLPHERTDVSVVVPNY NHGNFVSAITSALNQNVSVELLVLDNASTDDSWSQLQFFADYPQVRLRNWNIQVQHNWNNHATWLAT GRYVVMLSADDLILPGHLEQAVKRLDENPASSLYYTPCLWINEHDQPLGTNLNHPGHLESDYVGGRRDEISDLL KFDSYTPSAAVIRRETINRIGSMNLHLKGAIDWDLWIRIAEISP AFIRKQPGVCYRQHSNNNSVDFYASTAP LEDHIRIVESIIDRKVAVKYLKAKEEIIAHLNDRASSYPENQIQHLLSRINNIKDYLRKGAGPVISVIIPTKNRPGL LANALESITYQTFKDFEVVHNDDGGDIGIVDFSDQLQISYVRSOSGGAASRNRAKLAAGRIIAYLDDD DVYLDHLEKLVDAVYKGRSEKHYTNCEYLIQERKEGRUELGRRERYAGISYRAQLLVSNFIPTPTWSHTKELI DTIGDFDESLEILEDWDFLLRASKVTEFYQVNATTVEVRSRSDRDDHTLRANADKLLAYHQKIYAKHPVENESI LANRQSLNSLNRQDVTPKNENSYQGWVNARQPNELAVQILAERMMLQWSKQYQFMVIMVVMVVKQSQQ NILLANTIDSFCQQLYSGWKLVISDFEAPDESFINNEVLGWLTTETVEDENLLTQAFNGVLAEPVSDWVTILPV GTRLTSTALLKVGDRLLNGGACVIYTDHYVSDGMIKDPVLPKPAFNLDMLRSQDYIGSSIFFERTDSLAAVG GFASFPGARTYACFRMLDNYGPGQTIHLPPEVMTTFPENQPENSLRVAAMQLALEEHLHRNNISASIEEGYV TGTFVLVQYHHSEQPFVSIIPNKKHEFLAPCIETLMKVYQYPAFEVIVDNQSTDPTDLSYEEIESFEANNVK VIQYDNPNFSAQCNIAGAESARGDFILFNNDTEIVQANWLERMMQHAQRNDVGVVGARLVFPETVTIQH AGIVLGGKYPDEVFQPYMNFVDKDVSLNRTKVQNYSAVTGACLLVRKSLYQQVGGMNEQNLAVALYGD VDLCRLRQLHKSVMWTPFSTLVHHTGKTLNSNSDHEKHLMMVIQTRQEREYMLSHWLDIANDPYHRL DKSECNGTIDCTHTPLWDDIPSARPLQGMALVGGSGEYRVNMPFTLERSALAEIVLSNMITSKARLPSTEL ARNAPDVFEVWQNALADEFIRMLEMYKKYLPVSFRIQMLDLDLLEIPDASSFRHFKQKNWRDAKARLSKLF CDRLIVSTEPLRTFAEDMIDDIIVPNMLERSVWGDLSVKRRAGKKPRVGVWVAAQQHAGDLALMTDWKA TGHEVDWVFEQGMCPDDIRPYVAEVNTEWLTYDKYPOQIAALNLDIAIPLAINEAFNEAKSNLRLLEYGALG WPVICDIYPYQTNNAPVCRVPNDASAWIEAIRSHIADLADATAQKGDQLRQWVVDHYMIEDHAQEWSA LTRPAGK

WO 2015/175801		127		128		129		130		PCT/US2015/030823	
Desulfovibrio africanus; Desulfovibrio africanus PCS	WP_00598 4173.1	492830219	Glycosyl transferase family 11 [Desulfovibrio africanus]	25.68							
Akkermansia muciniphila CAG:154	WP_02219 6965.1	547786341	glycosyl transferase family 11 [Akkermansia muciniphila CAG:154]	25.66							
Dysgonomonas mossii; Dysgonomonas mossii DSM 22836	WP_00684 3524.1	493897667	protein [Dysgonomonas mossii]	25.66							
Prevotella oris; Prevotella oris F0302	WP_00437 2410.1	490506359	glycosyl transferase family 11 [Prevotella oris]	25.66							
Pseudogulbenkiana ferrooxidans 2002; Pseudogulbenkiana ferrooxidans	WP_00895 2440.1	496239055	glycosyl transferase family 11 [Pseudogulbenkiana ferrooxidans]	25.66							
Salmonella enterica	WP_00028 5641.1	446208786	protein [Salmonella enterica]	25.66							
Carnobacterium sp. WN1359	YP_008718 688.1	554649642	glycosyl transferase family 11 [Carnobacterium sp. WN1359]	25.59							

WO 2015/175801							MMKIIVNMACGLANRMFQYSYYLFLMHKGYNVKVDYNSAKLAHEKVAWNDIFPKARIEQASFSDLKSG GGSDVSKIRKYLPLFSSVVMPTAFDANLPVENKKLQYIIGVFQANMMVEAVEEDVKRCFKQFPFTDERNL KLQNEQSCSEVAIHVRKGDYQRIWYQNTCPYEQNAIRLSEKVNPNKLYVFTDNPWEVKEHFKDFPY TLVEGNPASGWGSHFDMQLMSVCKHNIIISNTSYSSWWSAFLNVHNEKIVIGPKVWFNPDSCSEFTSERILCK DWIAV	134
	WP_02196 4668.1	547227670	protein [Prevotella sp. CAG:1185]	25.58			MFQYAMASSVARRAGEILDLISWIRQMEKLSADDIYGLGIFSFDEKFTSNEVQKFLPSGKFSAKIYRAVN RRMPFSWRRVLEEGMGWHPQIMEIRRSVYFMGYWQSEKYSFDFIQEIRKDFTFREEVRQSIERRPVE KIRKSDAVSLHRRGDYAQNALGEILSFTMQYIDAARYSERVKTVPFFIFSDDIPWAKENLPLPYEVCYIDD NIQTNEREIGHKSGYEDMYLMTQCQHNIIANSFSWWSGAWLNHNPNKIWAPKKWCNGSFSNYADIVPE QWVKL	135
	WP_00964 5343.1	497331130	glycosyltransferase, family 11 [Selenomonas sp. CM52]	25.58			MEIVFIFNLGNQMSQYALYLSKRNLGCKVRYAYNIRSLSDHNGFELDRVFGITYPNLNFKNICINIYRLLFAN KYLFLVQKMIYVLRQMNVMYSIKEKNDYDYKILTRHKGIVLYGGWHSKGYFLSNADIIKDKFRFNISKLNS LVLYHRLSSNAVALHVRGDYMAPEHYNVFGCVCGIEYKAAIQIYQSLNPNVIFVNSDIEWVKNITGIQ MIFVDFNKKNSWMDMCLMSCCEHNIISNTFSWWSGAWLNHNPNKIWAPKKWCNGSFSNYADIVPE QWVKL	136
	WP_00748 6621.1	494751213	protein [Bacteroides nordii]	25.57			MKKDIIILRVWGGVGNQLFIYAFKVLISLITDCKVTILDIRTGFGANDGYKRVYRLGDFSLIPALRFYTLTSLFAQ RKMPYIRHLLAYKDFEEDQKPLETLDFFKIYSDKNLYLQGYWQYDFSSYRDVLLKDLRFEVEINNTLYY SDLIEKSNVAIHFRRIQYEPVISIDYKAIKAIYSENVENPTFFIFSDDINWCRENLSINGICFFVENFKDELYELK LMSQCNHFIANSNTFSWWSGAWLSVNADKKVIMPDGYTDVSMNGSIVHI	137
	WP_00563 5503.1	491855386	protein [Parabacteroides merdae]	25.54			MIIHQLKGGNLGNQMFQYALYKELKHRGRDVKIDDESFGIGDKLRVPVLDVRFVGYDRATKDEVIALTDSKMDI FSRIRRLTGRKTRFRIDEMEGIFDKILETENAYLVGYWQSEKYSFTSPEVIEQIEAFGRPOEIMHDSVSWST LQIECCESVSIHVRRTDYMDAEHIKHNLCSEKYYKNAISKIRREEHPNAVFFITDDKEWCKEHFKGPKFITVE LQEGEFTDVADMLLSRCKHHIIANSFSWWSAWLNDSPEKIVIAPSKWNNKMMDDIYTERMTKVAI	138
	WP_02276 8139.1	551024004	protein [Butyrivibrio sp. NC2007]	25.51			MIVVYSNAGLANRMFHYALYKALEVKGIDVYFDEKSYVPEWSFETTLMDVFPNIQYRESLQFKRASKKTF DKIVHCSNLFGRYVYVYRKYDDKLTLENTQDLCLGLWQSEKYSFMDVRQEIQKCFQYRSFVDDKNVKT AQOMLSNSVAIHVRKGDYQQRNIRWKNCTCTIDYRLAIDYIRMHVQNPVYVFTDNKDWVWVFTDLDY TLCWNPSTSGKQNYLDMQLMSCAKHNVIANSYSTSWWSGAWLNENSCKIVIAKRWVFNKIVTPDILPEQWI KI	139
	WP_00430 2233.1	490430100	protein [Bacteroides ovatus]	25.5				

PCT/US2015/030823

8

Mesotoga prima MesG1.Ag.4.2; Mesotoga prima	YP_006346 113.1	389844033	glycosyl transferase family protein [Mesotoga prima MesG1.Ag.4.2]	25.5	MRVVFGGGLGNQMFQYGLYCLKNNQNEVKADCTQYSTTPMNNGFELERLNLDAHANLDVSKLTG GNRLSPKVIWKLFRPKVYFEEKIPFSDPDVLKGNRRLKGVWQNMNYLEPCAKELRDVFTFPAFSSDNN KRLADEIAKVEAVGVHFRRGDFLKSSNLGLFGGICSDQYLRATQIMIENTVVEPVYVYFCDPQWAKNSFSD ARFTVIDWNIIGNSYRDMQMLSLCKHNIIANSTFSWWAAWLNRPNRVTIAPERMVNRDLDFSGIFPND WIRLQG	WO 2015/175801
Clostridium sp. KLE 1755	WP_02163 9228.1	545399562	glycosyltransferase, family 11 [Clostridium sp. KLE 1755]	25.49	MIVLKLQGLGNQMFAYAFARTIQEQKKKLLDTSDFQYDQKREYSLGHFILNENIEDSSGKFNLWYDQR KNPLLKVGFKFWPKFQQLKLFYVWDYAKYIPVDVSKKHKNILLHGLWQSDKYFSQSEIRKEFAVKDEP SQGNKAWLERISSANAVCVHRRGDFLAKGSVLLTCSNSYYLKAMEISKVNEPEFFIFSDIEDVKKIFEFP YQITLVNQSNPDYEEILRLMSKCKHFIIANSTFSWWSSILLSNEDKVIVAPRLWYSDGRDTSALMRDEWIIIDN E	140
Bacteroides plebeius CAG:211	WP_02205 2991.1	547321746	glycosyltransferase family 11 [Bacteroides plebeius CAG:211]	25.42	MDLVTLSGGLGNQMFQAFYWALKRGGKVFYKKNLAAKEHNGYELQTLFGVEEKCVDGGLWMTRLLGC PLLGKILHILPHKIRERVLYNYSIYLPFERNGLHWVGYWQSEKYFQDVADDIRRIFCFDHLNLNPAATSAALK CMSEQVAVSHRRGDYLLPCNVATYGGGLCTVEYENAIRYKERYPOAVFYVFSDDLWVYRENIPISAGKM VFVDWNRGKDSWQDMFLMSKCHNIIANSTFSWWGAWLNTHPEKLVIAPERWANCAPADALPDGWV RIEVSRR	141
Treponema lecithinolyticum; Treponema lecithinolyticum ATCC 700332	WP_02168 6002.1	545448980	glycosyltransferase, family 11 [Treponema lecithinolyticum]	25.4	MAIKVIGGGLGNQMFYAFACALQKCGHKVYVDTSLYRKATVHSGIDFCHNGLETERLFGIKFDEADTAD VRRLLSTSAEGLNRRIRRYFTKTHYIDTVFKYTPPELLSDKNDCYLEGYWQTEKFLPIEKDIRRLFTFRPTLSEKS AAVQSALQAQAAVLSASIHVRRGDFLTKTLNVCTETYYNNAIKYAVKKHAYSRFYFSDDIPWCREHLFCF NAHAVFIDWNTGNSWQDMALMSMCRNIIANSTFSWWAAWLNNASDKTVLAPAIWNRRLQLEVVDY YGYDSDIVPESWIRIPID	142
Bacteroides eggerthii; Bacteroides eggerthii DSM 20697	WP_00429 1980.1	490419682	glycosyl transferase [Bacteroides eggerthii]	25.34	MRLIKMTGGLGNQMFYAFYLRMKKRHTNTRIDLSMMHYNVHGYEMHVRVFNLPKTEFCINQPLKKVIE FLFFKKIYERKQDPSSLFPDKKYLWPLLYFKGYQSERFFADMENDIRIAFTFNSDLFNEKTQAMLTQIKHNE HAVSLHRRGDYLEPKHWKTTGSVCQLPYLNAITEMNKRQESYVYVFSDDIAWVKENLPLQAVFIDWVK GAESWQDMMLMSHCRHHIICNSTFSWWGAWLNPRENKTVMPEWRFQHCDDTPNIYPDGWIKVPVN	143
Bacteroides stercoris; Bacteroides stercoris ATCC 43183	WP_00565 6005.1	491891563	glycosyl transferase [Bacteroides stercoris]	25.34	MRFIKMTGGLGNQMFYAFYMRMKKHYSNTRIDLSMMHYKAHNGYEMHVRVFNLPPIEFIRINQPLKKVIEF LFFKKIYERKQVPSLVYDQKYPWPLLYFKGYQSERFFADMADDIRKAFATFNPRLSNRKTKEMSEQIDHDE NAVSIHVRGDYLEPKYWKTTGVCVQLPYLNAIAEMNKRISQPSYVYVFSDDIAWVKENLPLKAFIDWVK GAESWQDMMLMSRCHHHIICNSTFSWWGAWLNPRENKTVMPEWRFHRCETPDICPDKWIKVPINQPD SIQ	PCT/US2015/030823 4

Pseudorhodobacter ferrugineus	WP_02270 5649.1	550957292	alpha-1,2- fucosyltransferase [Pseudorhodobacter ferrugineus]	25.17	MIVMQIKGGLGNQMFQYAAAGRALSLQTMPLHLDRYYRREREHGYGLGAFNIEASPLDESLLPPLPRESPL AWLIWRLGRRGNLVRENGMGFNPTLSNVTKPAWITGVFQSERFYAAHAAATIRAEITPVAAPDLVNARWL AEIAAEPRVSLHVRGDDYVDAKAAAHGSCPTAYYERALAHITARMGTAPVYVAFSDDPAWVRENRLRP AEIRVPGHNDTAGNVEDLRLMSACRHHIVANSSFSWWGAWLNPRADKIVASPARWFAFADPAFTNPDIWPE AWARIEG	WO 2015/175801
	YP_002329 683.1	215487252	fucosyltransferase [Escherichia coli O127:H6 str. E2348/69]	25.16	MMYCCSLGGLGNQMFQYAAAYILKQHFPDTILVLDSSYFNQPKDQKQIRHLELDQKQIFDRFSSKDEKVKIN RLRKHKPILLNSFLQFTAKLCKNYSINDASYNPESIKNIDVACLSFYQDSKLLNEHRDILPLFEIRDDLRVL CHNLQYSLITDSKNTSIHVRGDDYVNNKHAACHGTLMDYYSAMEYSECGSQTFIIFDQVWAKEKFS KYSNCLVADADENKFSVIDMYLMSLCNNNIANSYSTSWWGAWLNRSEDKLVIAKQWYISGNECSLKNEN WIAM	158
	WP_01635 9991.1	511537894	protein [Lachnospiraceae bacterium 3_1_57FAA_CT1]	25.16	MIIKVMGGLGNQMQYALYEKFSIGKNVKLDISWFEEDSSVQEKVFARRSLELRQFKDLQFDTCSAEKEA LLGSGILGKLERKLIPARNKHFEYSDIYHSEVFNMSDAYLEGHWACEKYVHDIMPLLQEKIQIFPESANSQNT VKRMIKAENSVSIIHRRGDDYLDPENEAAMFGGICTNSYKAAEYIKSRVPDTHFYLFSDDTAYLRENYHGDY TIVDWNKGEDSFYDMELMSCCRHNICANSTFSFWGARLNRTPDKIVIRPAKHKNSEIEPQLLHELWDNW VIIDGGRIV	159
Butyrivibrio fibrisolvens	WP_02275 5397.1	551010878	glycosyl transferase [Butyrivibrio fibrisolvens]	25.09	MKPLVSLIVPVLNVEKYLEQCLTSISSQTYDNFEVLVWGKCIDNSENICKWCEKDHFRFIEPQLKSCLYARN VGIDAAKGEYIAFCDDCITSDFLSCFVDTALKNSSDIVETQFTLCDQNLSPYDYDRNHLGHILGHGFLEYTSA PSWWKYFVKRDIFTSNLHYPERFEGEDISMYSLFSYCNKIDYVEKPTLYRQVPSLLMNNPQGKRKRYESLF DIIDFVTNEFKTRLLFQKSWLKLFLQLEMHSASISDSATSDDDEAISMQRQEISGYLKKVPVKNTIFEVTALGWG GEIVSSIAKFNTHGVSSNNMFNRYFFELLESTRKKLEEMIINFSPDIFLIDISEADYLSYKGNLGTFFVKNW KIGSIFMKMIQTHSNSSIFLLENYMQQAPDHVDNTNLEIKMLYDDIKINHPDIICISAPADILNRSSEPELPCI YQLKL VSDKLHTMYSVPVINCETKGGNGQIFQYVFSKYIEKMTGYRPLHIGFFDYVKAIPGGTKRIFSLDKLF PDIEITSGKIP CSHVVEEKSFSINPGSDIFRYGYWQDIRYSDVKDEVLESFNVDTSMSKDVDFADTIRNANS IAMIHRRQDYLNENNVSLEFQSLSDYKSAVDMIRKEYADDLVLFISSDDPEYANSIADSFIEGFVMP LHKDY EDLYLTLAHHHIANSTFSLWGALLSARKDGIRIAPRNWFKGTPTNLYPDKWLLI	3
Anaeromusa acidaminophila	WP_01870 2959.1	517532751	protein [Anaeromusa acidaminophila]	25.08	MFCVRIYGGGLGNQMFQYALGRAMAKHYSETAAFDLSWYEQIKPGFEASVCQYNIELSRKDRPKAWYEPIL KRISRHTDKLEMMWFGLFFEKYHYDSTVFERGLCKKNITLDGYWQSYKFSIAEDDLRRELTIPKEREELIAISRS LPENSVSIHVRGDDYVSNPKANAMHGTCSWEYQAAIEKMTGLVKEPQYVAFSDDTWTKENLPLPNAMYI GRELGFDYELIIMSRCKHNIMANSTFSWWGAWLNSNPKNVVIAPRKWFRHKIKVNDLFPSSWVVL	1

WO 2015/175801	163	MDIVIFNGLGNQMSQAFYLAKKNDNLNCHVIFDPKSTNVHNGAELKRVFGIELNRNYLDKIISFYGYIFN KRIVNKLFLSLVGIIRMIYEPKNVDYREELLKPSNFSFVWGGW/HSEKYFKDIELEVKVFKPEVTNSPYFTEWF NKIFLDNNSVSHIRRGDYLDPSPDYQFNGVCTIDYKEKAILYKERILEPNFYFSDINWCMKTFGTENMY YVDCNKGKDSWRDMYLMSECRHHINANSTFSWAAWLSPPSYNGIVLHPKYFIKDIETKDYYPQKWIMIE	glycosyl transferase family 11 [Bacteroides sp. 2_1_16]	496044479	WP_00876 8986.1	Bacteroides sp. 2_1_16	25.08		
164		MDKVVVHLTGGLGNQM/QYALGRSISNRNCP/LLNTSFDYTDYDFKFCGLSRYNVKAFFIKNSYNNKYYRY VIRLLSRYGACVFGSYEKKIFSDYDEKVKRSCSVYGTWQSYGYFSDIRDLRDYEMVVGCEEEVEKYVSDIK RVDSVSLHRRGDYFDNKRQSHIGILTMEYKAMSLFPDSSVFVSDDIEWVRENLTNTNIVVW/LESDN PENEIYMSLCKNNIISNSTFSW/GAWLNNKKNYKVIAPRMWYKDNQSSDLMPSDWCLI	glycosyl transferase family protein [Chlorobium phaeobacteroid es BS1]	189500849	YP_001960 319.1	Chlorobium phaeobacteroides lorobium phaeobacteroides BS1	25.08		
165		MIVISMGGLGNQMFEYAFYTLKHLPKSEIKVDTKYAFPYSHNGIEVF/KIFGLNPPEANWKEVHSLVKTP IEGNKAHFIFLIRILKANLVREPTSCQKQDTEFYNSEFFELPQNKSFYLYGPFVNYNYFAAIHNEIMDLTY FPEITDVTNIEYKRIESSHSIHIRRGDYTEGVPLVPDARYREALVYINKKIEDPHFFVFTDDKDYCKSLFSDN QNFTIVEGNTGANSFRDMQLMSLCKHNIJANSTFSW/GAFLNKNSEKIVAPNIAFKDCSCPICPDWII	protein [Treponema bryantii]	551312724	WP_02293 2606.1	Treponema bryantii	25.08		
166		MVIKLFGLGNQMFIYAAKQIAJNQKLTFTDITGTFEDDSRFRVVELKQNLVSQESRRWMSFRYPLG RILRKISRKIGFCIPLVNFKFIVEKPYHFQNEIMRIASFSSYILEGYFQSYKYSKIEAQIREDFKFTKEVIGSVEKE ASFITNSRYTPVAIGVRRYSEMKGEGELAVWEHDYDAIKYIANKVPNLIFIV/SEDDIDWVKKNLKLDYPVVF VTSKKELAAIQDMYLMSLCNHIIISNSFFYWW/GAYLASTNNHIVIA/PSVFLNKDCTPIDWVII	LPS biosynthesis alpha-1,2- fucosyltransferase [Bacteroides fragilis 638R]	375358171	YP_005110 943.1	Bacteroides fragilis; Bacteroides fragilis 638R	25		
PCT/US2015/030823		MSGGLGNQMFEQYALYMKLTAMGREVKFDDINEYRGEKAW/IMLAVFGIEYPRATWDEIVAFDGSMDFS KRLKRLFRGRHP/IEVVEQGFYDPKVLSEFNMYLKGFSQSQRYFEDILEEVQETTFRPELKMNLPAPLYETTEK YLLRIEGCNAVGLHMYRGDSRNEELYDGTCTEKYEGAVRFIQDKCPDAKFFISNEPKWVKGWVISMKS QIREDMSREEIRALEDHVLIENNTETGYLDMFLMSRCRHHIISNSFSW/WAAFINENPDKLVTAPSRWVN GVPSEDVYVKGMTLIDEKGRVERTIKE	protein [Firmicutes bacterium CAG:534]	547951298	WP_02235 2105.1	Firmicutes bacterium CAG:534	25		
		MVIVKIGDGLGNQMFMVYCGYSVAKHDNDTLTLLTSDVDNSTLRTYDLDKFNIDFTDRESFTNKGFHKKVYK RLRRSLKVNVIYESTENCPCVLDVYRRKFRDKYLGHYFQNLVYFKTCKEDIMRQFTPKPEPSAKADELIHRA TENTCSVHVRGGDIKPLSIKYYKDALDKIGEAKKDMRFVFSNVRNLAEEYIKELGVDAEFWDLGEFTDIEELF LMKACRRHILSDSTFSRWAALLDEKSEEVFPFSPDADKIYMPPEWIMEEVDGNEEKR	glycosyl transferase family 11 [Firmicutes bacterium CAG:882]	547971670	WP_02236 8748.1	Firmicutes bacterium CAG:882	25		

Vibrio parahaemolyticus; Vibrio 10329; Vibrio parahaemolyticus 10296; Vibrio parahaemolyticus 12310; Vibrio parahaemolyticus 10290	WP_00549 6882.1	491639353	glycosyl transferase family 11 [Vibrio parahaemolyticus]	25	MVIVKSGGLGNQLFQYAGCAISNRSLCELLDTSFYPKQSLRKRYELDKFNIAKAVATQKVFSCGGGDDLLS RFLKLNLSLFFPNYIKESLVYLAELSHCKSGFLDGYWQNPQYFSDIKDELVKQIVPIMPLSSPALEWQNI INTKNCVSLHVRGDDYVNNHTNSVHGVCDLSYREAITNIHETVEKPKFFVFSDDISWCKDNLGSLGHFTYV DNTLSAIDDLMLMSFCEHHIANSTFSWWGAWLNDHGITIAPKRWFSSVERNNKDLFPEKWLIL	WO 2015/175801 168
Herbaspirillum frisingense; Herbaspirillum frisingense GSF30	WP_00646 3714.1	493509348	glycosyl transferase family protein [Herbaspirillum frisingense]	24.92	MIVSRLLGGLGNQMFQYAAAGRALALRRGVPAFIDSRAFAADYKTHAFGMQCFCADQTEAPSRLLPNPPAEGR LQRLRRFLPNPLRVYTEKTFTEAVLSLDPDGYLDGYWQSEKYFADFADDIRKDFAVKAAAPSAPNQAWLEL IGRTHSVSLHRRGDDYVSNAAAAAHHGTCDLGYVERAVAHQHQTGOAPELVFSDDLDDWVATNLQLPYT MHLVRDNDAAATNFEDLRLMTACRHHIVANSFSSWWGAWLDRSESITIAPARWFWADTPDARDLVPPQR WVRL	169
Rhizobium sp. CF080	WP_00775 9661.1	495034125	Glycosyl transferase family 11 [Rhizobium sp. CF080]	24.92	MIITRILGGLGNQMFQYAAAGRALAIAEAEKLDIEMGAYKLRPFALDQFNIAAIAQPDEVPAKPKRGLLR KFTSAFKPDRSSCERIVENGLTFDSRVPALRGSLHSGYWQSEQYFASSADAIRSDFSLSKPLGARQDVLARI GAATTPVSHVRRGDDYVTPNSANAVHGTCEPPWYHEAMRRMLDRAGDASFFVSDPEQWARDNLQSSRP MVFEPPQNGRGGEDMHLMAACHAHIIANSFSSWWGAWLNRPNKHVIAPRQWFRAPDKDDRDIVPA TWERL	170
Verrucomicrobium spinosum	WP_00995 9380.1	497645196	glycosyl transferase, family 11 [Verrucomicrobi um spinosum]	24.85	MVISHISEGLGNQMFQYAAAGRRRLSYHLGTTTLKDDYHYRLHPFRSFQDRFLITSPIATDAEISHLCPLEGLAR AIRARLPGLRGATRLRLGNLGLSPYQPRHLHSFKETPKQPLIGKVVSRHFHFDPDVLECPDNVCLVGYW QDERYFGEIRDILLRELTLSKPPAGATKAVLERIQRSSVSLHVRGDKTKSSYHCTSLLEYCLAAMSEMRARL QAPTEFFVSDDDWDWVREQIPCSSSVIHVDHNRAEDVSEDFRLMKSCDHIIASSSLSWWWAAWLGTTNENSF VFSPPADRWLNFNSHFTADVLPPLHWIQLDGSLLPAQ	171
Fibrella aestuarina; Fibrella aestuarina BUZ 2	YP_007319 049.1	436833833	glycosyl transferase family 11 [Fibrella aestuarina BUZ 2]	24.83	MTANRVLVNSPMVIAKITSGLGNQLFQYALGRHLALQGNLSLWFDRLYHFHQEYATDTPRKFKLDRFNVRYN LLDSSPWLYASKATRLPGRSLRLIDTRFEADHFDPVIRPAAPLTLWGFQWQSEKYFAQSTPQIRQELTFN RPLSDTFVGYQQQIEQAEVPSVHVRGDDYVTHPEFSQSGFVGLAYYQKALAHQLDLPNATLFFSDDPD WVRANIVTEQPHVFVQNSGPDADVDLQLMSLCHHHVIANSSFSWWGAWLNRPNRDKVVIQGRWFAN KPWDTDKLPLSGWLRL	2

PCT/US2015/030823

Coleofasciculus chthonoplastes; Coleo fasciculus chthonoplastes PCC 7420	WP_00610 0814.1	493031416	Glycosyl transferase family 11 [Coleofasciculus chthonoplastes]	24.73	MLSINKNLFVHIPKSCILKEVYIMISFPNLGKGVRLGNQMFOYAFRLSTARRLGKVFYCPAWSGDSLFTLN DQEEVVSQPEGITKQYRQGLNPGFSENAISIQDGTGTEISGFQSDKYDNDPDLVRQWFSLKKEKIASIRDRFSRL NFANSVGMHLRFVGDVVGQIKRPPMRRSYKKALSYPNQELILVFSDEPERTKMLDGLSGNLFSLGKHNY EDLYLMTKCQHFICSVSTFSWWGAWLGGERTVYIPKEGQYRPGYGRKAEGVSCSWIEVQSRIGFLDDY RLVSRLKRLPKSLMNFY	WO 2015/175801
						180
Bacteroides gallinarum	WP_01866 6797.1	517496220	glycosyl transferase [Bacteroides gallinarum]	24.66	MRLIKMTGGLGNQMFIYAFYLRMKKRHTNTRIDLSMMHYNVHHGYEMHHVFNLPKTEFCINQPLKKVIE FLEFKIYERKQDSSNLLPEDKKYFWPLLYFKGFYQSERFFADMENDIRKAFNSGLFNEKTQTMLKQJHNE HAVSLHVRRGDYLEPKHWKTTGVCQLPYINAIAMNRRIEQPFYVYFSDDIJAWVKENLPLPQAVFIDWN KGVESWQDMMMLMSHCHRHICNSTFSWWGAWLNPKENKTVMIPERWFQHCETPNIPYAGWKVPIN	181
Firmicutes bacterium CAG:882	WP_02236 7483.1	547967507	glycosyl transferase GT11 family [Firmicutes bacterium CAG:882]	24.66	MNNVEIMGGLGKQLFYAFSRYLQKLGKVNVLKDKDFTTQFPENNGITKREFVLDKYNTRYVAAAGEKTYR DYCDENDYRDDYAGSDEVLYEGYWQNFIDFYNNVRKEMQEELKLPKPEFIDNSMAAVEKD MSSCNSVALHIR RSDYLTQVNAQIFEQLTQDYASAVSIEQYTHEKPVLYFSDDPPEYAAENMKDFMGCRVTIMPPCEPYQDM YLMTRAKHNIANSTFSWWGATLNNANPDNITVAPSRWMKGRTVNLVYHKDWITL	182
Bacteroides xylanisolvans; Bacteroides xylanisolvans CL03T12C04	WP_00802 1494.1	495296741	protein [Bacteroides xylanisolvans]	24.6	MIAVNVNAGLANQMFIYAFGRGLMAKGLDVCDFQSNFKPRSQWAFELVRLQDAFSPIDIKVMPEGHFK WVFPSLPRNGLERREFQEFMKKWHNFIGDEVYIDEPYGVDPMEKCATRNCIYKGFQWQSEKFRHCEDDI RKQFTFLPFDELKNIEVAAKMSQENSVAIHRLKGGDDYMQSELMGKGLCTDYVMKAIDYMRKHNNPHFY VFTDNPVWVNDLPEFEYLVDWNEVSGKRNFRDMQLMSCAKHNIIGNSTYSWWAAWLNANQDKIVVG PKRFFNPINSFFSTCDIMCEDWISL	183
Geobacter sp. M18	YP_004197 726.1	322418503	glycosyl transferase family protein [Geobacter sp. M18]	24.58	MIGMVIFRAYNGLGNQMFIYALGRHLALLNEAELKIDTTAFADDPRLREYELHRLKVQGSATPDEIAFFREM ENTHPOAYLRLTQKSRLFDPAILSARGNIYLGFWQTEKYFADIREILLDEFEPVPAGEDSIKYLSHMKATNA VALHVRSDYVSNPMTLRLHHGVLPDYREAVRRIAGMVPDPVFFISDDPQWAKDNIRLEYPAFCVDAHD ASNGHEDLRLMRNCKHFIANSFSWWGAWLSQNTGKVVAPLKWFAKPEIDTRDIVPLQWIRI	
Ruegeria pomeroyi DSS-3	YP_168587.1	56698215	alpha-1,2-fucosyltransferase, [Ruegeria pomeroyi DSS-3]	24.57	MITTRLHGRNLGNQMFOYAAARGIAARLGTQVALDTRLAESRGEVLTFRFEDLDLAQPDQLPPLKGDGLLR HGAWRLGLLAPFRFRHGLGYNAAIETWDDGTLYLHGYWQSERYFAHIAARADFAFPFNSQNAEMAA RIGDTDAISLHVRRGDYVALAAHTLCDQRYAAALTRLEGVAGDPVYLFSDDPAWARDNLALPVQKVWV DFNGPETDFEDMRMLSLCRHNIIGNSSFSWWAAWLNNAHNAHPGKRIAPASWFGDAKLHNPDLPLPDWLKIE V	

Lachnospiraceae bacterium 28-4	WP_01629 1997.1	511037973	protein [Lachnospiraceae bacterium 28-4]	24.52	MIIQLAGGLGNQMQQYAMYQKLLSLGKKVKLDSWFEENRQKNVYARRELELNYFKKAEYEAETEEERKALVGEFGAGKIKGKLFPGTRKIFRETEMYPHPIEDFEDRYLYGFACEKYADIMEILQEQQVFPPSGNPNQKMAERIADESUSLHIRGVDYLDAAENMAMFGNICTEYAGAIKEMKKYPSAHFFVFSDDIPYAKETYSGEEFTVVDINRGKDSFFDIWLMVSGCRHNICANSTFSFWGARLNRNKGKVVMRPFHKNQSKFEPELIMHELWKGWVFIDNRGNIC	WO 2015/175801
Prevotella sp. CAG:1092	WP_02198 9703.1	547254188	glycosyl transferase family 11 [Prevotella sp. CAG:1092]	24.49	MRILVFTGGGLGNQMFEYAFYKHLKSCFPKESFYGHYGVKLEHYGLEINKWFDVTLPPAKWWTLPPVVGILFYLYKKLVPNSKWLDLQREWKHKDAKVFFPKFTKQYFPKENGWLKWVDEASICEKNKKLLQVHDEETCFVHVRRGDYLASNFKSIFEGCCTLDYYKRALEYMKNKNNPKVRFICFSDDLWMMRKNLPMDDSAIYVDWNTGTDSPLDMYMMISQCDNGIANSFSFYWGAYLGGKKTIVYPQKWVNNMEGNNPFIEMDEWLGM	186
Spirosoma luteum	WP_01861 8567.1	517447743	protein [Spirosoma luteum]	24.41	MVISLGGGLGNQLFOYAFGLKLAALQQLTELRLERHILLESKAIARLQYTPRTYELDTFGVEAPAAASLMDTVSCLSRVALSDKTALLRESTLTPNAINNLRVRDVVCLGYWQSEEFERPATEQLRKHLVFRKNPAQSRSMADTILSCQNAAFVHIRRGDYVTNTHANQHHLGLCDVSYRRACEYVKECIPDVQFFVFSDDPDWAKRELGIHLQPARFIDHNRGADSWQDMYMLMSLCRHAIVANSFSWGWALNPVAERLVVAPGQWFWFNQPVLSQQIIPPHWHCI	187
Marinomonas posidonica IVIA-Po-181; Marinomonas posidonica	YP_004480 472.1	333906886	glycosyl transferase family protein [Marinomonas posidonica IVIA-Po-181]	24.34	MIIVDLGGGLGNQMFOYACARSLSELNPLKVYVGLASQTVHNGYELNRVFGLDLEFATENDMQKNLGGFLSKPILRKIFSKKPLNLIKQNFPPENFVNSLSFYIKDSGFLQGYWQTEKYLNLHKSQILKDFCFVNMDDDETNSIANDIQSGHSIHHVRRGDYLTNLKAKAIHGHCSLDYLLKAIFFLQEKIGESRLFISSDDPEWVSENIAITRFS DVSVIQHNRGVKSFNDMRLMSMCDHHIIANSFSFWGWALNPQSNKKIAPKNWVFTDKMINTIDLIPSSWILK	188
Bacteroides; Bacteroides sp. 4_3_47FAA; Bacteroides sp. 3_1_40A; Bacteroides dorei 5_1_36/D4; Bacteroides vulgatus PC510; Bacteroides dorei CL03T12C01; Bacteroides vulgatus dnLKV7	WP_00583 9979.1	492425792	glycosyl transferase family 11 [Bacteroides]	24.32	MKIVVFKGGLGNQLFOYAFYKLSRKDETFYFNDAWVNVSHNGFELDKYFKTDDLKCSRFWIIILFKTILSKLYHWKIYVVGVSVEYQYPNHLFOAGYFLDKKYYDENTIDFKHLLSEKNQSLKDIQNSNSVGVHRRGDYMTKQNLVIFGNICTQKYYHDAIRITEKVNDAVFYVFSDDISWVQTHLIDIPNAVYVNVNWTGESSYDMYMLSSCKYNIIANSTFSYWAARLNKKTNNMVIYPSKWYNTFTPDIFPESWCGI	189

PCT/US2015/030823

WO 2015/175801	191	protein [Candidatus Pelagibacter ubique]	519013556	WP_02016 9431.1	Candidatus Pelagibacter ubique	MTIRIKLTGGLGNQMFQATGFAIAKKNVRLSLDKYINKRKLNFNGFELQKIFNIYSKVSFLNKTLSFKSINFTE ILNRIDTTFYNFKEPHFHYTSNINLPLKHSFELDGWQSELYNEFAIEIKRIFNFSGLDKSNLLVADDINRNSI SIHIRRGDFLLKQNNHHTDLKEYYLKAINETSIFKNPKYFIFSDDTSWTVDFVIDHPYIIVDINFGARSELD MYMLSLCKSNIAANSFSSWWSAWLNNKDKIYAPKNWFNDKSICTDDLIPESWNIL
						24.32
192	191	uncharacterized protein [Bacteroides sp. CAG:875]	547952428	WP_02235 3174.1	Bacteroides sp. CAG:875	MSVIINMACGLANRMFQYAFYLYLQKEGYDAVVDYFTRADLVHENVDWLIRFPEATRRATARDIRKMG GHDCFSRLRRKLLPMTTKVLETSGAFEIILPPKNRDSYLLGAFQSAKMVESVDAEVRRIFTPEFESGKNQYFQ TRLAQENSGLHIRKGKDYQERIWYKNTCGVEYYRKAVDLMKEKVDSPSYFTDNPWVKNLSWLEYKL VDGNPGSGWGSHCDMQLMSLCKHNISSSYSSWWGAYLNTLNKIVVCPRIWFNPSTKDFSSNPLAEG WISL
						24.29
193	192	protein [Butyrivibrio fibrisolvens]	551011911	WP_02275 6327.1	Butyrivibrio fibrisolvens	MIIILQGGGLGNQLFLYGLYKLNHLKRDVKMDIESGFEGDELKPKCLDCMNLEYAIAATRDVTDIRDSYMDI FSRIIRKITGRKTFDYEPEDGNYPDKVLEMTKAYLNGYFQSEKYFGDEESVKALKDELTKGKEDILTSTDLTKI YHDIKNSVSLHIRRGDYLTPGIITYGGICTDEYYDKAIAIRETPEARFFISNDIEWCKEKFAGDKNILEV NTGINLSEDNKIGKSDKIDISEYRDLAELYLMSACKHILANSSFSWWGAWLSDHEGMTIAPSKWLNKN MTDINTKDMILLI
						24.29
194	193	glycosyl transferase family protein [Roseburia hominis A2-183]	347532692	YP_004839 455.1	Roseburia hominis; Roseburia hominis A2-183	MVTVKIGDMGNQMYNACGYAAAKRSGEKLRLDISECDNSTLRDYELDHFRVWYDEKESFPNRTFWQKL YKRLRRDJRYHVIRERDMYAVDARVFPARRGRYLHGYWQCLGYFEEYLDLREMTTPAYEQTDAYRELM QQFTQTPTCALHVRGGDLGGPNRAYFQQAIAIRMQKEKPDVTIVFTNDLPKAKECLDDGEARMRYIAEFGE ALSDIEFFLMSACQNIISNSTYSTWAAAYLNTLPGRIVIVPKFHGVEQMALPDWVLDGGACQKGEIDAV
						24.22
PCT/US2015/030823	194	alpha-1,2- fucosyltransferase [Rhodopirellula europaea]	495934621	WP_00865 9200.1	Rhodopirellula europaea; Rhodopirellula europaea 6C	MATSVHPHLSGDKQALDSKAAQQVCSTQAASASDRACITISISGGLGNQMLQYAAAGRALSIHHDCPLQLDLK FYSSKRHRSEYELDAFPIQAQRWIKPSFFSQVLDKIQGESKAPTVEEQSKRFDRAFFDIELPARIRGYFFSEKYFL PYADQIRTELTPVPPLDQPARDMAQRLSEGMSLSLHVRRGDYVSNANARQRFWSCTSEYFEAAIEQMPAD STVFESDDIEWAKQNISSRPTVYVNDLKLGPETGLRDLWLMTHAKSHIIANSFSSWWGAWLSDGSEA NLTIAPKKWFNDPEIDDSDIVPTSWRRI
						24.16
PCT/US2015/030823	194	protein [Rudanelia lutea]	518832653	WP_01998 8573.1	Rudanelia lutea	MVIAKITSGLGNQLFYALGRHLAIQNQTRLWFDLRYHRTYETDTPRQFKLDRFSIDYDLDYSPWLYVSKA TRLLPGRSRLPFLDTKPEPHFLDPAVPNAKGAFTLDGFQWQSEGYFASNAATIRRELTFTRQPGPMYARYR QQIEQTQTPTVSVHIRRGDYVSHPEFSQSGALDDTYQTALAQINGQFPDATTLLVFSDDPEWVRQHMFRER PHVLVENTGPDADVDLQLMSLCHHHIIANSFSSWWGAWLNPDPKRVIAPKQWFRNKPWNTADLPAG WVRL
						24.16

Bacteroidetes;Capno cytophaga sp. oral taxon 329 str. FO087;Paraprevotella clara YIT 11840	WP_00861 8094.1	495893515	glycosyl transferase [Bacteroidetes]	24.15	197	WO 2015/175801	MRLKMTGGGLGNQMFYAMYKMKITFPDVRIDLSDMVHYQVHYGYEMNKVHFLPRTEFCINRSLKKIEFL LFKTLERKQGGSLVPYTRKYHWPWYFKGYQSEKYFAGIEKEVREAFVDIRRASRRSLRAMQEKADPHA VSIHVRGDDYLLKHKWKGALGCOSYYLNALAELEKRVKHPHYVSEDLNWWVRQNLPLIKAEFIDWKNKGE DSWQDMMMLMSHCRHHIICNSTFSWWGAWLNLPLDKVIAPERWTQITDSADVVPESWLKVSIG
	WP_01815 9152.1	516906936	protein [Smaragdicoccus niigatensis]	24.08			MADVVTLAGGLGNQLQTAYAKNLEARGHRVTLDTGTVRWTRGLHIDPQICGLKILNATPPAPVPGRLLAA TVLRALATRLRFGDGRIVRQTORTLEFDEQYLNLSNPGRYVEGYWQCERYFSDVGQTVRKVFLDMLGRH VSYNGLSRLPAMADPSSISLHVRRGDYVTANFIDPLALEYERALEELAVPSPRIFVFSDDLDWATRELGRICD VIPVEPDWTSHPGGEIFLMSQCSHHIANSFSSWWGAWLDGRTSSRVVAPRQWFSLETYSARDIVPDRWT KV
Bacteroides fragilis CAG:558	WP_02201 2576.1	547279005	family 11 glycosyltransfer ase [Bacteroides fragilis CAG:558]	24.05	198		MIHLGGGLGNQMFQAFARSLALQYNENISFNITLYKELKNEERFSLGHLNINTMCIVETPDENKRIWELF NKQIFHQKIAKILPASIRWWWMSNRNIYANVCGPYKYHPRHSQNTTIIHGGFQSWKYFKEHQSMIKAE LKVITPISEPNKKILKEIONSNSICVHIRRGDFLSAQFSPHLEVCNKDYKEKAIMISSQJENPTTFIFSNTHEDLV WIRKNYNIPQNSVVVDLNNPDYELRLMYNCKHFILSNSSFSWWAQYLSKSNKIIAPKIWDKRGIDFSDIY MPEWIIIK
	WP_02265 7592.1	550904402	protein [Desulfovibrio desulfuricans]	24.05			MSFIDVAAIQRMALVKVDGGLGSMWQYALSLAVGKSSFTVKHDLWFRHYAKDIRGIENRFFILNSVFT NINLRASENERLFFHIALNRYPDSICNFDPDILALKQPTLYGGYVNAQVVTSAKEIREAYVFAPAVEESNQA MLQTHAAPMPVAVHVRGDYIGSMHEVLTPRYFERAFKILAAALQPKPTFFVFSNGMEWTKKAFAGLPYD FVVDANDNDNDNVAGDLFLMTQCKHFTISNSSLSWWGAWLSQRAENKTVIMPSKWRGGKSPIPGECMRV EGWHMCPVE
Hoeftia phototrophica;Hoeftia phototrophica DFL- 43	WP_00719 9917.1	494373839	alpha-1,2- fucosyltransfera se, [Hoeftia phototrophica]	24.05	200		MHGGGLGNQLFQYAVGRAVALRTGSELLDTRFTSSNPFQYDLGHFSIQAKVANSELPPGKNRPLAYAW WRKFGSPRFVREQDLGYNARITIEADCYLHGYSQKYFEDIASILWKDLFRQAISGENASMAERIQSAP SVSMHIRRGDYLTSAKARSTHGAPDLGYYGRALGEIRARSGSDPVYLFSDDPDWVRNNMRMDANLVTVA INDGKTAFFEDLRMLSLCDHNIIVNSTFSWWGAWLNPDLKIVVAPKRWFADPKLSNPDIPTPGWLRLGD
	WP_00203 0616.1	487957217	glycosyl transferase, family 11 [Vibrio cholerae]	24.04			MKIISFGGLGNQLFQYAFYLYLKDN5DFGNIFLDFSFYEQNKRDVIRNFYGVDSLDIIKQSSYVRGKFLILKL INKRFFNNLLEFVDKENGIDETILSTNKVFFDGYWQSYRVKDYKSNIKELSFYDFKGNILEVRKKICQSSNV CMHVRRGDYVAEKNTKLHVGCSLQYRDLNKNKNDNSIDHIFISDDIDWVKNNISFDIPVTVDVFGQ SVPDYAEMILLFSCGKHKVIANSTFSWWGAWLSDRNGVIVSPKKWFAKEEKNYDEIFIEGSLRL

PCT/US2015/030823

WO 2015/175801							MIIVFRGGMGNQMFQYAFRLYLEMKGATLKADLSEFKCMKTHAGYELDKAFDLHPAEASYKEIRAVADYI PVMHRFPFSRKFVILYKKEKTRVEAEGPKSHISEEKYDFMSEDERLHLASSEDLMDGFWIKPDMYDDE VLKCTFTSKTLDKYGKGTIEDEHSCSVHVRCDYTGTLGDLKEYEAAEKILSEFADADVKFYVFSDDREKA LLSPFMKKMVFCDTPASHAYDDMYLMSRCRHHIANSTFSFWGARLSADKSGITICPKYEDKNNTANRLVHE GWQML
203						24.03	MIIMKFMGGLGNQYQYALGRKSELHNSFLASDIHIYKNDPREFVLDFENIKVHLPWKVILKNSDYALKF DKVFHTEFYHELVEKALESKDIPRKNLNLRLGWSGNRKYEDYDKISDEITLKEKFKTKDFNTVKNKNSDS VGIHRRGDYKVAHFNKPYGLLPPSYSAADVDFIGNRIEKSNIFFISDDTDVWKENLPLFKDSFFVSDIIGSVD YIEFELLKNCKHQIANDSTFSWWAARLNSNPAAKIVIKPRWFADDRQQAAYEEDSYIKEAIAKL
204						24	MKIVNIGGLGNQMFVYAMYLALKEAHPREEIILCRSYKGYPLHNGYELERIFGVEAPEAALSQARVAYPF FNYKSWQLMRHFLPLRKSMASGTTQIPFDYSEVTRNDNVYDGYWQNEKNFLSIRDKVIKAFTPFEFRDEK NKALSDKLKSVKTASCHIRRGDYLKDPYGVCSNDYITRAITELNQSVPNDMYCIFSDDIGWCKENFKFLIGDK EVFVDWNKQGESFYDMQLMSLCHYNIANSFSWWGAWLNNDDKVVVAPERWMNKLTLENDPICDN WKRIKVE
205						23.99	MRLIKMTGGLGNQMFYAFYLMKKLFPHTKIDLSMMHYHVVHGYEMNRVFAALPHTFEFCINRTLKLME FLCKVYERKQKNGSMEEAFKYYAWPLIYKFGYQSERFFADIEDDVRTKFCNMELINSRREMMKIIDAD EHAVSIHRRGDYLLPKFWANAGCVCQLPYKNAITELEKHESTPSFYVFSDDIEVWVQNLPLQNAHYIDWN QGNDSWQDMMLMMLSHCRNHIICNSTFSWWGAWLNPKNKTIVVPSRWFMEKETPYIYVPSWIKVPIN
206						23.99	MRLIKVTGGLGNQMFYAFYLMKKYYPKVRIDLSMMHYKVHYGYEMHRVFKLPHTFEFCINQPLKKIIEFLF FKKIYERKQAPNSLRAFEKKYFWPLLYKFGYQSERFFADIKDEVREAFTEFDRSKANSRSLDMLDLDKDNNAV SLHRRGDYLPQKHWAATTSVCQLPYQNAIAEMSKRVTSPSYIFSDDIWVRENPLQNAVVIDWNTGE DSWQDMMLMMLSHCKHHIICNSTFSWWGAWLNPISIDKTIVVPSRWFQYSETPDYIYPTGWIKVPVD
PCT/US2015/030823						23.97	MIIVRWGGLGNQLFYSGQYLEIETDKKVFYDVASFGTSDQLRKLKELCSFIDPIPLYNAYFTRYTGKVRNLF KALFQWSNTYLSMESMFDICLLEKARGKIFLQGYWQEEKYATYFPMQKVLSEWKNPNVLEIEENIRSAKISVS LHVRRGDYFSPKNINVYGVCTEKYEQADRANSEIEEDKQFVFSDDILWVKNHIVSLPESTVFPNHEISQFA YIYLSLCKVNIISNSTFSWWGAYLNLQHNQNLVIAPSRWTFTSNKTLALDSWTIKI

Desulfomicrobium baculatum; Desulfomicrobium baculatum DSM 4028	YP_003159045.1	256830317	glycosyl transferase family protein [Desulfomicrobium baculatum DSM 4028]	23.76	>gi 256830317 ref YP_003159045.1 glycosyl transferase family protein [Desulfomicrobium baculatum DSM 4028]MAKIVTRIMGGIGNQLFCYAAARRLALVNHAELVIDDVTGFSRDRVYRRRYMLDHFNISARKATNYE RMEPEYRRGLAKYISKLLPFEREYIEQERIEFDFLEVRTYNNIYIDGLWQSENKFDVIEDIRDDLIKIPPT DLENINIAKKIKNIQNTIAMHVRWFDLPGLNGNNVSTYYHRAIAMMEQINAPHYFLFSDNLEAVHSKLD LPEGRVTFVSNNDGDDNAYADLWLMSCCKHITANSTFSWWGAWLGEGRDSDVVLVPRFSPDGGVTSWC FTGLIPRWEQVSSIR	213
Prevotella pleuritidis; Prevotella pleuritidis F0068	WP_021584236.1	545304945	galactoside 2-alpha-L-fucosyltransferase [Prevotella pleuritidis]	23.76	MDIVLIFNGLGNQMSQAFYLAQRQRNNHTVYCVFGPRTQYSLDKLFDIPYRHNNAVLLVLRALDKAHFSN HRWLRLLRPTLQLLGVKMIVEPLSRDFDMRHFTHQKGVFVYRGWVHSELNFTAVADAVKRRFRFPEIQDA AVLAVIDRIKSCQSVSLHLRRGDYGLSEFQGVCTEAYYEHAIAFYEQIESPEYFVFSDDPTYAREQFGADPNF HIIDLNHGEDAWCDLLMMTQCRYNIA NSTFSWWGAWLNDNP SKIVVHPRYHLNGVETRDFYPRNWICIE	214
Bacteroides sp. 1_1_14	WP_008763191.1	496038684	glycosyl transferase, family 11 [Bacteroides sp. 1_1_14]	23.75	MKVIWFNGNLGNQVFCYKKEFLHNKYPNETIKYYSNSRSPKICVEQYFRLSLPDRIDSEKVRVFEFLGKFFR RIPLKFVPKWYCTRKSLNYESYFEHYLQDKSFEKEDSSWLKAKKPDNFSEKYLIFENLICNTNSVAVHIRRGD YIKPGSDYEDLSATDYEQAIKATEVYLDSEFFESDDLEFVKNNFKGDNIIYVDCNRGADSYLDILLMSQAK INIIANSTFSYWGAYMNHKKVMYSDLWFRNESGRQMIPNIMLDSWICITKRK	215
Agromyces subbeticus	WP_022893737.1	551273588	protein [Agromyces subbeticus]	23.65	MVGRVGIARRQAADVSTCDGEGLVAWRIRTGEIVLGLQGGIGNQLFEWAFAMALRSIGRRVLFDAVRCRG DRPLMIGPLIPASDWLAAPVGLAGATKAGLLSDRSWVPLVRQRRSGYDPSVLERLGGTSYLLGTFSQARY FDGVEHEVRAAVRALLEGMLTPSGRRFADELADPHRVAVHVRRGDYVSDPNAAVRHGVLGAGYYDQAL EHAAALGHVRRVWFSDDDLWVREHLARDDDLCPADATRDHGGEGEIALIASCATRIIANSSFSWWGGWLG APSSPAHPVIA PSTWFA DGHSDAAELVPRDWVRL	216
Prevotella salivae; Prevotella salivae DSM 15606	WP_007133870.1	494220705	alpha-1,2-fucosyltransferase [Prevotella salivae]	23.59	MIATTLFEGGLGNQMFYATAKALSLHYRTPMAFNLRQGFQDYKYQRHLELNHFQCQLPTAKWITFNKYGE LNIKIRISRRIGRNLLCPHYQFIKEKEPFHYEKRLFEFTNKNI FLEGYVWQSPRYFENYSDEIRDFQLKSILPHTITD ELQMLKGTGKPLVMLGIIRRYQEVKDKKDSYPLCNKDYAKAISHVQEQLPAPLVVFTQEQAWAMNLP TNANLYFVKEKDNAWATIADMYLMTQCQHAISNSTFYWWGAWLQHPNIENHIVVAPNFINRDCVCVDN WIID	PCT/US2015/030823

Carnobacterium sp. WN1359	YP_008718 687.1	554649641	glycosyl transferase [Carnobacteriu m sp. WN1359]	23.57		MIFVDLSEGLNQMFQAYSRYLQELYGGLTYLNTSSFKRNKSTRSYSLNFFLYENVKLP5KFRRVIYFY5K TIRMFIKVIRMINPYSDKYF5MIPYGFYVSSQVFKYLTVP7TKRHNFVMTGTWQTNKYFQ5INDKIKDELKV KTEPNELNKKLITEINSQ5CVHIRLGDYTNPEFDYLFHVCTSDYYLKGMDYIVSKVEPNFYIFS5SSDIEWI KNYNFKYKVYIDLNNPDFEDFRLMNYNCKHFISN5TFSWWAQFLSNNDKKIIVAPSKWQK5SNEAKDIY LDHWKLEIE	219
						MLIQIAGGLGNQMYYAMRKLKAGADRNIKLDTKWFDKQ5GLAKRKLLEYFTGLPLPVCSESERA RFTDRSVARKVVEKLVPGMG5RFT5SCMYHPEIFELKDKYIEGFACQKYDDIMGELQELFVPTHPDEEINI KNMINLMNEMEMVPSVSHIRRGDYLDPENAA5LFGNIATDAYYDSAMEYFKAIDPDTHTFYIFTNDPEYAREK YADPGRYTIVDHN5TKYSLDIDLQMLSHCRGNICAN5TFSFWGARLNRRKDKIPVRTLVMRN5NQPVTPELMH EYWP5WVLDKDGKVR	
Clostridium sp. KLE 1755	WP_02276 2290.1	551018062	glycosyl transferase [Butyrivibrio sp. AD3002]	23.55		MIVIRVMGGLGNQMYYALYKFKALGKETRLDTSWFDNASMQENVLARR5LELRFDDNLTYEACTPQER EALLGKEGFNKLKRLFP5KNKHFE5SEM5FHP5EIFKLDNVYLEGHWACEYYH5DIMPL5LQ5KIIFPKTDNIQN NMLKNK5N5SVSHIRRGDYLDPENAA5MFGGICTDSYYK5AEGYIRN5RTN5PH5YLF5DDPAYL5REHYKG EETVVDWNH5GAD5FYDMELM5SCCKHN5VCAN5TFSFWGARLN5RTEKK5VIRPAKH5KNSQ5QAEP5ERMHEL WENWV5IDE5GRIV	220
						MKFFVGGGLGNQLFQY5YYRYLKKKYP5ERILGNP5DSLKAHNGIEIDKWFDIELPPT5YLYNKL5GILLYRVNRF LYNHGYRLLFCNRVYP5Q5MKHFFQW5GDWQDY5IIKQINIF5FRSEL5PIGKENMEFLKKMETCNSISVHIRRG DYLKTDLHIYGGICT5KY5REAIF5MEQEVEEP5FFFF5DDCLYVET5FADIRNKI5ISHNRDD5R5F5DMYLM5AH AKNMILAN5T5F5C5WAA5YLN5RTAKI5IIPDRWVNT5DFS5KLEALPNEWIKIRV	
Paraprevotella xylaniphila; Paraprevotella xylaniphila YIT 11841	WP_00862 6629.1	495902050	glycosyl transferase [Paraprevotella xylaniphila]	23.47		MRLIKMTGGLGNQMFIYAMYLKMR5VFPDTRIDLSDMVHYRVHYGYEMNKVFNLP5RTEFRINR5LKKIIEFL LFK5TILERKQ5GL5P5YIRKYHWPW5YFKGYQ5EEYFAGVEKEV5EAFV5DVRVNRK5SLCAMQEIMADP5D AV5IHVRRGDY5LQ5GKH5W5K5G5CICQ5RSY5L5NAL5ELEK5RIV5PH5YV5F5EDL5DWVRQY5PLENAV5FIDW5NKG ED5WQDMMLM5SHCRH5IIC5NST5SWW5GAWLN5P5SPDKI5V5APERWTQT5TNSADV5VP5ESWLK5V5IG	221
						MTDRALIAIVK5GLGNQLFIYAAAR5MALRTGRQLYLD5AVRGYLD5DYG5R5FR5LNR5FPIE5AELMPEQWRVA STL5RHP5RAK5LV5RAL5NK5YLPE5AWR5FY5VAER5GDT5RP5GAL5WNH5GRNV5KRV5TL5MGY5WQDE5AYFL5DYAEL5LR5REL GPP5MP5DAPE5VR5ARGER5FAG5TES5FL5HV5RRCRY5PL5LDAGY5QK5AV5DL5ACAEL5NKP5VFM5IFG5DDI5EW5V5VNNI DFRG5AGY5ER5QDY5DES5DEL5ADL5WL5MTR5CR5HAI5ANS5FS5W5WAAW5L5GGA5AG5SR5HR5VW5AP5Q5SGL5ALK5CAK SWE5AV5DAQ5PE	
Thauera sp. 28	WP_00293 0798.1	489020296	glycosyl transferase family protein [Thauera sp. 28]	23.47		MTDRALIAIVK5GLGNQLFIYAAAR5MALRTGRQLYLD5AVRGYLD5DYG5R5FR5LNR5FPIE5AELMPEQWRVA STL5RHP5RAK5LV5RAL5NK5YLPE5AWR5FY5VAER5GDT5RP5GAL5WNH5GRNV5KRV5TL5MGY5WQDE5AYFL5DYAEL5LR5REL GPP5MP5DAPE5VR5ARGER5FAG5TES5FL5HV5RRCRY5PL5LDAGY5QK5AV5DL5ACAEL5NKP5VFM5IFG5DDI5EW5V5VNNI DFRG5AGY5ER5QDY5DES5DEL5ADL5WL5MTR5CR5HAI5ANS5FS5W5WAAW5L5GGA5AG5SR5HR5VW5AP5Q5SGL5ALK5CAK SWE5AV5DAQ5PE	222
PCT/US2015/030823							3

Subdoligranulum variabile; Subdoligranulum variabile DSM 15176	WP_007048308.1	494107522	alpha-1,2-fucosyltransferase [Subdoligranulum variabile]	23.44	MIYAEAGGLGNQMFIYAFARALGLRCGEAVTLLDRQDWRDGAHTACALEGLNLVPEVKILAEFGFAKRHLPRQNTAKALMIKYEQRQGLMARDWHDWERRCAPVLNLLGLHATDGYTPVRRGPARDFLAWGYFQSEAYFADFAPTIRAEIRAKQAPAGVWAEKIRAAACPVALLHRRGDYCRPENEILQVCSPAYARAAAAAAYPEATLFFVSDIDWAKEHLDTAGLPVWMPRGDAVGLNLMLALCRGFILSNSTYSWVAQYLAGGRTVWAPDRWFAHTKQTALYQPGWHLIETR	WO 2015/175801
Firmicutes bacterium CAG:24	WP_021916223.1	547127527	protein [Firmicutes bacterium CAG:24]	23.4	MIIVVMGGLGNMQQYALYRKLESIGKDARLDVSWFLDKERQTKVLASRKLELSWFENLPAKYCTQEEKQAILGKNLIGLKKLLGGSNRHFTESDMYHPEIFLEDAYLSGFWACEAYADILPMLRSQHFFPDPKEKGEWDLEAAAKNKETMERMKQETSVSIIHRRGDYLDAKNAEMFGGICTDAYEAAISYIKEQTPDAHFFVFSDDSYVKNAYPGKEFTVVDWNTGKNSLFDMLMSSCNHNICANSTFSFWGARLNPSPDKVMIRPSKHKNSQNIIVPEEMKRLWDGWVLDIGKGRII	225
Prevotella sp. CAG:474	WP_022310139.1	547906803	glycosyl transferase family 11 [Prevotella sp. CAG:474]	23.39	MIITKNGGLGNQLFEYACARNILQLKYNDVLYLDIEGFKSRPHYSLEKFKLSSDVRMLPEKDSKSLILLQAISKLNRLAFKLGPLEGTIYWKSSNYRPLKIKNTRGKLYLYGYWQSYVEYFENEAIKQELNVKTEIPIECSELLKEINKPHSICVHVRRGDYVSCGFLHCDEAYYNRGINHIFDKHPDSNVVVFSDIKVWVKANMNFDPVAYVEVDVPDYETLRIMYMCKHFVMSNSSFSWVAWSYLSDNKEKIVAPSYWLPANKDNKSMYLDNWTIL	226
Roseburia intestinalis; Roseburia intestinalis L1-82	WP_006855899.1	493910390	glycosyl transferase family 11 [Roseburia intestinalis]	23.38	intestinalis]MRGNRGMIAVKIGDGMGNQLFNYACGYAQARRDGDLSLVDISECDNSTLRDFELDKFHLKYDKKESPNRNLGQKIYKNLRRALKYHVIKEREVYHNRDHRDYDNDIDPRVYKKGLRNKLYGYWQHLYAYFEDYLDEITAMMTPAYEQSETVKKLQEEFKTPTCAVHVRGGDIMGPAGAYFKHAMIERMIEQEKPGVRYIVFTNDMERAEALAPVLESQKDKDAVGQAENRLEFVSEMGFSDVDEFFLMAACQNQLSNSTFSTWAAAYLNQNPDKTVMIPDDLLSERMRQKNWIIK	227
Bacteroides ovatus; Bacteroides ovatus ATCC 8483	WP_004296622.1	490424433	protein [Bacteroides ovatus]	23.29	MKIVLFTPLGLGNQMFQYLFYLYLRDYNPNQNIYGYNNRNLNKHNGLEVDKVFDIQLPPTHVISDASAFFIRALGGLGKYFYGKDLSPWKVYFDGYWQNKKEYFQNNVNDKMRREGFLNKKNDLILSLRNTNSVSVHVRRGDYCDSCKDLFLQSCPTQYYEASISVMKEKQKPVFFVFSDDIPWVKVNLNIPNAYYIDWNKKENSYLDMYLMSLCTASIIANSTFSFWGAMLGNKKELVIKPKKWIGDEIPIFPWSWLSL	228
Butyrivibrio sp. AE3009	WP_022779599.1	551035785	glycosyl transferase [Butyrivibrio sp. AE3009]	23.25	MLIIQIAGGLGNMQQYALYRKLLKYHPDGVRLDLSWFDSEVQKNMMLAKREFELAFKGLPYECKPEERAAFLDRNAAQKLSGKVLKGLRDNANPNVFEESRMFHPHPEIFELDNKYIIGYFACQKYDDIMGDLNLFEPFELDPELEKKNLISLISMEKENSVSVHRRGDYLDPENFKILGNIAATDEYESAMKYFEDRYEKVHFYFTSDHEYAREHFADESKYTIVDWNTGKDSLQDVRMLNHNHCLGNICANSTFSFWGARLNQRQDKVMIRYTKMIRNNQPVDPDTMHDYWKGWILIDETGREV	PCT/US2015/030823

WO 2015/175801	2						MTKNEKKLVKFQGGGLGNQLYEAFCWLRQQYSDYEVLDLSYYKIRSAHGELGIWNIFFPINIEVASNWDI IKYSDQIPIMYGGKADRLNSVRTNVNDRFSSKRKHSYYTEISNTDSEVINALLNNGIRYFDGWQNIIDYFKG NIEDLRNKLKFESEKCDKITDEMLRDNVSLHVRRGDYVGESEKEVGLSYYKAVEYVLDLDRVDQAKFFISD DKYAEATAFEWIDNKTWAGYDNELAHVDMMLMSRMKNINIANSTFSLWAAAYLNDSMNPLIVYVDPVSLD KKTFSWNGIK	231
							MDSQFLKHILSGGFGNQLFOYFFGEYLKEKYNCSIFFSEPALDINQLQIHRFFPALRISHNTELRPHYSTFQ QLAYRCMRKLLLPFLNRKVKIENGSNYQNSFNQDYCFDGYWQSYRSLAFTPSLQFEDQLINDISADYIN AIEQSEAVFLHIRRGDYLNRKENQKVFAECLNMFENANRIKIDIKNVHFFVFNDSIQWVSKHLKLNDEVT IQNEGSCDLKDFYLMTRCKHAISNSTFSWAAAYLINNSDKVIAPKHWNIDISMINNATKDLIPTWIRL	232
							MIITRLHGRGNQMFQYAAGRALADRAGVPLALDSRGAILRGEGLTRVFDLELADPVHLPLKQTNPLRYA IWRGIGQKVGAKPYFRERGLGYNPAFEDWGDNSYLHGYWQSQYFQNSAERISDFTTFAFSNQNAE MAARIAESTAISLHVRRGDYLTFAAHVLCDOQAYDAALAKVLDGLQGDPIVYVFSDDPQWAKDNLSLPCCK VVVDNFGPETDFEDMRMLMSLCQHNIIGNSSFSWAAAWLNLQTPGRRVAGPAKWFGDPKLSNPDIFFHDW LRISV	233
							MGNQLYAYATAKAMAVALNKKLVDPPILEAPQRHYDLGLFNIQDEDFGSPFVQWLVWVAVSVRLGKFF KTIMPFAWSYQMIRKDEEGFDESLQKQSRNIVIEGYWQSKYFESIRPTLLKELSKDKPNAINQKYLDEIESV NAVAVHIRRGDYVANPANAVALHGLCDMDYKKAIAIKDVENPYFFITDDPDWAEDNFKESEHQIKHNI GKODHEDFRLLTNCKYFIANSFSWGAWLSDYKNKIVISPNKWFNVDVPIITERIPESWIRV	234
							MITVRIDGGFGNQMFOYAFFLHLKTTITDNKISVDLNCYNPHGSGDIFTRFLAPEQAAPSEIKRHFHRSIYHL LRPLDSAGITNPPYREEDIDDLNSVLNKKRVYLRGYWQDKRYPSVKDQLIDCFDLGKMDMTGASAEINNV LEQASEESRSVGVHLRGDYGDPVYSGICTPEYEAFAKHVSEKIKDPVFHIFTNDISMIEKCGLSGKYDLKIT DINDEAHGWADLKLMSACRHHIISNSSFSSWAAFLGEATTEASADVINVIPEYMRQGVSAETLRCPWTT VTSGRVYPS	235
							MKIVICKGGLGNQLFEYCRYSLRHRHDNRGVLYHYDRRRRTKQHGCVWLDKAFHITLPLNEPLRVKLLVMVLK TLRLHLKRLRYREEDPRAVLIDDYSQHKQYITNAAEILNFRPFQLDYAEIQTTPFAVSVHVRRGDYLLANK SNFGVCSVHYLYLSAAVAVRERHPESRFFVSDMEWAKENLNLPCVFEVEHAQAQPDHADLYLMSLCKGH IIANSTFSFWGAYLSKGSSAIAIYPKQWFAEPTWNVDPDIFPAHWMAL	236

91

WO 2015/175801	241						MRIGILYCTGKYTVFWNHFTSCQEHFLREHEKHYYIFTDGEIAHLNCRVHRIEQHGLGWPDSLTKRPHM FERIADTLRQNSDFVFNANMVFRLDVGKEFLPTREQALVHRHPLRRPAWLLPYERRPESTAYIPYSGG SIYVCGGVNGGYTPYLDFAVAMLRNIDIDVERGIIARWHDESHINRFVIGRHYKIGHPGYVYPDRRLPFR IIRVDKASVGGHTFLRGQTPAPEEQSKTVAKLRSQKRCMPRAAQDEPIILARMGGLGNQMFIYAA ARVLAERQGAQLHLDTGKLSGDSIRQYDLPAFSDAPLWHIPCGCDRIVQAWFALRHVAAGCGMPKPTMQ VLRSGFHLDRFFSIRHSAYLIGYWQSPHYWRGHEDRVRSFDTLRFERPHLRLEAAVSPQNTISVHLRRG DFRAPKNSDKHLIDGSGYERARKLLEMTPTQSHFYFSDPEEAQRFLFAHWENTSFQPRRSQDEEDLLMSRC SASIANSSFSWWGAWLGRPKQHVIAPRMWFTRDVLMHTYTLDLFPEKWILL
	242						MILIHVMGGLGNQLYQYALYEKMKSLGKKVKLDTYAYNDAAGEDKEWRSLELDRFFPAIEYDKATSEDRTKLL DNSGLLTAKIRRKLLGRKDKTIRESEYMP EIFHMDDDVLYLGFWNCERYEYEDIPLLQDKLQFPISNNNPRNQ CIEQMOKENAVSHIRRTDYLTADGARYMGICTEDVYKAMAYIEERVSNPVYIFSDDDVEYAKQHYHQD NMHVVDWNSKADSIYDMQLMSCKCHNICANSTFSMWAARLNQNKEMIRPLHHDNYETTTATQVKQN WKNWILLDQNGQVCE
	243						MTMNIIRMSGGLSQMFQYALYLKLSMGKEVKFDDINEYRGEKARPIMLAVFIEYPRATWDEITSFTDG SMDLLRLRRKIFGRKAIEYEEQGFYDPNVLNFDSDMYLRGNFQSEKYEQDKEEVKRLYRFTLEDMLRPERLY KATKACLDGIESSESGLHMYRSDSRVDGELYDGICTGNYKGAVERFIQDKVPDAKFYIFSNPKVWVRGWV DLIQSQIQEGMSPSQVKEMEKRFVMVEANTEYTGyLDMMLMSCKKHNIISNSSFSWWSAWMINDHPEKV VWAPDRWSSDKEGNEIYTTGMTLVNKEGRVNTIHENSTVK
	244						MILSYITRGLGNQLFEYAYARSLLLKRGNELIINFSLVRAAGKEIEGDDNLRYFNVYSYTELDKDIVLSKGD LQLFYILFKDQKLFRIIKKEKWFSEFFRRFGIIFQDYLDNISNLIPIRTKNVFCYKVENPKYFDDIRSLILKEFTPRI PPLKNNDDQLYSVIESTNSVCISIRRGDFLCKDKFDRFLVCDKEYLEAMEEAKKRISNSTFIFFSDDIEWVRENIH SDVPCYYESGKDPVWEKRLMYSCKHFIISNSTFSWAAQYLSRNEEKVVIAPDRWSNVNVPGEKSFLLNSFIKI PIGILP
	PCT/US2015/030823						MIYVEINGRLGNMFEIAAASLTDDEVTLWCKGDWQLNCIKMYSDTLFNKNYPVKSLPNNIRIYEETFFHPI PYKENQDILLIKGYFYQYKYLDRKVLKLYPCPMPVKDLIEKRGDILSQYTVVSNVRRGDYLNLPHRHPFVGK KFLERAMLWFGDKVHYIISDDIEWCKAHFKQFDNVHYLTNSVPLLDLYIQTACHHNIISNSSFSWWGAYLN NHPQKVIAPHRWFGMSTNINTQDLLPPEWMIQCQVEPKVFLKALPLHAKYLLKRVLK

Prevotella sp. oral taxon 299 str. F0039; Prevotella sp. oral taxon 299	YP_008444 280.1	532354444	protein HMPREF0669_0 0176 [Prevotella sp. oral taxon 299 str. F0039]	22.9	MDSQLLKHILSGGFGNQLFQYFFGEYLKEKYNCSFFSEPALDINQLQIHRFFPTLRISHNTELRRFHYAFTQ QLAYRCMRKLLLEPFLNRKVIENGSNYQNSFNDTYCFDGYWQSYRLSAFTPSLQFEDQLINDISADYIN AIEQSEAVFLHIRRGDYLKNGKQVFAECPLNYFENAVNKIKGNKTYHFFFSNDIEWVKCHLKLNNNEVTF IQNEGSSCDLKDFYLMTRCKHAIISNSTFSWAAAYLINNNDKKVIAPKRWYNDSLMMNNATKDLIPTWIRL	WO 2015/175801 247
Paraprevotella xylaniphila; Paraprevotella xylaniphila YIT 11841	WP_00862 8783.1	495904204	alpha-1,2- fucose transferase [Paraprevotella xylaniphila]	22.87	MKIVCLKGLGNQMFEYCRFRDLMDSGNGKVYLFYDRRRRLKQHDGLRLSDCFEELPSCPWGIRLVVWVGL KICRAIGVLRKLYDDEKDPADVLIDDYSQHRRIPIPNARRYFSFRQFLAELQSGVQVMIRAVDYPVSVHVRRGDY LHPNSSFVLCGVDFYRQAIAYVRKRPDARFFFSDDMEWVRNLMWEDAVVVEHTELMPPDYMDLYLM TLCRGHIISNSTFSFWGAYLAVDGNMGKMYPRRWFRDPTWTIPFSEEWVGL	247
Dethiosulfovibrio peptidovorans; Dethiosulfovibrio peptidovorans DSM 11002	WP_00565 8864.1	491897177	glycosyl transferase family 11 [Dethiosulfovibri o peptidovorans]	22.84	MFQYAFGRALALDLGLDLKLDISNFGSDSRPFLSGIYSLTKNIPFGCYLSTSTRLKVKMTKKLRRWGVWGMGMD KNMPPGVLEPFPVPLVSLDEVLSKLSHLFVDGYWQSEKYFSRYSDVIRSDFRVIESSAFLAWKKRMLSEPG GSISVHVRRGDYVTDSSANRVHGVLPYLYLRAKELINTISDGLVFYVTDTPVWARNLCLGDKTIYVSGEDL KDYELALMSCDDHHVAVANSSFSWGAWLGDQDTSTVTIAPGRWFRKMDSSFVIPDNWIKIWT	248
Lachnospiraceae bacterium 10-1	WP_01622 9292.1	510896192	protein [Lachnospiraceae bacterium 10- 1]	22.83	MIIIVVMGGLGNQLQYALYRKFVRMGKEARLDSWFLDKEKRGVLAERELELDYFDRLIYETCTPEEQLI GSEGVAGKLKRKFLPGRIRWFHESKIYHPPELLQMENMYLSGVFACEKYADILYDLREKIQFPVNDHPKNIM AQEMQERESVHLRRGDYLDKNTAMFGNICTDAYCYCAIEYMKTLCSKPHFYIFSDDIPIYVVRQRTGEEYT VVDINHGRDSFEDMWLMSRCRHNICANSTFSFWGARLNSDNKIMIRPTIHKNSQVFKVEEMLWPG WKFSIPDGGIK	249
Treponema maltophilum; Treponema maltophilum ATCC 51939	WP_01652 5279.1	513872223	protein [Treponema maltophilum]	22.82	MFCAAFVEALKHAGQKVFVDTSLYKNGTVRSIGDFCHNGLETEHLFGIKFDEADKADVHRLSTSAEGLNRIIR RKYFTKTHYIDTVFRYTPVLSKSDSDRYLEGFWQTEKYFLPIESDITLFRFRQPLSEKSAVQSAQOQEPAS LSASIHVRRGDFLHTKLTNVCCTETYNNAIEYAAKAYAVAFYVFSDDIQWCREHLNFFGARSVIDWNIGAD SWQDMVLMSCMCRNIANSFSSFWAAWLNAASDKIVLAPAIWNRRLQLEYADRYGYDYSDVIPETWIRIP I	250
Bacteroides massiliensis; Bacteroides massiliensis dnLKV3	WP_01627 6676.1	511022363	protein [Bacteroides massiliensis]	22.79	MKLVSTAGLGNQLFQYCFYRYLLNKEPNEKIYGYNNKKWLKHHGIIIEHFFDVKLPRSTRWVNLGYQYLRYY KCFSCGVSKDDDFEMNRTMFVGYWQDQCFSGINISYKKNLVISEKNTWLLGELKCNVSAIHFRRGDYMLP QFKKIFGEVCTVYKYLKIRKVEEKISEPVFVFSDDIDWVKQNTFNKYVYVDWNGQNSFWDMYLMSSQC SANIIANSTFSFWGAYLNNKNNPFVYIPQKVVWRTNLKQPNIFPKTWMAAL	PCT/US2015/030823

WO 2015/175801					253					254					255					PCT/US2015/030823																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																											
Enterococcus faecium;Enterococcus faecium DO;Enterococcus faecium EnGen0035	YP_006376 560.1	389869137	family 11 glycosyltransferase [Enterococcus faecium DO]	22.71	MIVLTGGGLGNQMFQGYARYIQIHREKFIYINDSEVKEADRFNSLGNLNTVNIKVLPRISKPLNETERLV RKIMVRLFGVAGFNESAIFOSLNKFGIYHPVSVYKFYESLKTGPKIKIEGGFQSWKYLETCPEIKQELRVKYEPM GENLRLLNISQSESVCHIRRGDYLSPKYKHLNVCDYQYFESMNYISKLNPTFFIESNTSDDLWDWIKEN YSLPGKIVYVKNDNPDYEELRLMYSCKHFISNSTFSWWAQYLSNNSGVIAPEIWNRLNHDGIADLYMPNWN ITMKVNR					MDVVVIFNGLGNQMSQAYYLA																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																					</

Bacteroides coprophilus; Bacteroides coprophilus DSM 18228 = JCM 13818	WP_00814 4634.1	495419937	protein [Bacteroides coprophilus]	22.56		MGFINMACGLANRMFQSYSLFLKKQGYKVTVDYFYSRAKLAHEKVAWNSIFPYAEIKQASRLKVFELWGGG SDLSKVRRRYFSPSTNVRTTTGAFDASLPANTARNEYIIGVFLNASIVEAVDDEIKKCFTELPTDEMNLRLKK EIEECESVAIHVRKGDYQSRWYQNTCSMEYRKAILQMKELQHSKFYFTDNVDWVKENFQEIYTLVE GNPADGYGSHFDMQLMSLCKHNIISNSTYSWWSAFLNRNPEKVVVIAPEIWFNPDSDEFRSDRALCKGWI VL	WO 2015/175801 258
						MKIVCLKGGLGNQMFEYCRFRDLMESGHDEVLYFDHRRRLKQHNGIRLSDCFEELPSCPWGKLVVWGLK ICRAVGVLKRLYDDEKPEAVLIDDYSQHRRFIPNARRYFFRQFLAELQSGFVQVMIRAVDYPVSVHVRRGDYL HPSNSSFLGCVDFYQQAIAVVRKRDPDARFFFSDDMEWVRNLMWEDAVVVEHTELLPDDYVDLYLMTL CRGHIISNSTFSFWGAYLAVDNGMGMKIYPRRWRFRDPTWTSPPIFSEEWVGL	
Bacteroidetes; Capnocytophaga sp. oral taxon 329 str. F0087; Paraprevotella clara YIT 11840	WP_00861 9736.1	495895157	alpha-1,2-fucosyltransferase [Bacteroidetes]	22.53		MLIQIAGGLGNMQQYAVYTKLREMGKDKVLDLSWFDQVQKQNM LAPREFELPIFGGTDYEECSAYERD ALLKQGAFAAIAIGKVLKGLRDEANPKVFEKEMYPHPEFELEDKVIKGYFACQKYGDIMDKLOEKFIPE HSDPDLHARNMALVERMEREPSVSVHRRGDYLDPSNVEILGNIAETEQQYQAMDYFTVKEPDTHEYFTSD HEYAREKFSDESKYITVDWNNKNSVQDMLMLSHCKGNICANSTFSFWGARLNKRDPKTVIRTYKMRNN QVNPQIMHDYWKGLMDEKGSII	259
						MKILVFTGGLGNQMFAFYLYLRLFPQERFYGLYKKLSEHYGLEIDKWKVSLPRQPWWVLPVTGLFYL YKQCVNSKWLNLNQEICKNPRAIIVFFPFKTKKYIPDDNIWLEWKVDESGLSEKNRLLSEIRSSDCCCFVHVR RGDYLSPTEKSLFEGCCTLSYQRAKSMKEISPVKVFCSDDIQWVKQNLGNRAVFDWNSGTDSPLD MYLMSQCRYGIMANSTFSYWGARLGRKKRIYYPQKWWNHGTLGDPDIPNTWVKI	
Blautia hydrogenotrophica DSM 10507; Blautia; Blautia hydrogenotrophica CAG:147	WP_00594 4761.1	492742598	protein [Blautia]	22.44		MEIHVYLTGRNLGNLQFYAFARHLQKEYGGKIICNIYELLEHRSEKAAWVPGKFNEMSNYKLNDSILEDIKLP WFADFSNPIIRIVKKVIPRIYFNLMSKGYLLWQKNYSINIPAIARNNEIIVNGWWQDVRFFHDVEAELSNEIVP TTKPISENEYLYNIAERENSVCSIRGGNYLVPKVKKKLFVCDKEYFYNAIELIKSVRNAIFVFSDDLEWVKS KLEEKPECKFYYESGKDTVEEKLRMMTCKKHFIISNSFSWWAAQYLAKNENKIVIPDAWFTNGDKNGLYI DDWILPTQTKDM	261
						MITVLLNGGLGNQLFQYAAAGRALAEKHDEVLLDLSRLQHPKPGDTPRCFELAPFNKASILAEEGRQPLGSY QACMHRLLKASIPILWGSILKEQCGCFDPIIFRAPSSCILDGFQWQSECVFKQITSLQQLSILKAPSPALRKAS SVLSDATVAVHVRRGDYVTNPAASAFHIGCSQDYQAAVANILTSYDPSQLVFSDDPAWCCQEHLDLQGP RLAADFGLNGSAEELVLSRCAHQIANSFSWWGAWLNPSPHKLVVAPCRWFTDPAITNDLLPETVWRPLP	
Geobacter lovleyi; Geobacter lovleyi SZ	YP_001952 981.1	189425804	glycoside hydrolase family protein [Geobacter lovleyi SZ]	22.44			PCT/US2015/030823 2

Providencia alcalifaciens	AFH02807. 1	383289327	glycosyltransferase [Providencia alcalifaciens]	22.26	MKINGKSSMIKQKKIISHLIGLGNLQFYATSYALAKENNAKIVIDDRFLKFKYKLGGRYLDKLNIGEKISS IDKLLFPLILCKLSQKENFIFKTKKFILEKKTSSFKYLTFSDEKHTKMLGYWQNAIYFQYFSELKEMFVPLDIS QEQLDSLQIHAQQSVALHVRRGDIYSNKALAMHGICSIDYYKNSIQHINAKLEKPEFYIFSNKDLWCEENLT PLFDGNGFHIVENNSQIEDLWLSQCQHHIIANSTFSWWGAWLANSDSQVITPDWFWFNKEIDIPSPVLSHWL KLKK	WO 2015/175801 270
Salmonella enterica	AFW04804. 1	411146173	glycosyltransferase [Salmonella enterica]	22.26	MFSCLSGGLGNQMFOQSAAYILKKNICHQAQLIIDDSYFYCQPKQDTPRIFEINQFNIVFDRVTTDEKRAISKL RKFKKIPLPLFKSNVITEFLFGKSLTDEDFYKVLKKNQFTVKMNACLSLYQDSSLINKYRDLLPLFTINDELLO VCQQLDSYGFIHETNTTSLHRRGDYVYTNPHAAKFGHTLSMNNYSQAMNYYVHKLKQGLFIIFSDDDVQWA AEKFGGRSDCYVNNVNCQFSAIDMYLMSLCNNIIANSTYSWWGAWLNKSEELVIAPRKWFAEDKESLL AVNDWISI	270
Sulfurospirillum deleyianum; Sulfurosp irillum deleyianum DSM 6946	YP_003304 829.1	268680398	protein Sdel_1779 [Sulfurospirillum deleyianum DSM 6946]	22.18	MIHKIMGGSLASQLHKYSVGRALSILKYNTLKLDFWFDNISGSDTIREYHLDKYNVVAKIATEQEIKQFKPNKY LLKINNLFQKFTNWKINRYNRYCNEFSISLENFNLLPDNIYVEGEWSDRYSFHSKIELOKELTLKSEYMDSTNHF LAKQSSDFAHDDNASKLHCTCSLEYKALQYISKNLKMLLIFSDDLWLKPNFNFLDNVEFEVEGFQDY EEFLMTLSKHNIANSFGSLFFAWLNHNHKKIISLSEWVFEELNKYIIDNIKKNIULENLE	271
Pseudovibrio sp. FO- BEG1	YP_005080 114.1	374329930	alpha-1,2- fucosyltransferase [Pseudovibrio sp. FO-BEG1]	22.15	MSVASQVRISGAARRRKLPTLIVRIRGGIGNQLFQYALGRKIALETGMKLFDRSEYDQYFNRSYCLNLFKT QGLSATESEMSAVLWPAOSFGQTVKLCRKFPYQRYREDELLOQDSETPVLKQSAYLDGYWQTWEIPFSI MEQLRDEITLKPMVLERLKLQRIKSGPSAALHVRGYDYSQAHNLQNFGLCSAGYKGMDFLTERVPGLT FYFSDSPERAREVVPQENNVYFSDPMQDGDHEDLMVMSSCDHIVTANSTFSWWAAFLNGNEDKHVIA PLKWFKNPNLDLSLIVPPHWQRL	272
Prevotella sp. oral taxon 472 str. F0295; Prevotella sp. oral taxon 472	WP_00923 6633.1	496529942	alpha-1,2- fucosyltransferase [Prevotella sp. oral taxon 472]	22.11	MKIVCKGGLGNQLFEYCRYHGLLRQHNNHGVYLYHYDRRTKQHGGVWLDKAFILTLPTWPVRVKLMVM ALKMLRKLHLKRLYREDDPRAVLIDDYSQHKKQFITNAAEILNFRPQAQLDYVDIEITSEPFASVHVRRGDYLL PANKANFGVCSVHYLLSAAVAVRERHPDARFFVSDDIEWAKMNLNLPNCVFEHAAQPOPDHADLYLMSL CKGHIIANSTFSFWGAYLSMSSAIAIYPKQWFAEPTWNAPDIFLGHWIAL	273
Butyrivibrio fibrisolvens	WP_02275 2732.1	551008155	glycosyl transferase [Butyrivibrio fibrisolvens]	22.08	NILIRVAGGLGNQMQQYAMRYRKLKSLGKEVKLDLSWFDVENQEQQLAPRKCELKYFDGVDFEECTDAERA YFTKRSILTALNKVFPATCKFEETEMFHPEIYSFKDKYLEGFLCNKYDDILPFIQNEIVFPKHSDPKMQKRN EELMERMDGWHTASIHLRRGDYITEPQNEALFGNIATDAYDAIRYVLDKDYQTHFYFNSNDPEYAREHYS DESRYTIVTNGDGNLSDLMELMSHCYRICANSTFSFWGARLNKRSDEKEMIRTFKMRNNOEVTAREMTD YWKDWIILDEKGNRIF	PCT/US2015/030823

Lewinella persica	WP_02057 1066.1	522059857	protein [Lewinella persica]	22.04	MVISRLHSLGNQMFQYAFARRIQLQNVKLRIDLSILLDSRPPDGYIKREYDLDFIKLSPAYHCNPTSLRILYA PGKYRWSQVVRDLARKGYPVYMEKFSVDNTLLDSPDNVYQGYWQSERYFSEVANTIRKDFAFQHSIQP QSESLAREIRKEDSVCLNIRRKDYLASPTHNVTDETYENCICQMRERFSGARFFLESDDLWVWCREFFADFHD VVIVGHDHAGPKFGNYLQLMAQCHHYIIPNSTFAWWAAWLGERTGSVINAPERWFGTDEFDYRDVWPER WLKVPN

OTHER EMBODIMENTS

While the invention has been described in conjunction with the detailed description thereof, the foregoing description is intended to illustrate and not limit the scope of the invention, which is defined by the scope of the appended claims. Other aspects, advantages, and modifications are within the scope of the following claims.

The patent and scientific literature referred to herein establishes the knowledge that is available to those with skill in the art. All United States patents and published or unpublished United States patent applications cited herein are incorporated by reference. All published foreign patents and patent applications cited herein are hereby incorporated by reference. Genbank and NCBI submissions indicated by accession number cited herein are hereby incorporated by reference. All other published references, documents, manuscripts and scientific literature cited herein are hereby incorporated by reference.

While this invention has been particularly shown and described with references to preferred embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the scope of the invention encompassed by the appended claims.

What is claimed is:

The claims defining the invention are as follows:

1. A method for producing a fucosylated oligosaccharide in a bacterium comprising providing bacterium comprising an exogenous lactose-utilizing $\alpha(1,2)$ fucosyltransferase enzyme, wherein the amino acid sequence of said enzyme comprises at least 90% identity to FutW (SEQ ID NO: 17), FutP (SEQ ID NO: 11), FutO (SEQ ID NO: 10), FutQ (SEQ ID NO: 12), FutR (SEQ ID NO: 13), FutS (SEQ ID NO: 14), FutU (SEQ ID NO: 15), FutV (SEQ ID NO: 16), FutX (SEQ ID NO: 18), FutZ (SEQ ID NO: 20) or FutZA (SEQ ID NO: 21), wherein said fucosylated oligosaccharide comprises 2'-fucosyllactose (2'-FL), lactodifucotetraose (LDFT), Lacto-N-fucopentaose I (LNF I), or lacto-N-difucohexaose I (LDFH I).
2. The method of claim 1, wherein said $\alpha(1,2)$ fucosyltransferase enzyme comprises any one selected from *Parabacteroides johnsonii* FutX, *Lachnospiraceae* sp. FutQ, *Prevotella melaninogenica* FutO, *Prevotellaa* sp. FutW, *Bacteroides* sp. FutZA, *Tannerella* sp. FutS, *Bacteroides caccae* FutU, *Salmonella enterica* FutZ, *Methanosphaerula palustries* FutR, *Butyrivibrio* FutV, *Clostridium bolteae* FutP, or a functional variant or fragment thereof.
3. The method of any one of the preceding claims, wherein said $\alpha(1,2)$ fucosyltransferase enzyme comprises any one of amino acid sequences SEQ ID NOs: 10-18, 20, 21, and 292, or a functional variant or fragment thereof.
4. The method of any one of the preceding claims, further comprising retrieving the fucosylated oligosaccharide from said bacterium or from a culture supernatant of said bacterium.
5. The method of any one of the preceding claims, wherein the bacterium further comprises an exogenous lactose-utilizing $\alpha(1,3)$ fucosyltransferase enzyme and/or an exogenous lactose-utilizing $\alpha(1,4)$ fucosyltransferase enzyme.

6. The method of claim 5, wherein the exogenous lactose-utilizing $\alpha(1,3)$ fucosyltransferase enzyme comprises a *Helicobacter pylori* 26695 *futA* gene.
7. The method of claim 5, wherein the exogenous lactose-utilizing $\alpha(1,4)$ fucosyltransferase enzyme comprises a *Helicobacter pylori* UA948 FucTa gene or a *Helicobacter pylori* strain DMS6709 FucT III gene.
8. The method of any one of the preceding claims, wherein said bacterium further comprises a reduced level of β -galactosidase activity, a defective colanic acid synthesis pathway, an inactivated ATP-dependent intracellular protease, an inactivated *lacA*, or a combination thereof.
9. The method of claim 8, wherein said method further comprises culturing said bacterium in the presence of tryptophan and in the absence of thymidine.
10. The method of claim 8, wherein said reduced level of β -galactosidase activity comprises a deleted or inactivated endogenous *lacZ* gene and/or a deleted or inactivated endogenous *lacI* gene of said bacterium.
11. The method of claim 10, wherein said reduced level of β -galactosidase activity further comprises an exogenous *lacZ* gene or variant thereof, wherein said exogenous *lacZ* gene or variant thereof comprises an β -galactosidase activity level less than wild-type bacterium.
12. The method of claim 8, wherein said reduced level of β -galactosidase activity comprises an activity level less than wild-type bacterium.
13. The method of claim 12, wherein said reduced level of β -galactosidase activity comprises less than 6,000 units of β -galactosidase activity.
14. The method of claim 12, wherein said reduced level of β -galactosidase activity comprises less than 1,000 units of β -galactosidase activity.

15. The method of claim 8, wherein said bacterium comprises a *lacIq* gene promoter immediately upstream of a *lacY* gene.
16. The method of claim 8, wherein said defective colanic acid synthesis pathway comprises the inactivation of the *wcaJ* gene of said bacterium is deleted.
17. The method of claim 8, wherein said inactivated ATP-dependent intracellular protease is a null mutation, inactivating mutation, or deletion of an endogenous *lon* gene.
18. The method of claim 17, wherein said inactivating mutation of an endogenous *lon* gene comprises the insertion of a functional *E. coli lacZ⁺* gene.
19. The method of claim 8, wherein said bacterium further comprises a functional lactose permease gene.
20. The method of claim 19, wherein said bacterium comprises *E. coli lacY*.
21. The method of claim 8, wherein said bacterium further comprises an exogenous *E. coli rcsA* or *E. coli rcsB* gene.
22. The method of claim 8, wherein said bacterium further comprises a mutation in the *thyA* gene.
23. The method of claim 8, wherein said bacterium accumulates intracellular lactose in the presence of exogenous lactose.
24. The method of claim 8, wherein said bacterium accumulates intracellular GDP-fucose.
25. The method of any one of the preceding claims, wherein said bacterium is a member of the *Bacillus*, *Pantoea*, *Lactobacillus*, *Lactococcus*, *Streptococcus*, *Propionibacterium*, *Enterococcus*, *Bifidobacterium*, *Sporolactobacillus*, *Micromonospora*, *Micrococcus*, *Rhodococcus*, or *Pseudomonas* genus.

26. The method of any one of claims 1 to 24, wherein said bacterium is selected from the group consisting of *Bacillus licheniformis*, *Bacillus subtilis*, *Bacillus coagulans*, *Bacillus thermophilus*, *Bacillus laterosporus*, *Bacillus megaterium*, *Bacillus mycoides*, *Bacillus pumilus*, *Bacillus lentus*, *Bacillus cereus*, and *Bacillus circulans*, *Erwinia herbicola* (*Pantoea agglomerans*), *Citrobacter freundii*, *Pantoea citrea*, *Pectobacterium carotovorum*, *Xanthomonas campestris*, *Lactobacillus acidophilus*, *Lactobacillus salivarius*, *Lactobacillus plantarum*, *Lactobacillus helveticus*, *Lactobacillus delbrueckii*, *Lactobacillus rhamnosus*, *Lactobacillus bulgaricus*, *Lactobacillus crispatus*, *Lactobacillus gasseri*, *Lactobacillus casei*, *Lactobacillus reuteri*, *Lactobacillus jensenii*, *Lactococcus lactis*, *Streptococcus thermophiles*, *Propionibacterium freudenreichii*, *Enterococcus faecium*, *Enterococcus thermophiles*, *Bifidobacterium longum*, *Bifidobacterium infantis*, *Bifidobacterium bifidum*, *Pseudomonas fluorescens* and *Pseudomonas aeruginosa*.

27. The method of any one of claims 1 to 24, wherein said bacterium is *Escherichia coli*.

28. The method of any one of the preceding claims, wherein said bacterium comprises a nucleic acid construct comprising an isolated nucleic acid encoding said $\alpha(1,2)$ fucosyltransferase enzyme.

29. The method of claim 28, wherein said nucleic acid is operably linked to one or more heterologous control sequences that direct the production of the enzyme in the bacterium.

30. The method of claim 29, wherein said heterologous control sequence comprises a bacterial promoter and operator, a bacterial ribosome binding site, a bacterial transcriptional terminator, or a plasmid selectable marker.

31. A nucleic acid construct comprising an isolated nucleic acid operably linked to one or more heterologous control sequences, wherein the nucleic acid encodes a lactose-utilizing $\alpha(1,2)$ fucosyltransferase enzyme for the production of said enzyme in a host bacteria production strain, wherein the amino acid sequence of said enzyme comprises at least 90% identity to FutW (SEQ ID NO: 17), FutP (SEQ ID NO: 11), FutO (SEQ ID NO: 10), FutQ (SEQ ID NO: 12), FutR (SEQ ID NO: 13), FutS (SEQ ID NO: 14), FutU (SEQ ID NO: 15), FutV (SEQ ID NO: 16), FutX (SEQ ID NO: 18), FutZ (SEQ ID NO: 20) or FutZA (SEQ ID

NO: 21), wherein the lactose-utilizing $\alpha(1,2)$ fucosyltransferase enzyme is capable of producing a fucosylated oligosaccharide comprising 2'-fucosyllactose (2'-FL), lactodifucotetraose (LDFT), Lacto-N-fucopentaose I (LNF I), or lacto-N-difucohexaose I (LDFH I), and wherein said one or more heterologous control sequences direct the production of the enzyme in the bacterium.

32. The construct of claim 31, wherein said $\alpha(1,2)$ fucosyltransferase enzyme comprises any one selected from *Parabacteroides johnsonii* FutX, *Lachnospiraceae* sp. FutQ, *Prevotella melaninogenica* FutO, *Prevotellaa* sp. FutW, *Bacteroides* sp. FutZA, *Tannerella* sp. FutS, *Bacteroides caccae* FutU, *Salmonella enterica* FutZ, *Methanosphaerula palustries* FutR, *Butyrivibrio* FutV, *Clostridium bolteae* FutP, or a functional variant or fragment thereof.

33. The construct of any one of claims 31 or 32, wherein said $\alpha(1,2)$ fucosyltransferase enzyme comprises any one of amino acid sequences SEQ ID NOs: 10-18, 20, 21, and 292, or a functional variant or fragment thereof.

34. The construct of any one of claims 31-33, wherein said heterologous control sequence comprises a bacterial promoter and operator, a bacterial ribosome binding site, a bacterial transcriptional terminator, a plasmid selectable marker, and/or an origin of replication.

35. An isolated bacterium comprising an isolated nucleic acid encoding a lactose-accepting $\alpha(1,2)$ fucosyltransferase enzyme, wherein the amino acid sequence of said enzyme encoded by said nucleic acid comprises at least 90% identity to FutW (SEQ ID NO: 17), FutP (SEQ ID NO: 11), FutO (SEQ ID NO: 10), FutQ (SEQ ID NO: 12), FutR (SEQ ID NO: 13), FutS (SEQ ID NO: 14), FutU (SEQ ID NO: 15), FutV (SEQ ID NO: 16), FutX (SEQ ID NO: 18), FutZ (SEQ ID NO: 20) or FutZA (SEQ ID NO: 21), wherein the lactose-utilizing $\alpha(1,2)$ fucosyltransferase enzyme is capable of producing a fucosylated oligosaccharide comprising 2'-fucosyllactose (2'-FL), lactodifucotetraose (LDFT), Lacto-N-fucopentaose I (LNF I), or lacto-N-difucohexaose I (LDFH I).

36. The isolated bacterium of claim 35, wherein said $\alpha(1,2)$ fucosyltransferase enzyme comprises any one selected from *Parabacteroides johnsonii* FutX, *Lachnospiraceae* sp.

FutQ, *Prevotella melaninogenica* FutO, *Prevotellaa sp.* FutW, *Bacteroides sp.* FutZA, *Tannerella sp.* FutS, *Bacteroides caccae* FutU, *Salmonella enterica* FutZ, *Methanosphaerula palustries* FutR, *Butyrivibrio* FutV, *Clostridium bolteae* FutP, or a functional variant or fragment thereof.

37. The isolated bacterium of any one of claims 35 or 36, wherein said $\alpha(1,2)$ fucosyltransferase enzyme comprises any one of amino acid sequences SEQ ID NOs: 10-18, 20, 21, and 292, or a functional variant or fragment thereof.

38. The isolated bacterium of any one of claims 35-37, further comprising a $\alpha(1,3)$ fucosyltransferase enzyme and/or an $\alpha(1,4)$ fucosyltransferase enzyme.

39. The isolated bacterium of any one of claims 35-38, wherein said bacterium is *Escherichia coli*.

40. The isolated bacterium of any one of claims 35-39, wherein said bacterium further comprises reduced level of β -galactosidase activity, a defective colanic acid synthesis pathway, an inactivated adenosine-5'-triphosphate (ATP)-dependent intracellular protease, an inactivated endogenous *lacA* gene, or any combination thereof.

41. The isolated bacterium of claim 40, wherein said bacterium comprises the genotype $\Delta ampC::P_{trp}^B cI$, $\Delta(lacI-lacZ)::FRT$, $P_{lacIq} lacY^+$, $\Delta wcaJ::FRT$, $thyA::Tn10$, $\Delta lon:(npt3, lacZ^+)$.

Synthetic routes for neutral hMOS

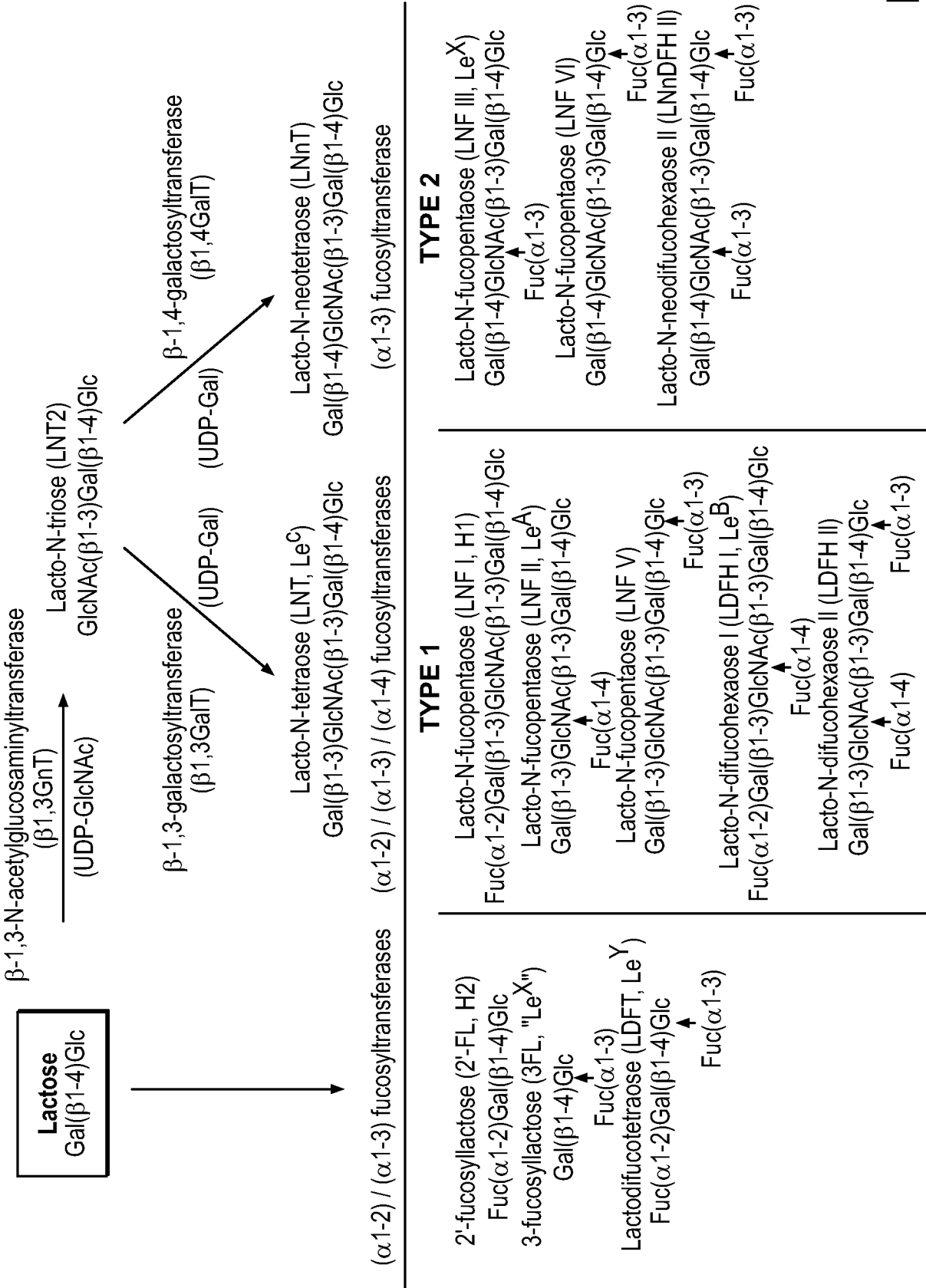


FIG. 1

Metabolic engineering for 2'-FL production in *E. coli*

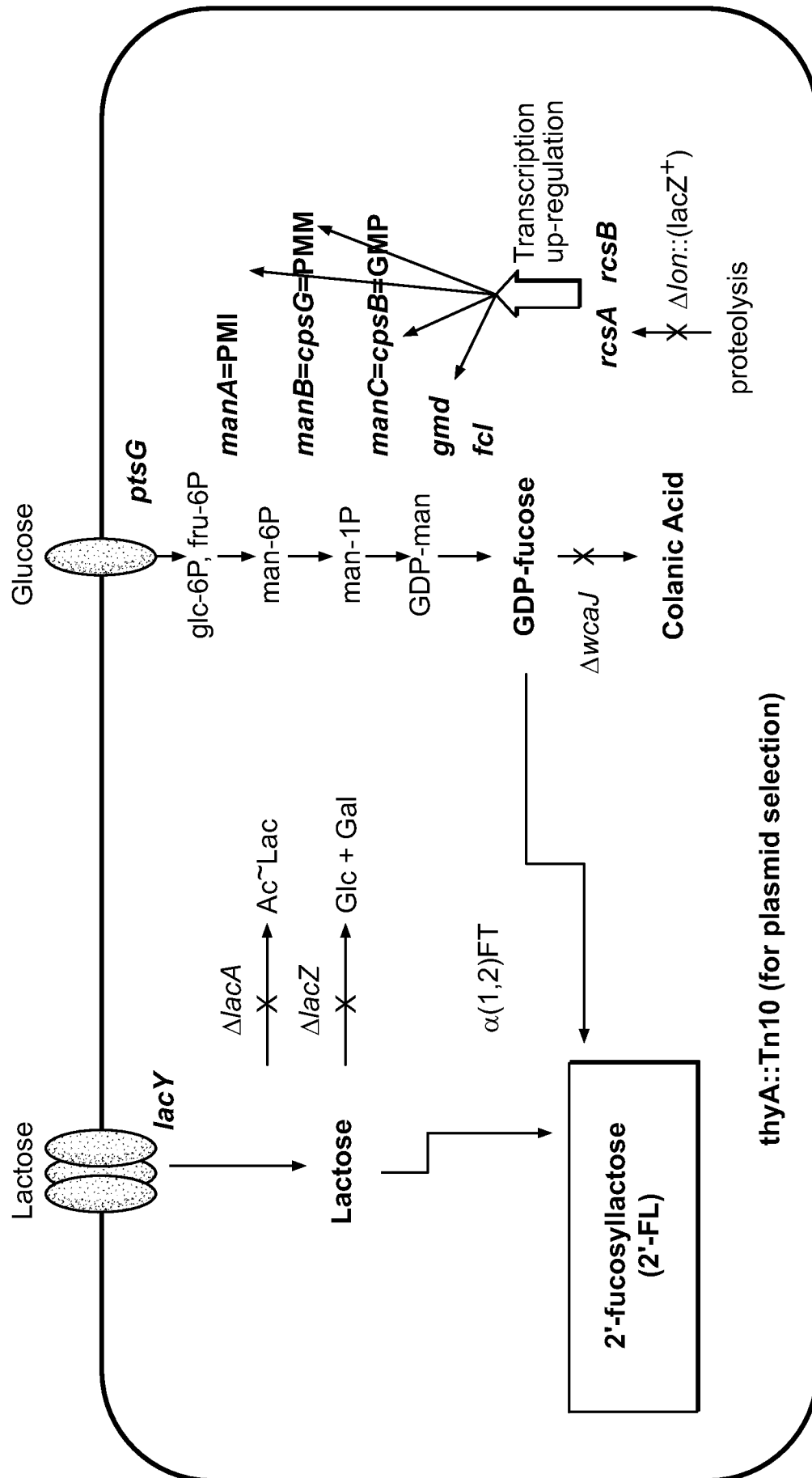


FIG. 2

3/19

	1	2	3	4
H.pylori futC	1	70.10	23.97	22.57
H.mustelae futL	2	70.10	27.04	21.36
Bacteroides vulgatus futN	3	23.97	27.04	31.23
E. coli O126 wbgL	4	22.57	21.36	31.23

FIG. 3A

H.pylori futC	MAFKVVQI	CGGLGNQM	FQYAF	AKSLQKHS	NTPVLL	DI	TSFDWS	DRKMQL	100																					
H.mustelae futL	MDFKIVQ	VHGGGLGNQM	FQYAF	AKSLQTHL	NI	PVLL	DTTWFDY	GNRELGLH	100																					
Bacteroides vulgatus futN	M - - RL	IKVTGGLGNQM	F IYAF - YRR	MKKYYPKVR	IDLSD	MMHYKVHYGE			95																					
E. coli 0126 wbgL	M - - SI	IRLQGGGLGNQL	FQFS	EGYALSK	INGT	PL YFD	ISHYA	ENDDHGGYR	103																					
H.pylori futC	LKPSRL	TYFFGYFQD	PRYFDA	ISPLIKQ	TFTLPP	PPENN	KNNKKE	EEYQ	200																					
H.mustelae futL	FEPSRI	AYFHGYFQD	PRYFED	ISPLIKQ	TFTLPH	PTHEH	- - - - -	AEQYS	193																					
Bacteroides vulgatus futN	- - - WPL	LYFKGFYQ	SERFFAD	IKDEV	RESFT	FDK - - - - -	NKANSRS		190																					
E. coli 0126 wbgL	AQKWSK	KYIGYWQ	SEHFFHKH	ILDLKE	FF - IPK - - - - -	NVS	EQAN		199																					
H.pylori futC	FVFCED	LEFT - QN	LDLGYP	FMDMT	TRDKE	EEAYWD	MLLMQ	SCQHG I IANS	100																					
H.mustelae futL	FLFCED	LEFV - QN	LDLGYP	FVDMT	TRDGA - -	AHWD	MMLMQ	SCKHG I ITNS	100																					
Bacteroides vulgatus futN	YIFSD	DLAWVK	ENLP - -	LQNAV	YIDWNT	DEL	SWQD	MMLMSHCKHH I ICNS	100																					
E. coli 0126 wbgL	FIFSD	DLFWCK	ENIETLL	SKKYN	IYYSED	LSQE	EDLWL	MSLANHH I IANS	100																					
	LFPI	DLPY	SAKE I A	AKMQHL - - - - -	PKLVR	DALCKMG	FDRVSQE I	VFYEYEPKL 100																						
	LFPI	DLQC	SAQQ I AA	AHMQNL - - - - -	PRLVR	GALRRMGL	GRVSKE I	VFYEYMPKL 100																						
	MHRV	FNLP	HTFC INQP -	LKKVIEFL - - - - -	FFKKI	YERKQ	APNSLRAF - -	EKK - YF 95																						
	LNNL -	QIPEEY	LQXYTRKJ	NNIYKFL	VRGSR	LYPE I	FLFLGFC	NEFHAYGYDFE - YI 103																						
	CKLS	L	L	LAA - KNS	VFVHIR	RRGDYVG - - - - -	IGCQL	GIDYQ	KKALEYMAKRV	NMEL 200																				
	RKLS	Q	L	LAA - KNS	VFVHIR	RRGDYMR - - - - -	LGWQL	DISYQL	RAIAYMAKRV	QNL EL 193																				
	LNM	LE	I	LDK	ENAV	SLHIR	RRGDY	LQPK	HWAT	TG	SV	CQLPY	YQNA	I	AEMS	RRVA	SPSY 190													
	L	L	A	AK	L	ES - QSS	LSIH	IR	RGD	Y	TKNT	AT	L	THG	VCS	LEY	YK	KALNK	IRD	LAM	IRDV 199									
	TS	SW	WA	A	YL	EN	PEKI	I	IGP	KHWL	FGHEN	I - - - -	LCKE	WVK	IESH	FVKS	QKYNA *	301												
	TS	SW	WA	A	YL	IKN	PEKI	I	IGP	SHWI	YGNEN	I - - - -	LCKD	WVK	IESQ	FETKS *	- - - - -	287												
	TS	SW	WG	AWL	NP	ND	KTV	I	VP	SRWF	QHSE	APDI	YP - -	TGW	IKVP	VS - - - - -	- - - - -	282												
	S	FS	WW	G	A	L	GT	S	ASQ	I	V	I	YPT	PWYD	I	TPK	NTY	I	P	I	VN	HW	IN	V	D	K	H	SSC *	- - - - -	298

FIG.

FIG. 3B

FIG. 4A	FIG. 4B
FIG. 4C	FIG. 4D
FIG. 4E	FIG. 4F

FIG. 4

	H. pylori futC H. mustelae futL Bacteroides vulgatus futN E. coli 0126 wbgL	MAFKVVQICGGLGNQMFAQAKSLQKHSNTP--VLLDITSFWDSDR MDFKIVQVHGGLGNQMFAQAKSLQTHLNIP--VLLDTTWFDYGNR M--RLIKVTGGLGNQMF YAFYLRMKKY--PKVR DLSDDMMH-- M--SIRLQGGGLGNQLFQFSFGYALSK NGTP--LYFD SHYAEN-- M-K- VK LGGLGNQMFAQALYLSLQES-FPKERVADLSSF-- m-v-iikmmg gnqMQYALYKAFEQK-HID--VYADLAWYKNKSV M-V- VQLSGGLGNQMFEYALYLSLKAK-GKEVKID-DVTCYEGPGT M- - VRLKGGGLGNQLSQYALGRK AHL-HNTE-LKLDTTWFTT SS M-VR VE IGGLGNQMFAQAFSLYLKNKSH WDRLYVD EAMKT-- M-K- VK LGGLGNQMFAQALFLSLKER-FPHEQVM DTSCF-- M- - QLKGGGLGNQMFAQALYKSLKKR-GKEVKID-DKTGFVNDKL M--RLVKM GGGLGNQMF YAFYLQMRKRF--SNVR DLTDDMMH-- M--RL KM GGGLGNQMF YAFYLKMKHHY--PDTN DLSDDMVH-- M--RLFGGLGNQLFQYAF FALSRRQGG--KARLETSSYEHDCK M-YSC--LSGGLGNQMFAQAAAY LKQYFQSTTLVLDDSYYSQPKR M--RL KMTGGGLGNQMF YAFYLRMKKRY--PKVR DLSDDMVH-- M--FQYALYKAFEQK-HID--VYADLAWYKNKSV
Prevotella melanogenica FutO YP_003814512.1		
Clostridium bolteae+13 FutP WP_002570768.1		
Lachnospiraceae sp. FutQ WP_009251343.1		
Methanospaerula palustris FutR YP_002467213.1		
Tannerella sp. FutS WP_021929367.1		
Bacteroides caccae FutU WP_005675707.1		
Butyrivibrio FutV WP_022772718.1		
Prevotella sp. FutW WP_022481266.1		
Parabacteroides johnsonii FutX WP_008155883.1		
Akkermansia muciniphila FutY YP_001877555		
Salmonella enterica FutZ WP_023214330		
Bacteroides sp. FutZA WP_022161880.1		
Clostridium bolteae FutP WP_002570768.1		

FIG. 4A

	KMQL	-----	ELFP	DLPYASAKE	A	AKMQHL	PKLVRDALKCM	-----	GFDRVSQE	VF	94					
	ELGL	-----	HLFP	DLQCASAQQ	AAAHMQNL	PRLVRGALRRM	-----	GLGRVSKE	VF	94						
	----	YKVHYGYEMHRVFNLP	----	TEFC	INQPLKKV	IEFLF	-FKKIY	-----	ERKQAPN	85						
	----	DDHGGYRLNNL	----	Q	IPPEYLQYYTPK	INN	YKFLVRGSR	LYPE	FLFLGFCNEFHAY	96						
	----	HGYHLHNGFELEN	FSVTAQKASAAD	IMR	AYYYPNYL	LWR	GKRFLPRR	KGM	-----	CLESSL	99					
	----	KFELYN	----	FG	IKINVASEKD	INRLSDCQADF	-VSR	RRK	FGKKKSF	-----	VSEKNDS	93				
	RPR	-----	QLDVFG	IYDRASREEL	TEMTDASMD	-ALS	RVRRLTG	-RRTK	-----	AYRERD	94					
	DTPRTYRLNN	----	YN	IGT	ASAKE	QL	IERGRAQQGRGYLLSK	SDL	LTPMYRRT	-----	YVRERMH	102				
	----	YDRHYGLELEKVFNL	SLCP	SNRL	HRNLQ	----	KRSFAKH	FVKS	-----	LYEHSEC	90					
	----	RNYPLHNGFEVDR	FAQKAPVASWRN	LK	VAYPPNYR	FWK	GKY	LPKR	KTM	-----	CVERKNF	99				
	RIP	-----	VL	SRWGV	EYDRATDEE		NL	TDSKMD	-LFSR	RRK	LTG	-RKTF	-----	R	DEESG	94
	----	YNVHYGYELHKVFGLPR	----	TEFC	MNQPLKKV	LEFLF	-FRT	IY	-----	ERKQH	-G	84				
	----	YKVHNGYEMNR	FDLSQ	----	TEFC	INRTLKK	LEFLF	-FKKIY	-----	ERRQDPS	85					
	RVC	----	ELHHFRVSLP	EGGPPWA	----	FR	KSR	PAC	LRS	LFAAP	KYPH	-----	FREEKRH	89		
	DTVRSLELNG	----	FN	SYDRFSFAD	----	EKEK	KL	RKF	KRN	PPKQ	SE	DS	ALFGKYALSDRA	104		
	----	YVHHGYEMHRVFNLP	----	TEFC	INQPLKKV	IEFLF	-FKKIY	-----	ERKQDPN	85						
	----	KFELYN	----	FG	IKINVASEKD	INRLSDCQADF	-VSR	RRK	FGKKKSF	-----	VSEKNDS	80				

FIG. 4B

H.pylori futC	EYEPKLLKPSR-LTYFFGYFQDPRYF--DAISPLIKQTFTLPPPPEN
H.mustelae futL	EYMPELFEPSR-IAYFHGYFQDPRYF--EDISPLIKQTFTLPHPTEH
Bacteroides vulgatus futN	SLRAFEKKYFWPLLYFKGFYQSERFF--ADIKDEVRESFTF--DK
E.coli 0126 wbgL	GYDFEYIAQKWKSKKYIGYWQSEHFHKHIL--DLKEFFI-----P
Prevotella melaninogenica FutO YP_003814512.1	RFDESVL R-QEGNRYFDGYWQDERYF--AA YREKVLKAFTFPAFKRA
Clostridium bolteae+13 FutP WP_002570768.1	CYENDIL R-M-DNVYLSGYWQTEKYF--SNTREKLEDYSF-ALVNS
Lachnospiraceae sp. FutQ WP_009251343.1	NFDPLVME--KDPALLEGCFQSDKYF--RDCEGRVREAYRFRGIESG
Methanosphaerula palustris FutR YP_002467213.1	TFDKAILT-VPDENVLDGYWQTEKYF--KDIEELRRREVTLKDEPDS
Tannerella sp. FutS WP_021929367.1	EFDEPVYRGLRPYRYRGYWQNEGYFV--DIEPMIREAFQFNVNILS
Bacteroides caccae FutU WP_005675707.1	SFDAAVLT-RKGDCYYDGYWQH EEF--CDMKETIWEAFSFPPEVDG
Butyrivibrio FutV WP_022772718.1	KFNPEILE--KENAYLVGYWQCDKYFDDKDVVREIREAF EKKPQE--
Prevotella sp. FutW WP_022481266.1	RMEPYTCQYVWPLVYFKGFYQSERYF--SEVKDEVRECFTF--NP
Parabacteroides johnsonii FutX WP_008155883.1	TLYPYEKRYFWPLLYFKGFYQSERFF--FDIKDDVRKA FSF--NL
Akkermansia muciniphilia FutY YP_001877555	GFDPGLAAPRRRHTYFKGYFQTEQYFL--HCREQLCREFR LKTPLTP
Salmonella enterica FutZ WP_023214330	FYTFETIKNIDKACLSFYQDAD---LLNKYKQLLPLFELR--DD
Bacteroides sp. FutZA WP_022161880.1	SLRAFEKKYLWPLLYFKGFYQSERFF--ADIKDEV RKAFTF--DS
Clostridium bolteae FutP WP_002570768.1	CYENDIL R-M-DNVYLSGYWQTEKYF--SNTREKLEDYSF-ALVNS

FIG. 4C

NKNNKKEEYQCKLSL|LAAKNSVFEVH|RRGDYVG-----|GCQLG|DYQKKALEYMAKRV-----P 196
-----AEQYSRKLSQ|LAAKNSVFEVH|RRGDYMR-----|LWQLD|SYQLRA|AYMAKRV-----Q 189
NK-----ANSRSLNMLE|LDKDENAVALSH|RRGDY|Q-PKHWA|TGSVCQLPYYQNA|AEMSRRV-----AS 187
KN-----VSEQANLLAAK|LESQSSLS|H|RRGDY|K-NKTATLTHGVCSEYK|KALNK|RD|LAM|----- 196
EN-----L---SLL-----EKL|DENS|ALHVR|RRGDYVG--NNL---YQG|CDLDY|YRTA|EKMCAHV-----TP 194
QV-----SEWEDS|-----R---NKN|SVS|H|RRGDY|Q-GEL---YGG|CTSLY|YAEA|EY|KMRV-----PN 187
AF-----PLPEDYLRLEKQ|IEDCQSVSVH|RRGDYLD-ESHGGLYTG|CTEAYYKEA|FARMERLV-----PG 198
|IN-----LEMAER|-----QAC---HSVSLHVR|RRGDYVS-NPTTQQFHGCCS|DY|YNRA|SL|EEKV-----DD 201
KK-----TKA|AS-----KMRREL|SVS|HVR|RRGDYEN-PEAKAMHGG|CSLDY|YHKA|DF|RQRL-----DN 192
RN-----KE|GALL-----QASD---SASLHVR|RRGDYVN-HPL---FRG|CDLDY|YKRA|HYMEERV-----NP 195
LM-----TDASSWSTLQ-Q|IECCESVSLHVR|RTDYVD-EEH|H|HN-|CTEKY|YKNA|DRVRKQY-----PS 196
AL-----ANRSSQMMMEQ|QNDPQAVS|H|RRGDY|LN-PKH|YDT|GC|CQLP|YKHAVSE|KKYV-----SN 186
NI-----ANPFSLFLKQ|EVDDQAVS|H|RRGDY|LL-PRHWANTGSVCQLP|YKNA|AEMENR|-----TG 187
EN-----AR|LED|-----RSCCS|SLH|RR|TDYLSNP-----YLSPPPLEY|LRSMAEMEGR|LRAAGAPQE 191
LL-----DICKNLELYSL|QRSNNTTALH|RRGDYVT-NQHA|AKYHGVLD|SY|YNHAMEY|VERE-----RG 205
SK-----VNARSAELLRRLDADANAVSLH|RRGDY|LQ-PQHWAT|TGSVCQLP|YQNA|AEMNRRV-----AA 187
QV-----SEWEDS|-----R---NKN|SVS|H|RRGDY|LQ-GEL---YGG|CTSLY|YAEA|EY|KMRV-----PN 174

FIG. 4D

H. pylori futC	NMELFVFCEDLEFT-QNLDLGY - - - FMDMTTRDKEEEAYWDMMLMQ
H. mustelae futL	NLELFLFCEDLEFV-QNLDLGY - - - FVDMTTTRDGA - - AHWDMMLMQ
Bacteroides vulgatus futN	PS - YY FSDD AWWKENL - - - PLQNAVY DWNTDE - DSWQDMMMLMS
E. coli 0126 wbgL	- RDVF FSDD FWCKEN ETL LSKY N YYS EDLSQEE - - - DLWLMS
Prevotella melaninogenica FutO YP_003814512.1	SL - FC FSND TWCQQHLQPYLKAP - VVYV TWNTGV - ESYRDMQLMS
Clostridium bolteae+13 FutP WP_002570768.1	AK - FFVFSDDVEWVKQED - - - FKGFV VDRNEYS - SALSDMYLMS
Lachnospiraceae sp. FutQ WP_009251343.1	AR - FFLFSNDPEWTRHFE - - - SKNCVLVEGSTED - TGYMDLYLMS
Methanosphaerula palustris FutR YP_002467213.1	PS - FF FSDDL PWAKENLD - - PGEKT - FVAHNGPE - KEYCDLWLMS
Tannerella sp. FutS WP_021929367.1	N CFYLFSDD NWVEENLQL - - - ENRC DWNQGE - DSWQDMYLMS
Bacteroides caccae FutU WP_005675707.1	QL - YCVFSNDMAWCESHLRALPGKEVYVDWNKGA - ESYVDMRLMS
Butyrivibrio FutV WP_022772718.1	AV - FF FTDDKEWCRDHFk - - - GPNF VVELEEGDGT AEMRLMS
Prevotella sp. FutW WP_022481266.1	PH - FYVFSDDL DWVKANL - - - PLENAQY DWNKGA - DSWQDMMMLMS
Parabacteroides johnsonii FutX WP_008155883.1	PS - YYVFSDD SWVKEN - - - PLKKAVYV TWNKGE - DSWQDMMMLMS
Akkermansia muciniphilia FutY YP_001877555	SLRYF FSDD EWARQNLRPALP - - - HVHVD NDG - GTGYFDLELMR
Salmonella enterica FutZ WP_023214330	KQNF FSDDVRWAQKAFLE - - - NDNCYV NNSDYDFSA DMYLMS
Bacteroides sp. FutZA WP_022161880.1	PS - YYVFSDD AWWKEN - - - PLQNAVY DWNKGE - ESWQDMMMLMS
Clostridium bolteae FutP WP_002570768.1	AK - FFVFSDDVEWVKQED - - - FKGFV VDRNEYS - SALSDMYLMS

FIG. 4E

```
SCQHG|ANSTYSWWAAYL|ENPEK|||GPKHWF|GHEN|L|-----CKEWWK|ESHFEVKSQKYNA|301
|SCKHG|TNSTYSWWAAYL|KNPEK|||GPSHW|YGNEN|L|-----CKDWWK|ESQFETKS*|-----287
|HCKHH|CNSTFSWWGAWL|PNMDKTV|VPSRWFQ|-----HSEAPD|YPT|GW|KV|-----PVS*|-----282
|LANHH|ANSSFSWWGAYL|GTSASQ|V|YPTPWYD|TPKNTY|-----P|VNHWN|NVDKHSSC*|-----298
|CCAHN|ANSSFSWWGAWL|QNREKVV|APKKNL|N|-----MEECHFTLPA|S|-----W|-----K|-----288
|LCKHN|ANSSFSWWAAYL|NRNEEK|V|APRRWLN|GKCT|-----PD|WCK|-----KW|-----R|-----278
|RCRHN|ANSSFSWWGAWL|NENPEKV|APAKWLN|GR|-----ECRD|YTE|-----RM|-----RL|-----289
|LCQHH|ANSSFSWWGAWL|GQDAEKMV|APRRWAL|SESFDTS|D|PD|S|-----W|-----T|-----295
|CCRHH|ANSSFSWWAAYL|NPNKNK|V|LTPNKWFN|-----HTDAVG|VPK|-----SW|K|-----PVF|-----287
|LCRHN|ANSSFSWWGAWL|NRNPQKV|V|APERWMNSP|ED|-----PV|-----SD|-----KW|-----KL|-----289
|RCKHH|ANSSFSWWAAYL|NDSPEK|V|APQKW|N|NR|-----DMDD|YTE|-----RMTK|-----AL|-----290
|CCKHH|CNSTFSWWAAYL|NPSVEKTV|IMPEQWTS|-----RQDSVDFV|ASCGRWVRV|-----KTE|-----282
|HCRHH|CNSTFSWWGAWL|NPRKEK|V|APCRWFQ|-----HKETPDMYPK|-----EW|KV|-----P|N|-----281
|NCRHH|ANSTFSWWAAYL|NEHAEK|V|APR|W|FNREEGDRYHTDDAL|P|-----GSWLR|-----290
|LCKNN|ANSTYSWWGAWL|NKYEDKLV|SPKQWFLGNNETS|-----LRNASW|TL|-----298
|HCRHH|CNSTFSWWGAWL|DPHEDK|V|VPNRWFQ|-----HCETPN|YPA|-----GWVKV|-----A|N|-----281
|LCKHN|ANSSFSWWAAYL|NRNEEK|V|APRRWLN|GKCT|-----PD|WCK|-----KW|-----R|-----265
```

FIG. 4F

11/19

FIG. 5A

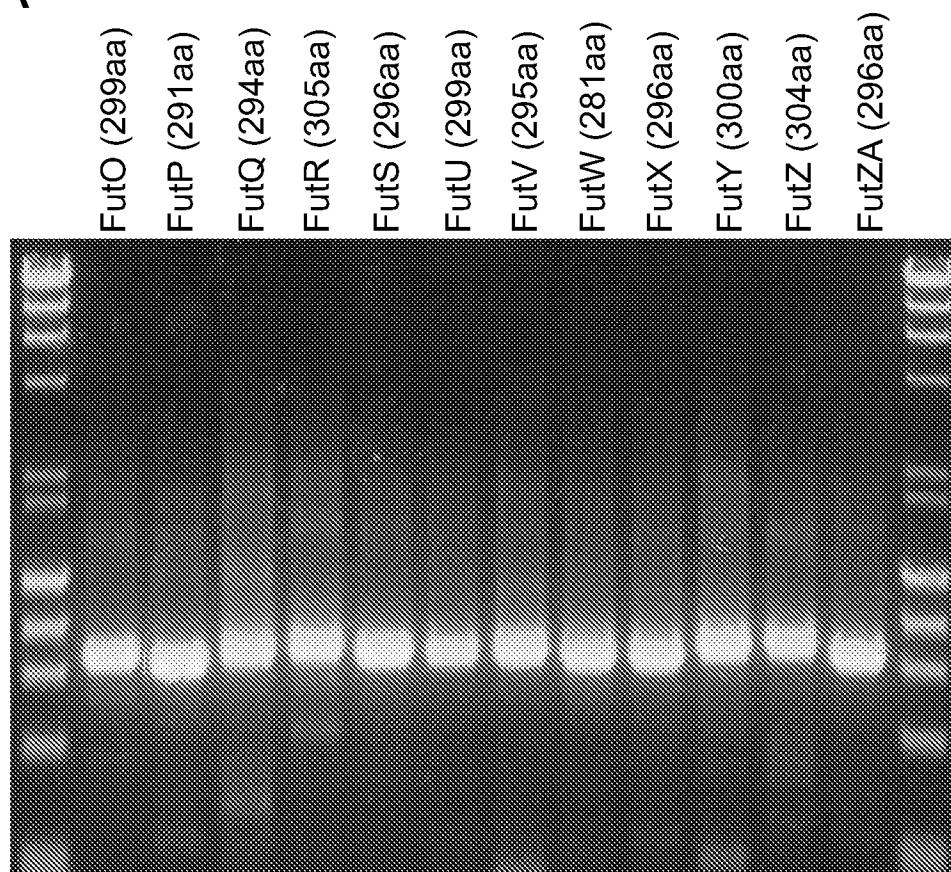
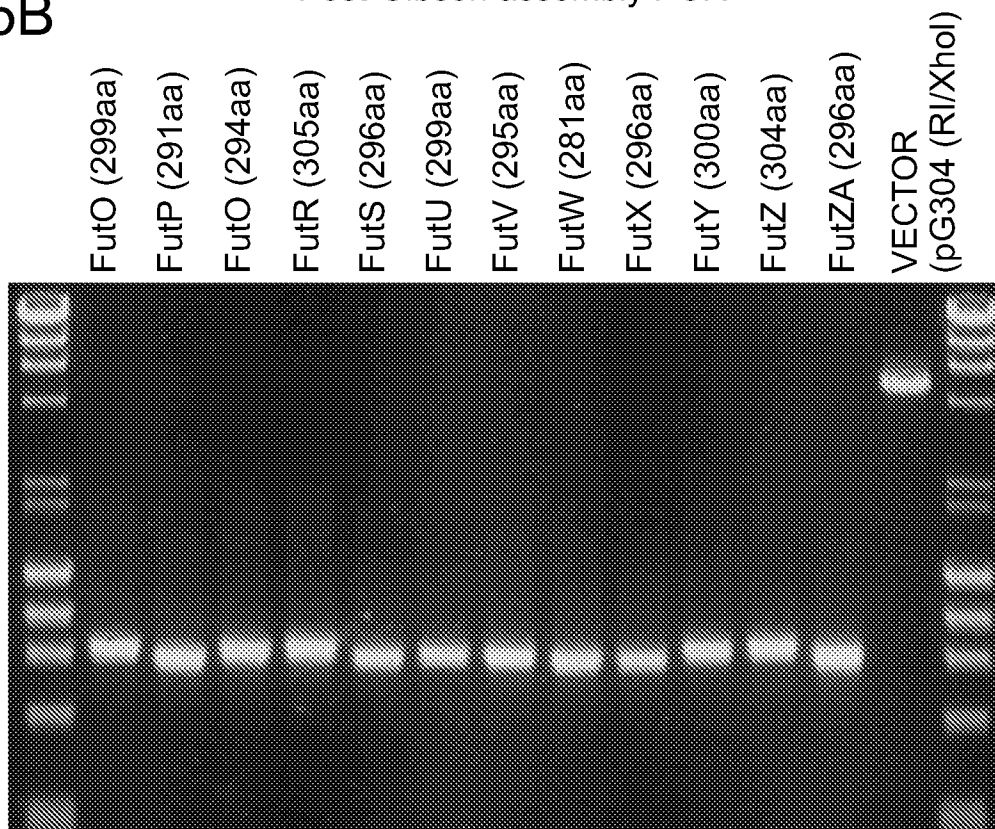
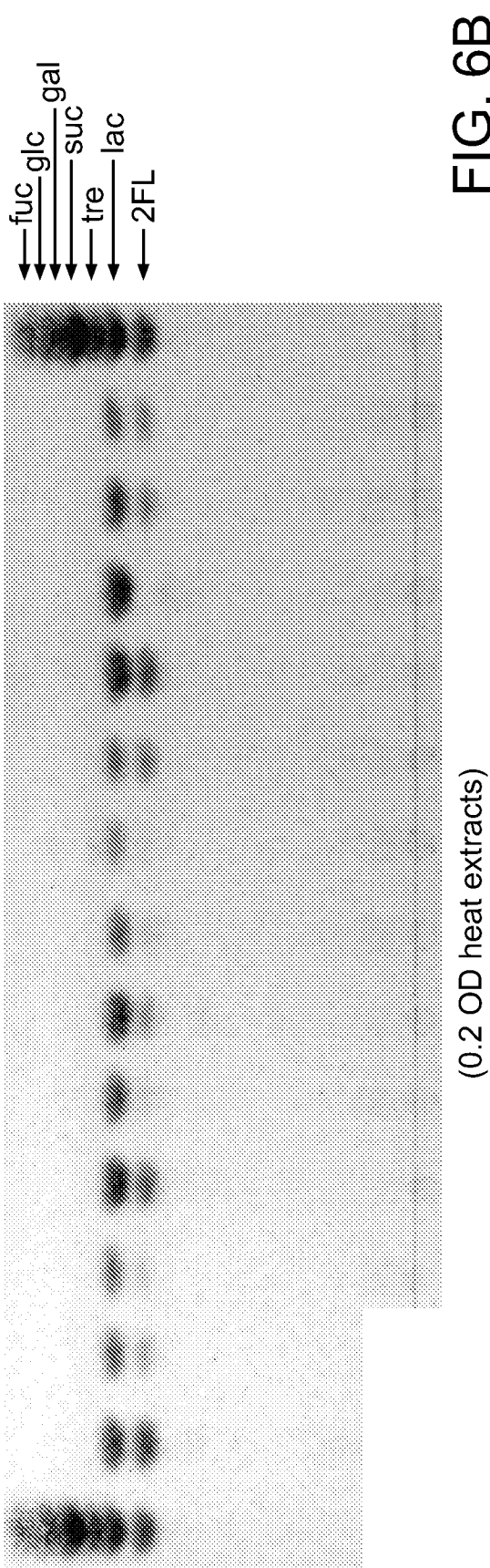
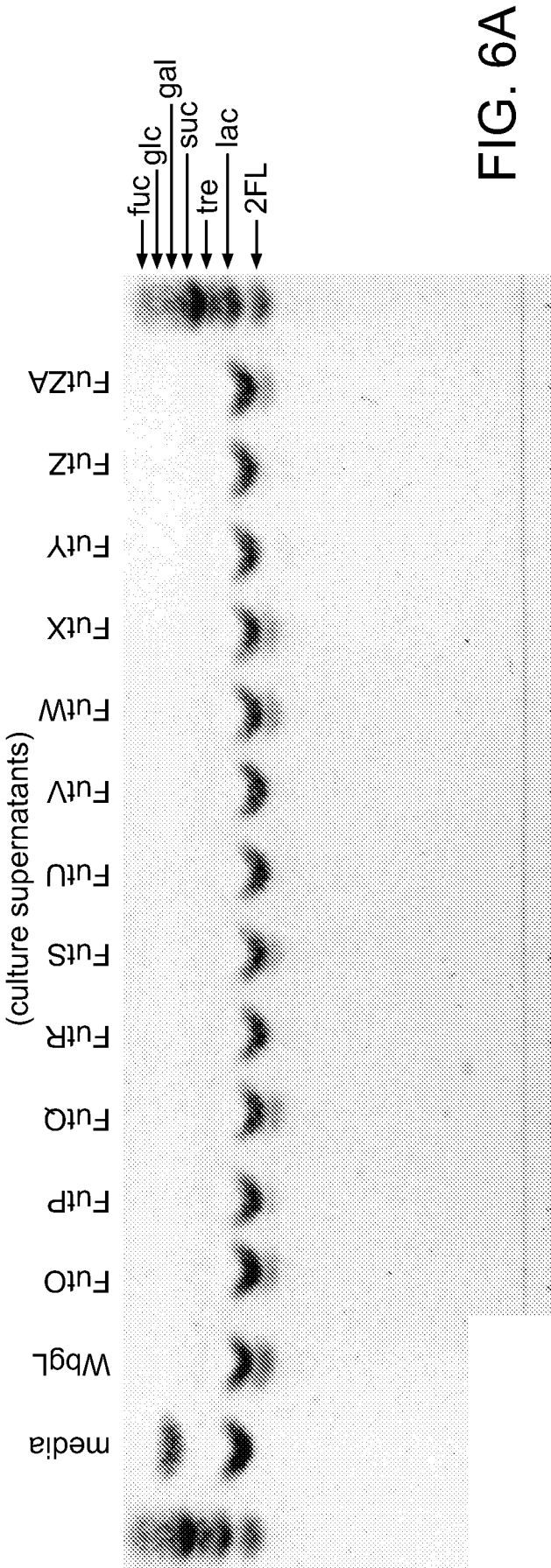


FIG. 5B

Post-Gibson assembly PCR



Gel-purified RI/XhoI syngene fragment



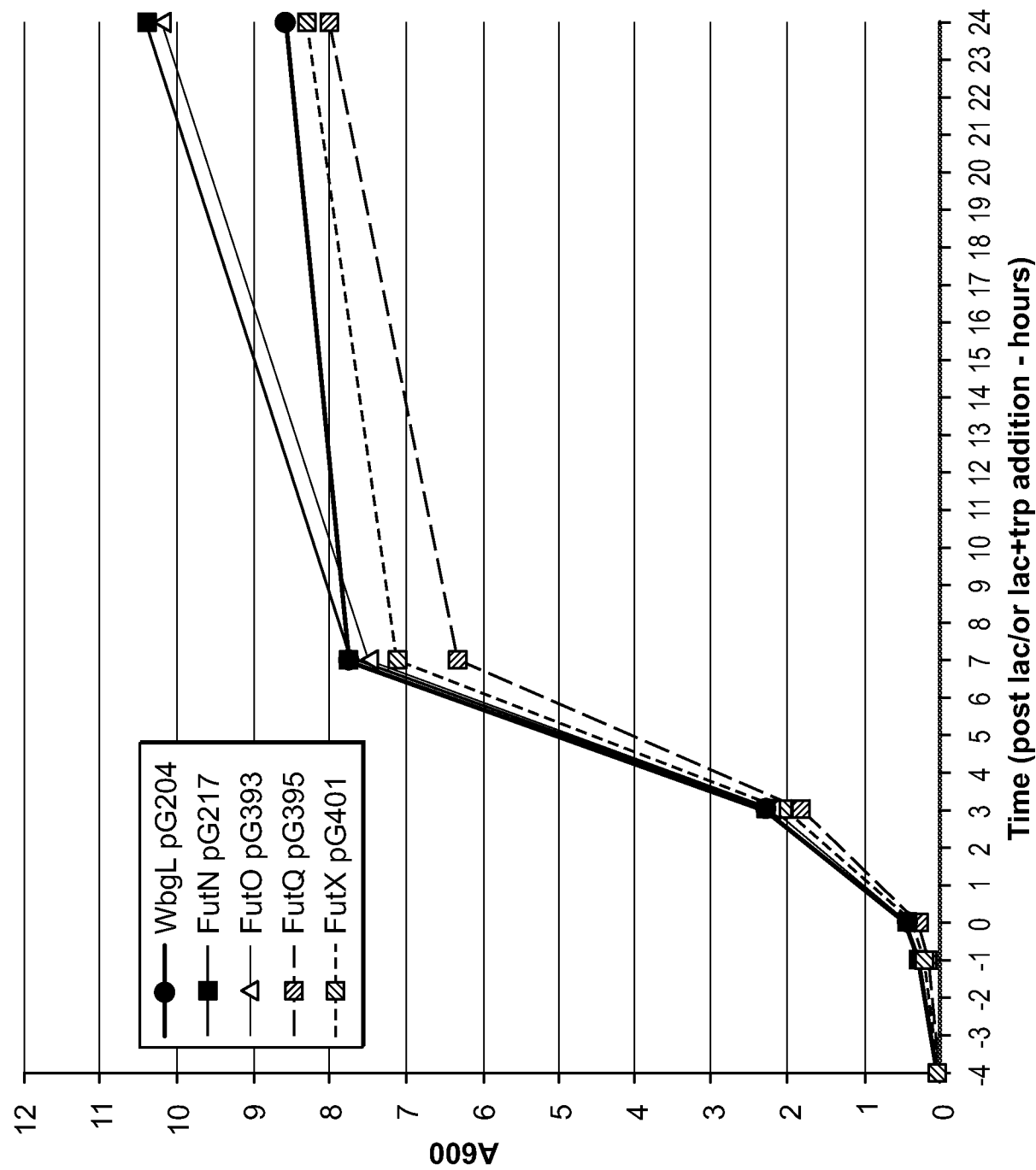


FIG. 7

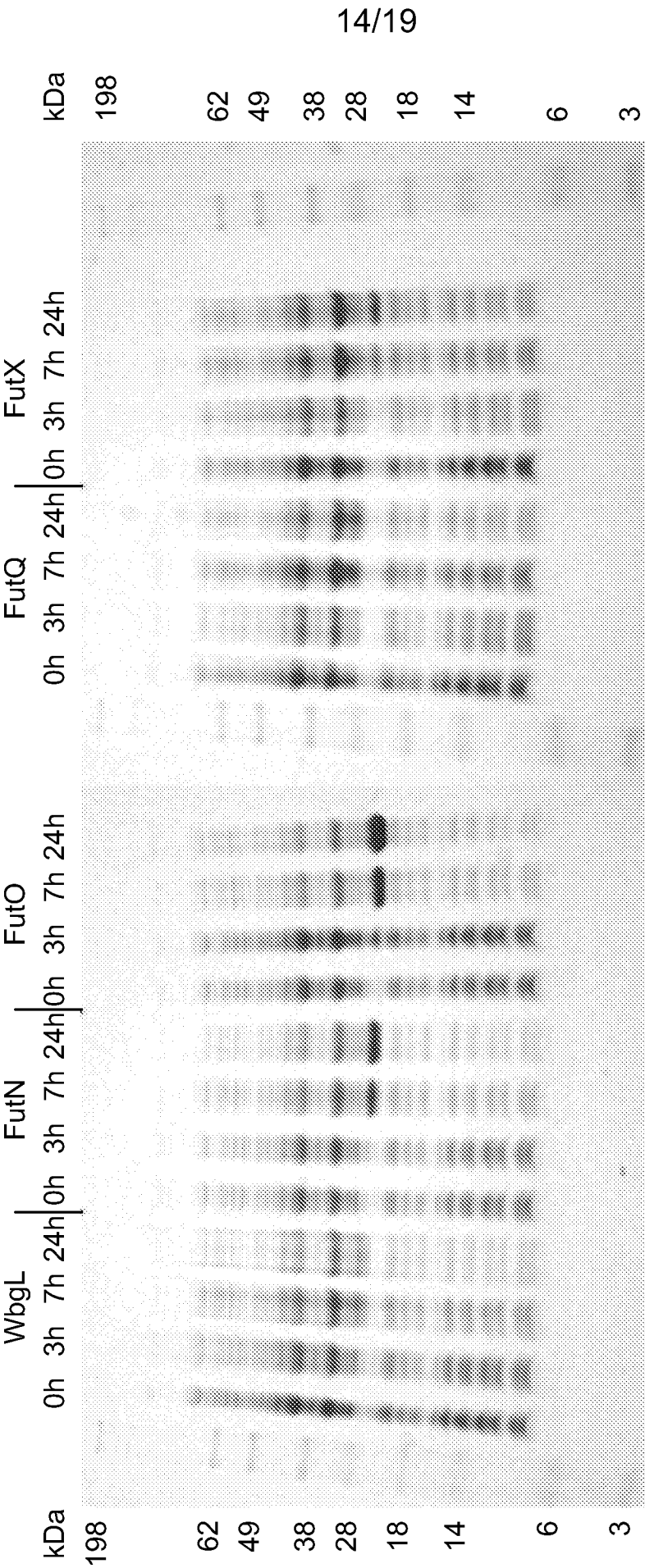


FIG. 8

FIG. 9A

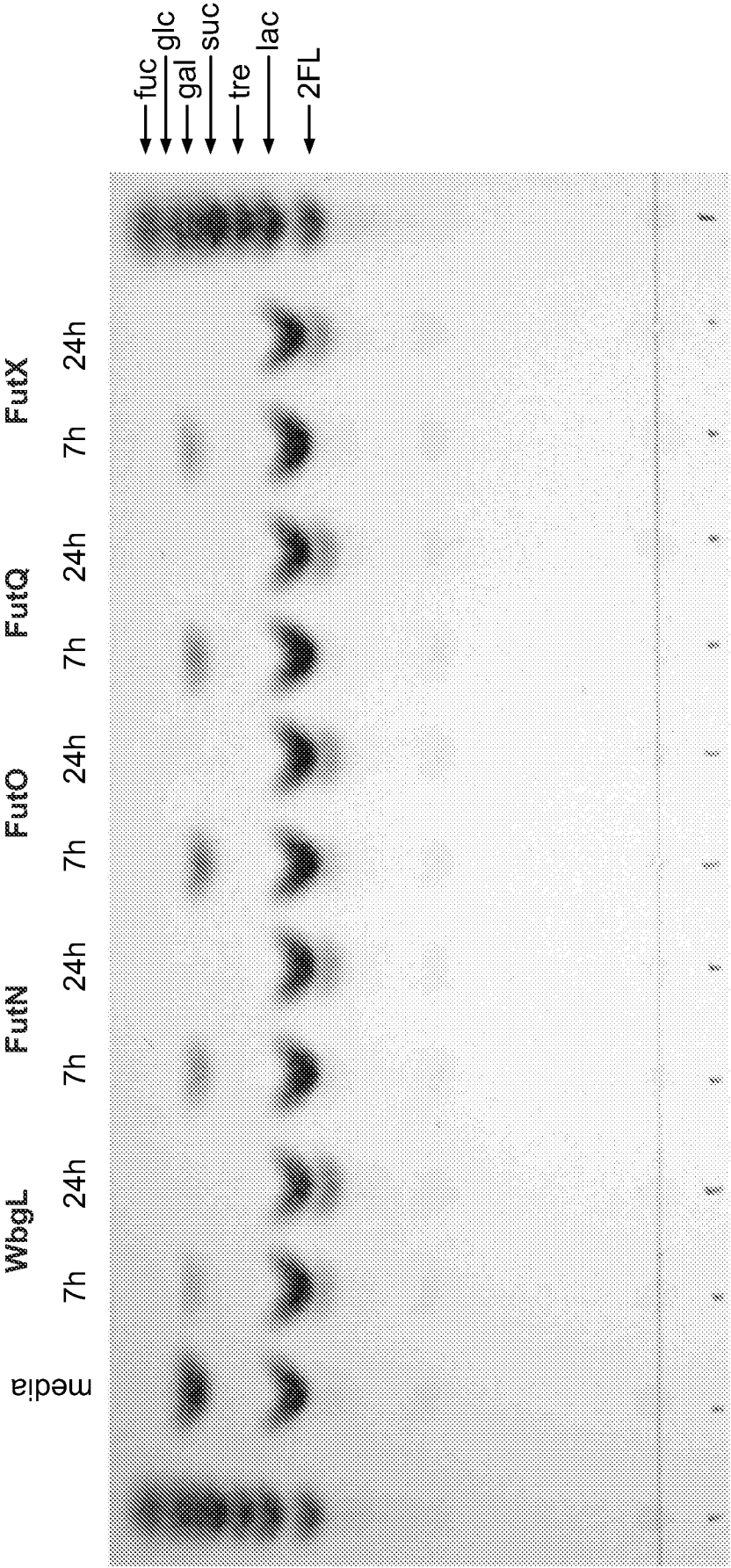
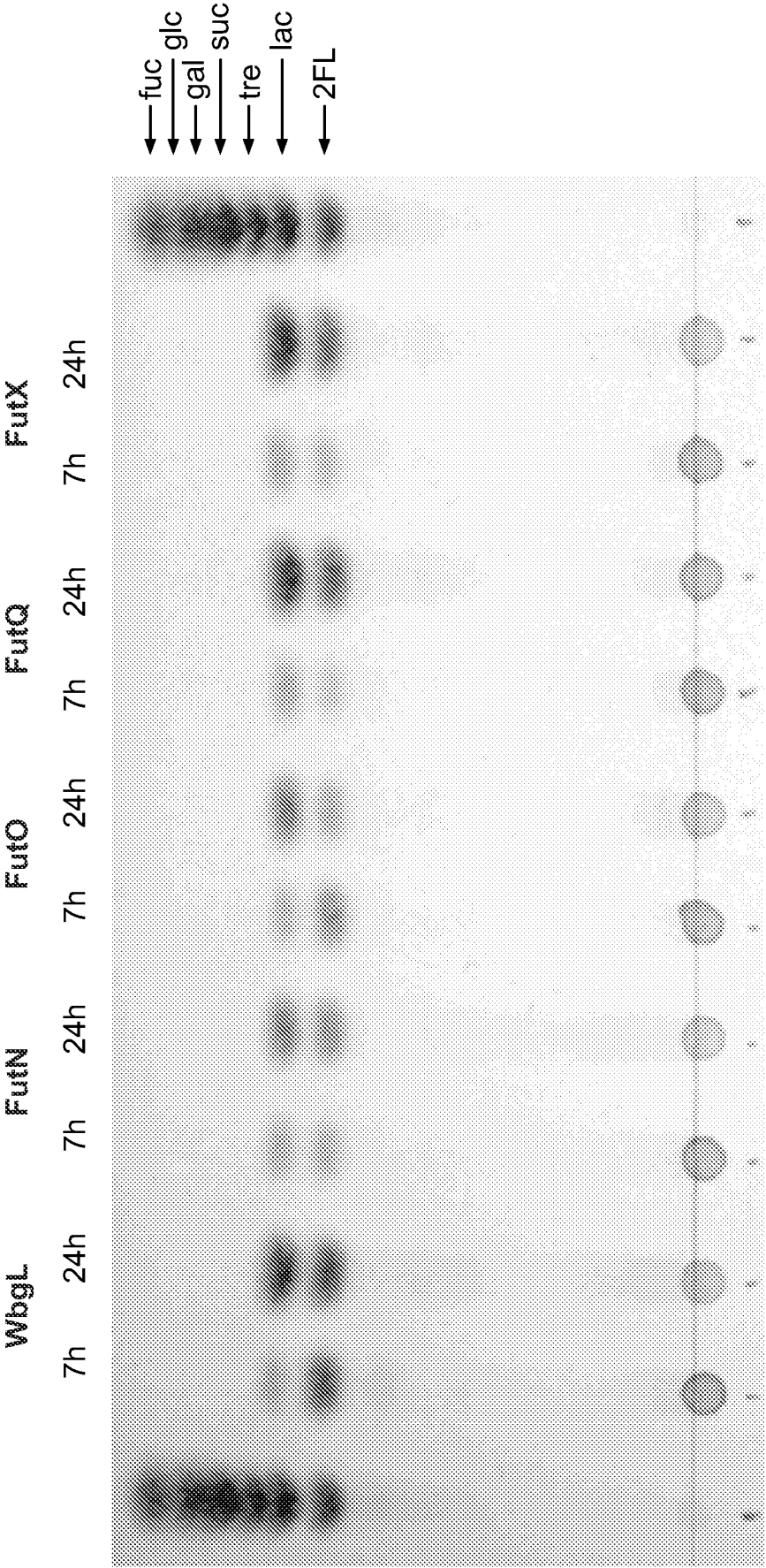
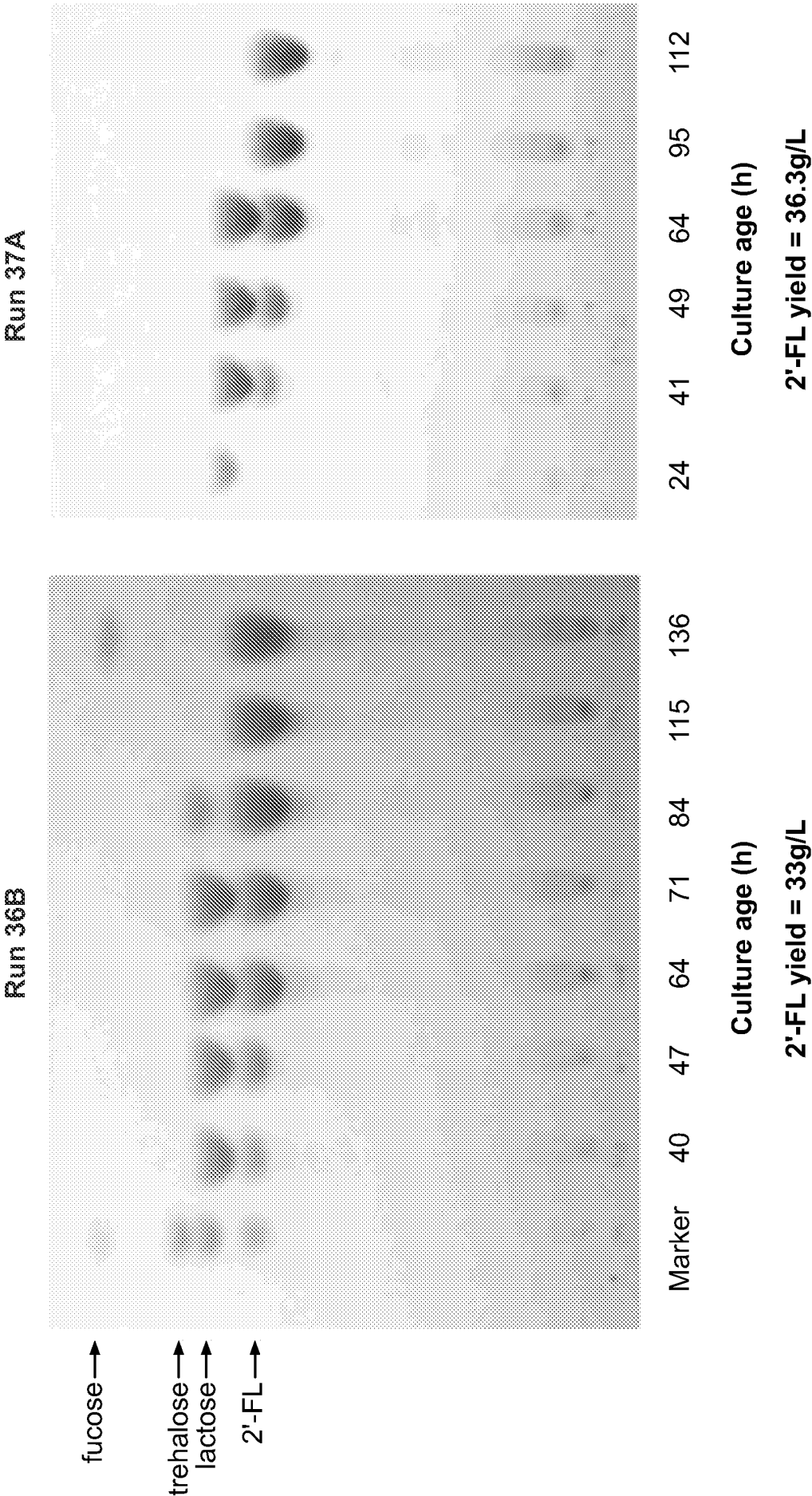


FIG. 9B





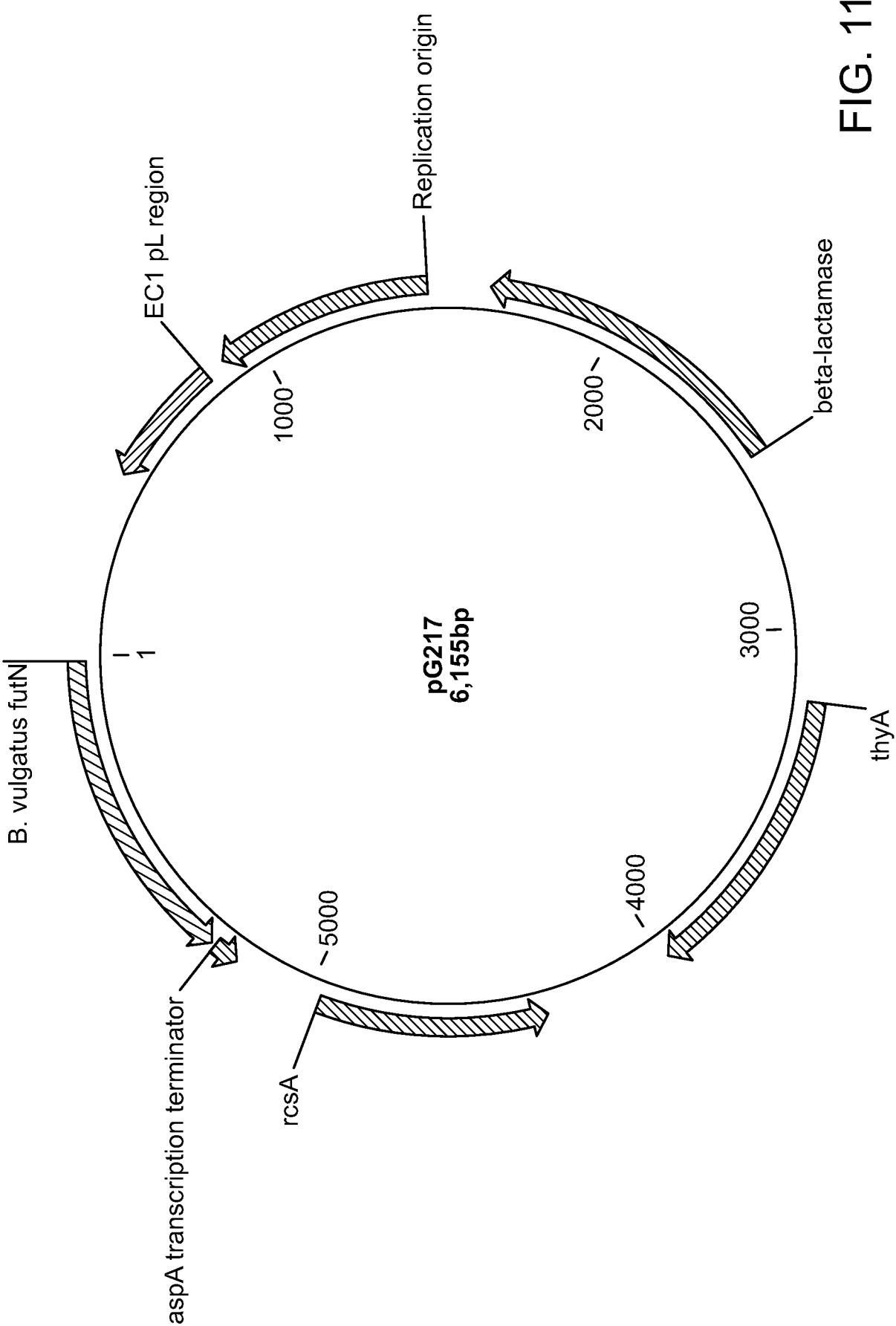
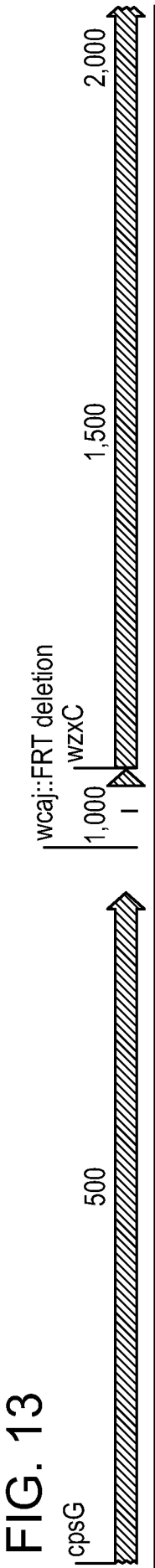
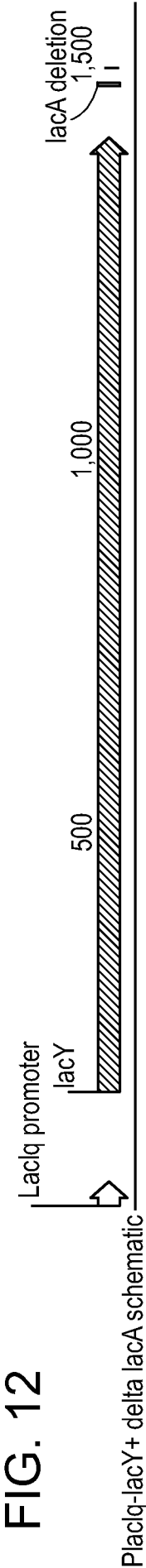
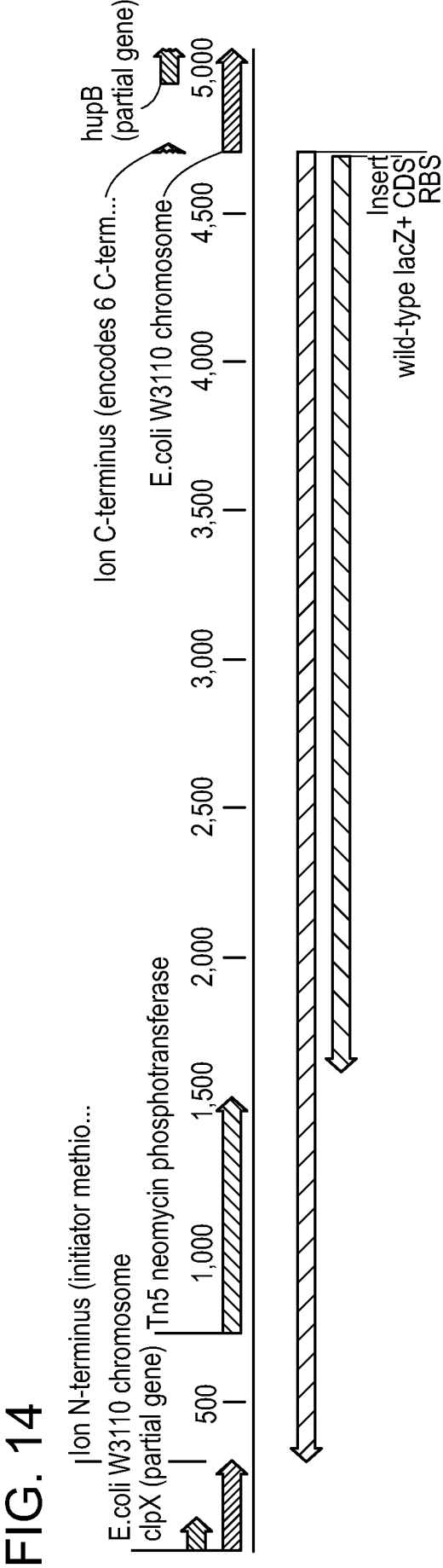


FIG. 11



19/19



37847-517 Sequence Listing_ST25 (003)
SEQUENCE LISTING

2015259088 28 May 2018

<110> Glycosyn LLC
McCoy, John
Heidtman, Matthew Ian
Merighi, Massimo Ian

<120> Alpha (1,2) Fucosyltransferase Syngenes For Use in the Production
of Fucosylated Oligosaccharides

<130> 37847-517001WO

<140> PCT/US2015/030823

<141> 2015-05-14

<160> 294

<170> PatentIn version 3.5

<210> 1

<211> 300

<212> PRT

<213> Helicobacter pylori

<400> 1

Met Ala Phe Lys Val Val Gln Ile Cys Gly Gly Leu Gly Asn Gln Met
1 5 10 15

Phe Gln Tyr Ala Phe Ala Lys Ser Leu Gln Lys His Ser Asn Thr Pro
20 25 30

Val Leu Leu Asp Ile Thr Ser Phe Asp Trp Ser Asp Arg Lys Met Gln
35 40 45

Leu Glu Leu Phe Pro Ile Asn Leu Pro Tyr Ala Ser Ala Lys Glu Ile
50 55 60

Ala Ile Ala Lys Met Gln His Leu Pro Lys Leu Val Arg Asp Ala Leu
65 70 75 80

Lys Cys Met Gly Phe Asp Arg Val Ser Gln Glu Ile Val Phe Glu Tyr
85 90 95

Glu Pro Glu Leu Leu Lys Pro Ser Arg Leu Thr Tyr Phe Tyr Gly Tyr
100 105 110

37847-517 Sequence Listing_ST25 (003)

Phe Gln Asp Pro Arg Tyr Phe Asp Ala Ile Ser Pro Leu Ile Lys Gln
 115 120 125

Thr Phe Thr Leu Pro Pro Pro Pro Glu Asn Asn Lys Asn Asn Asn Lys
 130 135 140

Lys Glu Glu Glu Tyr His Arg Lys Leu Ser Leu Ile Leu Ala Ala Lys
 145 150 155 160

Asn Ser Val Phe Val His Ile Arg Arg Gly Asp Tyr Val Gly Ile Gly
 165 170 175

Cys Gln Leu Gly Ile Asp Tyr Gln Lys Lys Ala Leu Glu Tyr Met Ala
 180 185 190

Lys Arg Val Pro Asn Met Glu Leu Phe Val Phe Cys Glu Asp Leu Glu
 195 200 205

Phe Thr Gln Asn Leu Asp Leu Gly Tyr Pro Phe Met Asp Met Thr Thr
 210 215 220

Arg Asn Lys Glu Glu Glu Ala Tyr Trp Asp Met Leu Leu Met Gln Ser
 225 230 235 240

Cys Gln His Gly Ile Ile Ala Asn Ser Thr Tyr Ser Trp Trp Ala Ala
 245 250 255

Tyr Leu Ile Glu Asn Pro Glu Lys Ile Ile Ile Gly Pro Lys His Trp
 260 265 270

Leu Phe Gly His Glu Asn Ile Leu Cys Lys Glu Trp Val Lys Ile Glu
 275 280 285

Ser His Phe Glu Val Lys Ser Gln Lys Tyr Asn Ala
 290 295 300

<210> 2
 <211> 286

37847-517 Sequence Listing_ST25 (003)

<212> PRT

<213> *Helicobacter mustelae*

<400> 2

Met Asp Phe Lys Ile Val Gln Val His Gly Gly Leu Gly Asn Gln Met
 1 5 10 15

Phe Gln Tyr Ala Phe Ala Lys Ser Leu Gln Thr His Leu Asn Ile Pro
 20 25 30

Val Leu Leu Asp Thr Thr Trp Phe Asp Tyr Gly Asn Arg Glu Leu Gly
 35 40 45

Leu His Leu Phe Pro Ile Asp Leu Gln Cys Ala Ser Ala Gln Gln Ile
 50 55 60

Ala Ala Ala His Met Gln Asn Leu Pro Arg Leu Val Arg Gly Ala Leu
 65 70 75 80

Arg Arg Met Gly Leu Gly Arg Val Ser Lys Glu Ile Val Phe Glu Tyr
 85 90 95

Met Pro Glu Leu Phe Glu Pro Ser Arg Ile Ala Tyr Phe His Gly Tyr
 100 105 110

Phe Gln Asp Pro Arg Tyr Phe Glu Asp Ile Ser Pro Leu Ile Lys Gln
 115 120 125

Thr Phe Thr Leu Pro His Pro Thr Glu His Ala Glu Gln Tyr Ser Arg
 130 135 140

Lys Leu Ser Gln Ile Leu Ala Ala Lys Asn Ser Val Phe Val His Ile
 145 150 155 160

Arg Arg Gly Asp Tyr Met Arg Leu Gly Trp Gln Leu Asp Ile Ser Tyr
 165 170 175

Gln Leu Arg Ala Ile Ala Tyr Met Ala Lys Arg Val Gln Asn Leu Glu
 180 185 190

37847-517 Sequence Listing_ST25 (003)

Leu Phe Leu Phe Cys Glu Asp Leu Glu Phe Val Gln Asn Leu Asp Leu
195 200 205

Gly Tyr Pro Phe Val Asp Met Thr Thr Arg Asp Gly Ala Ala His Trp
210 215 220

Asp Met Met Leu Met Gln Ser Cys Lys His Gly Ile Ile Thr Asn Ser
225 230 235 240

Thr Tyr Ser Trp Trp Ala Ala Tyr Leu Ile Lys Asn Pro Glu Lys Ile
245 250 255

Ile Ile Gly Pro Ser His Trp Ile Tyr Gly Asn Glu Asn Ile Leu Cys
260 265 270

Lys Asp Trp Val Lys Ile Glu Ser Gln Phe Glu Thr Lys Ser
275 280 285

<210> 3

<211> 281

<212> PRT

<213> Bacteroides vulgatus

<400> 3

Met Arg Leu Ile Lys Val Thr Gly Gly Leu Gly Asn Gln Met Phe Ile
1 5 10 15

Tyr Ala Phe Tyr Leu Arg Met Lys Lys Tyr Tyr Pro Lys Val Arg Ile
20 25 30

Asp Leu Ser Asp Met Met His Tyr Lys Val His Tyr Gly Tyr Glu Met
35 40 45

His Arg Val Phe Asn Leu Pro His Thr Glu Phe Cys Ile Asn Gln Pro
50 55 60

Leu Lys Lys Val Ile Glu Phe Leu Phe Phe Lys Lys Ile Tyr Glu Arg
65 70 75 80

37847-517 Sequence Listing_ST25 (003)

Lys	Gln	Ala	Pro	Asn	Ser	Leu	Arg	Ala	Phe	Glu	Lys	Lys	Tyr	Phe	Trp	
				85					90					95		
Pro	Leu	Leu	Tyr	Phe	Lys	Gly	Phe	Tyr	Gln	Ser	Glu	Arg	Phe	Phe	Ala	
			100					105					110			
Asp	Ile	Lys	Asp	Glu	Val	Arg	Glu	Ser	Phe	Thr	Phe	Asp	Lys	Asn	Lys	
		115					120					125				
Ala	Asn	Ser	Arg	Ser	Leu	Asn	Met	Leu	Glu	Ile	Leu	Asp	Lys	Asp	Glu	
	130					135					140					
Asn	Ala	Val	Ser	Leu	His	Ile	Arg	Arg	Gly	Asp	Tyr	Leu	Gln	Pro	Lys	
145					150					155					160	
His	Trp	Ala	Thr	Thr	Gly	Ser	Val	Cys	Gln	Leu	Pro	Tyr	Tyr	Gln	Asn	
				165					170					175		
Ala	Ile	Ala	Glu	Met	Ser	Arg	Arg	Val	Ala	Ser	Pro	Ser	Tyr	Tyr	Ile	
			180					185					190			
Phe	Ser	Asp	Asp	Ile	Ala	Trp	Val	Lys	Glu	Asn	Leu	Pro	Leu	Gln	Asn	
		195					200					205				
Ala	Val	Tyr	Ile	Asp	Trp	Asn	Thr	Asp	Glu	Asp	Ser	Trp	Gln	Asp	Met	
	210					215					220					
Met	Leu	Met	Ser	His	Cys	Lys	His	His	Ile	Ile	Cys	Asn	Ser	Thr	Phe	
225					230					235					240	
Ser	Trp	Trp	Gly	Ala	Trp	Leu	Asn	Pro	Asn	Met	Asp	Lys	Thr	Val	Ile	
				245					250					255		
Val	Pro	Ser	Arg	Trp	Phe	Gln	His	Ser	Glu	Ala	Pro	Asp	Ile	Tyr	Pro	
			260					265					270			
Thr	Gly	Trp	Ile	Lys	Val	Pro	Val	Ser								
		275					280									

37847-517 Sequence Listing_ST25 (003)

<210> 4
 <211> 297
 <212> PRT
 <213> Escherichia coli

<400> 4

Met Ser Ile Ile Arg Leu Gln Gly Gly Leu Gly Asn Gln Leu Phe Gln
 1 5 10 15

Phe Ser Phe Gly Tyr Ala Leu Ser Lys Ile Asn Gly Thr Pro Leu Tyr
 20 25 30

Phe Asp Ile Ser His Tyr Ala Glu Asn Asp Asp His Gly Gly Tyr Arg
 35 40 45

Leu Asn Asn Leu Gln Ile Pro Glu Glu Tyr Leu Gln Tyr Tyr Thr Pro
 50 55 60

Lys Ile Asn Asn Ile Tyr Lys Leu Leu Val Arg Gly Ser Arg Leu Tyr
 65 70 75 80

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

Gly Tyr Asp Phe Glu Tyr Ile Ala Gln Lys Trp Lys Ser Lys Lys Tyr
 100 105 110

Ile Gly Tyr Trp Gln Ser Glu His Phe Phe His Lys His Ile Leu Asp
 115 120 125

Leu Lys Glu Phe Phe Ile Pro Lys Asn Val Ser Glu Gln Ala Asn Leu
 130 135 140

Leu Ala Ala Lys Ile Leu Glu Ser Gln Ser Ser Leu Ser Ile His Ile
 145 150 155 160

Arg Arg Gly Asp Tyr Ile Lys Asn Lys Thr Ala Thr Leu Thr His Gly
 165 170 175

Val Cys Ser Leu Glu Tyr Tyr Lys Lys Ala Leu Asn Lys Ile Arg Asp

37847-517 Sequence Listing_ST25 (003)

180

185

190

Leu Ala Met Ile Arg Asp Val Phe Ile Phe Ser Asp Asp Ile Phe Trp
195 200 205

Cys Lys Glu Asn Ile Glu Thr Leu Leu Ser Lys Lys Tyr Asn Ile Tyr
210 215 220

Tyr Ser Glu Asp Leu Ser Gln Glu Glu Asp Leu Trp Leu Met Ser Leu
225 230 235 240

Ala Asn His His Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala
245 250 255

Tyr Leu Gly Ser Ser Ala Ser Gln Ile Val Ile Tyr Pro Thr Pro Trp
260 265 270

Tyr Asp Ile Thr Pro Lys Asn Thr Tyr Ile Pro Ile Val Asn His Trp
275 280 285

Ile Asn Val Asp Lys His Ser Ser Cys
290 295

<210> 5
<211> 293
<212> PRT
<213> Helicobacter bilis

<400> 5

Met Gly Asp Tyr Lys Ile Val Glu Leu Thr Cys Gly Leu Gly Asn Gln
1 5 10 15

Met Phe Gln Tyr Ala Phe Ala Lys Ala Leu Gln Lys His Leu Gln Val
20 25 30

Pro Val Leu Leu Asp Lys Thr Trp Tyr Asp Thr Gln Asp Asn Ser Thr
35 40 45

Gln Phe Ser Leu Asp Ile Phe Asn Val Asp Leu Glu Tyr Ala Thr Asn
50 55 60

37847-517 Sequence Listing_ST25 (003)

Thr Gln Ile Glu Lys Ala Lys Ala Arg Val Ser Lys Leu Pro Gly Leu
65 70 75 80

Leu Arg Lys Met Phe Gly Leu Lys Lys His Asn Ile Ala Tyr Ser Gln
85 90 95

Ser Phe Asp Phe His Asp Glu Tyr Leu Leu Pro Asn Asp Phe Thr Tyr
100 105 110

Phe Ser Gly Phe Phe Gln Asn Ala Lys Tyr Leu Lys Gly Leu Glu Gln
115 120 125

Glu Leu Lys Ser Ile Phe Tyr Tyr Asp Ser Asn Asn Phe Ser Asn Phe
130 135 140

Gly Lys Gln Arg Leu Glu Leu Ile Leu Gln Ala Lys Asn Ser Ile Phe
145 150 155 160

Ile His Ile Arg Arg Gly Asp Tyr Cys Lys Ile Gly Trp Glu Leu Gly
165 170 175

Met Asp Tyr Tyr Lys Arg Ala Ile Gln Tyr Ile Met Asp Arg Val Glu
180 185 190

Glu Pro Lys Phe Phe Ile Phe Gly Ala Thr Asp Met Ser Phe Thr Glu
195 200 205

Gln Phe Gln Lys Asn Leu Gly Leu Asn Glu Asn Asn Ser Ala Asn Leu
210 215 220

Ser Glu Lys Thr Ile Thr Gln Asp Asn Gln His Glu Asp Met Phe Leu
225 230 235 240

Met Cys Tyr Cys Lys His Ala Ile Leu Ala Asn Ser Ser Tyr Ser Phe
245 250 255

Trp Ser Ala Tyr Leu Asn Asn Asp Ala Asn Asn Ile Val Ile Ala Pro
260 265 270

37847-517 Sequence Listing_ST25 (003)

Thr Pro Trp Leu Leu Asp Asn Asp Asn Ile Ile Cys Asp Asp Trp Ile
 275 280 285

Lys Ile Ser Ser Lys
 290

<210> 6
 <211> 283
 <212> PRT
 <213> Escherichia coli

<400> 6

Met Glu Val Lys Ile Ile Gly Gly Leu Gly Asn Gln Met Phe Gln Tyr
 1 5 10 15

Ala Thr Ala Phe Ala Ile Ala Lys Arg Thr His Gln Asn Leu Thr Val
 20 25 30

Asp Ile Ser Asp Ala Val Lys Tyr Lys Thr His Pro Leu Arg Leu Val
 35 40 45

Glu Leu Ser Cys Ser Ser Glu Phe Val Lys Lys Ala Trp Pro Phe Glu
 50 55 60

Lys Tyr Leu Phe Ser Glu Lys Ile Pro His Phe Met Lys Lys Gly Met
 65 70 75 80

Phe Arg Lys His Tyr Val Glu Lys Ser Leu Glu Tyr Asp Pro Asp Ile
 85 90 95

Asp Thr Lys Ser Ile Asn Lys Lys Ile Val Gly Tyr Phe Gln Thr Glu
 100 105 110

Lys Tyr Phe Lys Glu Phe Arg His Glu Leu Ile Lys Glu Phe Gln Pro
 115 120 125

Lys Thr Lys Phe Asn Ser Tyr Gln Asn Glu Leu Leu Asn Leu Ile Lys
 130 135 140

37847-517 Sequence Listing_ST25 (003)

Glu Asn Asp Thr Cys Ser Leu His Ile Arg Arg Gly Asp Tyr Val Ser
145 150 155 160

Ser Lys Ile Ala Asn Glu Thr His Gly Thr Cys Ser Glu Lys Tyr Phe
165 170 175

Glu Arg Ala Ile Asp Tyr Leu Met Asn Lys Gly Val Ile Asn Lys Lys
180 185 190

Thr Leu Leu Phe Ile Phe Ser Asp Asp Ile Lys Trp Cys Arg Glu Asn
195 200 205

Ile Phe Phe Asn Asn Gln Ile Cys Phe Val Gln Gly Asp Ala Tyr His
210 215 220

Val Glu Leu Asp Met Leu Leu Met Ser Lys Cys Lys Asn Asn Ile Ile
225 230 235 240

Ser Asn Ser Ser Phe Ser Trp Trp Ala Ala Trp Leu Asn Glu Asn Lys
245 250 255

Asn Lys Thr Val Ile Ala Pro Ser Lys Trp Phe Lys Lys Asp Ile Lys
260 265 270

His Asp Ile Ile Pro Glu Ser Trp Val Lys Leu
275 280

<210> 7

<211> 281

<212> PRT

<213> *Vibrio cholerae*

<400> 7

Met Ile Val Met Lys Ile Ser Gly Gly Leu Gly Asn Gln Leu Phe Gln
1 5 10 15

Tyr Ala Val Gly Arg Ala Ile Ala Ile Gln Tyr Gly Val Pro Leu Lys
20 25 30

2015259088 28 May 2018

Leu Asp Gln Phe Asn Ile Asn Ala Asp Ile Ala Asn Glu Asp Glu Ile
50 55 60

Leu Gly Trp Leu Lys Lys Asn Thr Tyr Tyr Ala Glu Lys Gln Arg Thr
85 90 95

Tyr Trp Gln Asn Glu Gln Tyr Phe Ser Gln Ile Arg Ala Val Leu Leu
115 120 125

Gln Ile Lys Ile Gln Gln Thr His Ala Val Ser Ile His Val Arg Arg
145 150 155 160

Tyr Lys Arg Ala Val Asp Tyr Ile Lys Glu Lys Ile Glu Ala Pro Val
180 185 190

Phe Ile Asp Ser Pro Val Phe Ile Glu Asp Thr Gln Thr Glu Ile Asp
210 215 220

Asp Leu Met Leu Met Cys Gln Cys Gln His Asn Ile Val Ala Asn Ser
225 230 235 240

37847-517 Sequence Listing_ST25 (003)

Ser Phe Ser Trp Trp Ala Ala Trp Leu Asn Ser Asn Val Asp Lys Ile
245 250 255

Val Ile Ala Pro Lys Thr Trp Met Ala Glu Asn Pro Lys Gly Tyr Lys
260 265 270

Trp Val Pro Asp Ser Trp Arg Glu Ile
275 280

<210> 8

<211> 289

<212> PRT

<213> Bacteroides fragilis

<400> 8

Met Ile Val Ser Ser Leu Arg Gly Gly Leu Gly Asn Gln Met Phe Ile
1 5 10 15

Tyr Ala Met Val Lys Ala Met Ala Leu Arg Asn Asn Val Pro Phe Ala
20 25 30

Phe Asn Leu Thr Thr Asp Phe Ala Asn Asp Glu Val Tyr Lys Arg Lys
35 40 45

Leu Leu Leu Ser Tyr Phe Ala Leu Asp Leu Pro Glu Asn Lys Lys Leu
50 55 60

Thr Phe Asp Phe Ser Tyr Gly Asn Tyr Tyr Arg Arg Leu Ser Arg Asn
65 70 75 80

Leu Gly Cys His Ile Leu His Pro Ser Tyr Arg Tyr Ile Cys Glu Glu
85 90 95

Arg Pro Pro His Phe Glu Ser Arg Leu Ile Ser Ser Lys Ile Thr Asn
100 105 110

Ala Phe Leu Glu Gly Tyr Trp Gln Ser Glu Lys Tyr Phe Leu Asp Tyr
115 120 125

Lys Gln Glu Ile Lys Glu Asp Phe Val Ile Gln Lys Lys Leu Glu Tyr

37847-517 Sequence Listing_ST25 (003)

130

135

140

Thr Ser Tyr Leu Glu Leu Glu Glu Ile Lys Leu Leu Asp Lys Asn Ala
145 150 155 160

Ile Met Ile Gly Val Arg Arg Tyr Gln Glu Ser Asp Val Ala Pro Gly
165 170 175

Gly Val Leu Glu Asp Asp Tyr Tyr Lys Cys Ala Met Asp Ile Met Ala
180 185 190

Ser Lys Val Thr Ser Pro Val Phe Phe Cys Phe Ser Gln Asp Leu Glu
195 200 205

Trp Val Glu Lys His Leu Ala Gly Lys Tyr Pro Val Arg Leu Ile Ser
210 215 220

Lys Lys Glu Asp Asp Ser Gly Thr Ile Asp Asp Met Phe Leu Met Met
225 230 235 240

His Phe Arg Asn Tyr Ile Ile Ser Asn Ser Ser Phe Tyr Trp Trp Gly
245 250 255

Ala Trp Leu Ser Lys Tyr Asp Asp Lys Leu Val Ile Ala Pro Gly Asn
260 265 270

Phe Ile Asn Lys Asp Ser Val Pro Glu Ser Trp Phe Lys Leu Asn Val
275 280 285

Arg

<210> 9

<211> 290

<212> PRT

<213> Escherichia coli

<400> 9

Met Ser Ile Val Val Ala Arg Leu Ala Gly Gly Leu Gly Asn Gln Met
1 5 10 15

37847-517 Sequence Listing_ST25 (003)

Phe Gln Tyr Ala Lys Gly Tyr Ala Glu Ser Val Glu Arg Asn Ser Tyr
20 25 30

Leu Lys Leu Asp Leu Arg Gly Tyr Lys Asn Tyr Thr Leu His Gly Gly
35 40 45

Phe Arg Leu Asp Lys Leu Asn Ile Asp Asn Thr Phe Val Met Ser Lys
50 55 60

Lys Glu Met Cys Ile Phe Pro Asn Phe Ile Val Arg Ala Ile Asn Lys
65 70 75 80

Phe Pro Lys Leu Ser Leu Cys Ser Lys Arg Phe Glu Ser Glu Gln Tyr
85 90 95

Ser Lys Lys Ile Asn Gly Ser Met Lys Gly Ser Val Glu Phe Ile Gly
100 105 110

Phe Trp Gln Asn Glu Arg Tyr Phe Leu Glu His Lys Glu Lys Leu Arg
115 120 125

Glu Ile Phe Thr Pro Ile Asn Ile Asn Leu Asp Ala Lys Glu Leu Ser
130 135 140

Asp Val Ile Arg Cys Thr Asn Ser Val Ser Val His Ile Arg Arg Gly
145 150 155 160

Asp Tyr Val Ser Asn Val Glu Ala Leu Lys Ile His Gly Leu Cys Thr
165 170 175

Glu Arg Tyr Tyr Ile Asp Ser Ile Arg Tyr Leu Lys Glu Arg Phe Asn
180 185 190

Asn Leu Val Phe Phe Val Phe Ser Asp Asp Ile Glu Trp Cys Lys Lys
195 200 205

Tyr Lys Asn Glu Ile Phe Ser Arg Ser Asp Asp Val Lys Phe Ile Glu
210 215 220

37847-517 Sequence Listing_ST25 (003)

Gly Asn Thr Gln Glu Val Asp Met Trp Leu Met Ser Asn Ala Lys Tyr
225 230 235 240

His Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala Trp Leu Lys
245 250 255

Asn Tyr Asp Leu Gly Ile Thr Ile Ala Pro Thr Pro Trp Phe Glu Arg
260 265 270

Glu Glu Leu Asn Ser Phe Asp Pro Cys Pro Glu Lys Trp Val Arg Ile
275 280 285

Glu Lys
290

<210> 10

<211> 288

<212> PRT

<213> *Prevotella melaninogenica*

<400> 10

Met Lys Ile Val Lys Ile Leu Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Leu Tyr Leu Ser Leu Gln Glu Ser Phe Pro Lys Glu Arg Val
20 25 30

Ala Leu Asp Leu Ser Ser Phe His Gly Tyr His Leu His Asn Gly Phe
35 40 45

Glu Leu Glu Asn Ile Phe Ser Val Thr Ala Gln Lys Ala Ser Ala Ala
50 55 60

Asp Ile Met Arg Ile Ala Tyr Tyr Tyr Pro Asn Tyr Leu Leu Trp Arg
65 70 75 80

Ile Gly Lys Arg Phe Leu Pro Arg Arg Lys Gly Met Cys Leu Glu Ser
85 90 95

37847-517 Sequence Listing_ST25 (003)

Ser Ser Leu Arg Phe Asp Glu Ser Val Leu Arg Gln Glu Gly Asn Arg
 100 105 110

Tyr Phe Asp Gly Tyr Trp Gln Asp Glu Arg Tyr Phe Ala Ala Tyr Arg
 115 120 125

Glu Lys Val Leu Lys Ala Phe Thr Phe Pro Ala Phe Lys Arg Ala Glu
 130 135 140

Asn Leu Ser Leu Leu Glu Lys Leu Asp Glu Asn Ser Ile Ala Leu His
 145 150 155 160

Val Arg Arg Gly Asp Tyr Val Gly Asn Asn Leu Tyr Gln Gly Ile Cys
 165 170 175

Asp Leu Asp Tyr Tyr Arg Thr Ala Ile Glu Lys Met Cys Ala His Val
 180 185 190

Thr Pro Ser Leu Phe Cys Ile Phe Ser Asn Asp Ile Thr Trp Cys Gln
 195 200 205

Gln His Leu Gln Pro Tyr Leu Lys Ala Pro Val Val Tyr Val Thr Trp
 210 215 220

Asn Thr Gly Val Glu Ser Tyr Arg Asp Met Gln Leu Met Ser Cys Cys
 225 230 235 240

Ala His Asn Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala Trp
 245 250 255

Leu Asn Gln Asn Arg Glu Lys Val Val Ile Ala Pro Lys Lys Trp Leu
 260 265 270

Asn Met Glu Glu Cys His Phe Thr Leu Pro Ala Ser Trp Ile Lys Ile
 275 280 285

<210> 11
 <211> 265
 <212> PRT

37847-517 Sequence Listing_ST25 (003)

<213> Clostridium bolteae

<400> 11

Met Phe Gln Tyr Ala Leu Tyr Lys Ala Phe Glu Gln Lys His Ile Asp
 1 5 10 15

Val Tyr Ala Asp Leu Ala Trp Tyr Lys Asn Lys Ser Val Lys Phe Glu
 20 25 30

Leu Tyr Asn Phe Gly Ile Lys Ile Asn Val Ala Ser Glu Lys Asp Ile
 35 40 45

Asn Arg Leu Ser Asp Cys Gln Ala Asp Phe Val Ser Arg Ile Arg Arg
 50 55 60

Lys Ile Phe Gly Lys Lys Lys Ser Phe Val Ser Glu Lys Asn Asp Ser
 65 70 75 80

Cys Tyr Glu Asn Asp Ile Leu Arg Met Asp Asn Val Tyr Leu Ser Gly
 85 90 95

Tyr Trp Gln Thr Glu Lys Tyr Phe Ser Asn Thr Arg Glu Lys Leu Leu
 100 105 110

Glu Asp Tyr Ser Phe Ala Leu Val Asn Ser Gln Val Ser Glu Trp Glu
 115 120 125

Asp Ser Ile Arg Asn Lys Asn Ser Val Ser Ile His Ile Arg Arg Gly
 130 135 140

Asp Tyr Leu Gln Gly Glu Leu Tyr Gly Gly Ile Cys Thr Ser Leu Tyr
 145 150 155 160

Tyr Ala Glu Ala Ile Glu Tyr Ile Lys Met Arg Val Pro Asn Ala Lys
 165 170 175

Phe Phe Val Phe Ser Asp Asp Val Glu Trp Val Lys Gln Gln Glu Asp
 180 185 190

37847-517 Sequence Listing_ST25 (003)

Phe Lys Gly Phe Val Ile Val Asp Arg Asn Glu Tyr Ser Ser Ala Leu
 195 200 205

Ser Asp Met Tyr Leu Met Ser Leu Cys Lys His Asn Ile Ile Ala Asn
 210 215 220

Ser Ser Phe Ser Trp Trp Ala Ala Trp Leu Asn Arg Asn Glu Glu Lys
 225 230 235 240

Ile Val Ile Ala Pro Arg Arg Trp Leu Asn Gly Lys Cys Thr Pro Asp
 245 250 255

Ile Trp Cys Lys Lys Trp Ile Arg Ile
 260 265

<210> 12

<211> 289

<212> PRT

<213> Lachnospiraceae bacterium 3_1_57FAA_CT1

<400> 12

Met Val Ile Val Gln Leu Ser Gly Gly Leu Gly Asn Gln Met Phe Glu
 1 5 10 15

Tyr Ala Leu Tyr Leu Ser Leu Lys Ala Lys Gly Lys Glu Val Lys Ile
 20 25 30

Asp Asp Val Thr Cys Tyr Glu Gly Pro Gly Thr Arg Pro Arg Gln Leu
 35 40 45

Asp Val Phe Gly Ile Thr Tyr Asp Arg Ala Ser Arg Glu Glu Leu Thr
 50 55 60

Glu Met Thr Asp Ala Ser Met Asp Ala Leu Ser Arg Val Arg Arg Lys
 65 70 75 80

Leu Thr Gly Arg Arg Thr Lys Ala Tyr Arg Glu Arg Asp Ile Asn Phe
 85 90 95

Asp Pro Leu Val Met Glu Lys Asp Pro Ala Leu Leu Glu Gly Cys Phe

37847-517 Sequence Listing_ST25 (003)

100

105

110

Gln Ser Asp Lys Tyr Phe Arg Asp Cys Glu Gly Arg Val Arg Glu Ala
115 120 125

Tyr Arg Phe Arg Gly Ile Glu Ser Gly Ala Phe Pro Leu Pro Glu Asp
130 135 140

Tyr Leu Arg Leu Glu Lys Gln Ile Glu Asp Cys Gln Ser Val Ser Val
145 150 155 160

His Ile Arg Arg Gly Asp Tyr Leu Asp Glu Ser His Gly Gly Leu Tyr
165 170 175

Thr Gly Ile Cys Thr Glu Ala Tyr Tyr Lys Glu Ala Phe Ala Arg Met
180 185 190

Glu Arg Leu Val Pro Gly Ala Arg Phe Phe Leu Phe Ser Asn Asp Pro
195 200 205

Glu Trp Thr Arg Glu His Phe Glu Ser Lys Asn Cys Val Leu Val Glu
210 215 220

Gly Ser Thr Glu Asp Thr Gly Tyr Met Asp Leu Tyr Leu Met Ser Arg
225 230 235 240

Cys Arg His Asn Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala
245 250 255

Trp Leu Asn Glu Asn Pro Glu Lys Lys Val Ile Ala Pro Ala Lys Trp
260 265 270

Leu Asn Gly Arg Glu Cys Arg Asp Ile Tyr Thr Glu Arg Met Ile Arg
275 280 285

Leu

<210> 13

37847-517 Sequence Listing_ST25 (003)

<211> 295
 <212> PRT
 <213> Methanosphaerula palustris

<400> 13

Met Ile Ile Val Arg Leu Lys Gly Gly Leu Gly Asn Gln Leu Ser Gln
 1 5 10 15

Tyr Ala Leu Gly Arg Lys Ile Ala His Leu His Asn Thr Glu Leu Lys
 20 25 30

Leu Asp Thr Thr Trp Phe Thr Thr Ile Ser Ser Asp Thr Pro Arg Thr
 35 40 45

Tyr Arg Leu Asn Asn Tyr Asn Ile Ile Gly Thr Ile Ala Ser Ala Lys
 50 55 60

Glu Ile Gln Leu Ile Glu Arg Gly Arg Ala Gln Gly Arg Gly Tyr Leu
 65 70 75 80

Leu Ser Lys Ile Ser Asp Leu Leu Thr Pro Met Tyr Arg Arg Thr Tyr
 85 90 95

Val Arg Glu Arg Met His Thr Phe Asp Lys Ala Ile Leu Thr Val Pro
 100 105 110

Asp Asn Val Tyr Leu Asp Gly Tyr Trp Gln Thr Glu Lys Tyr Phe Lys
 115 120 125

Asp Ile Glu Glu Ile Leu Arg Arg Glu Val Thr Leu Lys Asp Glu Pro
 130 135 140

Asp Ser Ile Asn Leu Glu Met Ala Glu Arg Ile Gln Ala Cys His Ser
 145 150 155 160

Val Ser Leu His Val Arg Arg Gly Asp Tyr Val Ser Asn Pro Thr Thr
 165 170 175

Gln Gln Phe His Gly Cys Cys Ser Ile Asp Tyr Tyr Asn Arg Ala Ile
 180 185 190

37847-517 Sequence Listing_ST25 (003)

28 May 2018
2015259088

Ser Leu Ile Glu Glu Lys Val Asp Asp Pro Ser Phe Phe Ile Phe Ser
195 200 205

Asp Asp Leu Pro Trp Ala Lys Glu Asn Leu Asp Ile Pro Gly Glu Lys
210 215 220

Thr Phe Val Ala His Asn Gly Pro Glu Lys Glu Tyr Cys Asp Leu Trp
225 230 235 240

Leu Met Ser Leu Cys Gln His His Ile Ile Ala Asn Ser Ser Phe Ser
245 250 255

Trp Trp Gly Ala Trp Leu Gly Gln Asp Ala Glu Lys Met Val Ile Ala
260 265 270

Pro Arg Arg Trp Ala Leu Ser Glu Ser Phe Asp Thr Ser Asp Ile Ile
275 280 285

Pro Asp Ser Trp Ile Thr Ile
290 295

<210> 14
<211> 287
<212> PRT
<213> Tannerella sp. CAG:118

<400> 14

Met Val Arg Ile Val Glu Ile Ile Gly Gly Leu Gly Asn Gln Met Phe
1 5 10 15

Gln Tyr Ala Phe Ser Leu Tyr Leu Lys Asn Lys Ser His Ile Trp Asp
20 25 30

Arg Leu Tyr Val Asp Ile Glu Ala Met Lys Thr Tyr Asp Arg His Tyr
35 40 45

Gly Leu Glu Leu Glu Lys Val Phe Asn Leu Ser Leu Cys Pro Ile Ser
50 55 60

37847-517 Sequence Listing_ST25 (003)

Asn Arg Leu His Arg Asn Leu Gln Lys Arg Ser Phe Ala Lys His Phe
65 70 75 80

Val Lys Ser Leu Tyr Glu His Ser Glu Cys Glu Phe Asp Glu Pro Val
85 90 95

Tyr Arg Gly Leu Arg Pro Tyr Arg Tyr Tyr Arg Gly Tyr Trp Gln Asn
100 105 110

Glu Gly Tyr Phe Val Asp Ile Glu Pro Met Ile Arg Glu Ala Phe Gln
115 120 125

Phe Asn Val Asn Ile Leu Ser Lys Lys Thr Lys Ala Ile Ala Ser Lys
130 135 140

Met Arg Arg Glu Leu Ser Val Ser Ile His Val Arg Arg Gly Asp Tyr
145 150 155 160

Glu Asn Leu Pro Glu Ala Lys Ala Met His Gly Gly Ile Cys Ser Leu
165 170 175

Asp Tyr Tyr His Lys Ala Ile Asp Phe Ile Arg Gln Arg Leu Asp Asn
180 185 190

Asn Ile Cys Phe Tyr Leu Phe Ser Asp Asp Ile Asn Trp Val Glu Glu
195 200 205

Asn Leu Gln Leu Glu Asn Arg Cys Ile Ile Asp Trp Asn Gln Gly Glu
210 215 220

Asp Ser Trp Gln Asp Met Tyr Leu Met Ser Cys Cys Arg His His Ile
225 230 235 240

Ile Ala Asn Ser Ser Phe Ser Trp Trp Ala Ala Trp Leu Asn Pro Asn
245 250 255

Lys Asn Lys Ile Val Leu Thr Pro Asn Lys Trp Phe Asn His Thr Asp
260 265 270

37847-517 Sequence Listing_ST25 (003)

Ala Val Gly Ile Val Pro Lys Ser Trp Ile Lys Ile Pro Val Phe
 275 280 285

<210> 15
 <211> 289
 <212> PRT
 <213> Bacteroides caccae

<400> 15

Met Lys Ile Val Lys Ile Leu Gly Gly Leu Gly Asn Gln Met Phe Gln
 1 5 10 15

Tyr Ala Leu Phe Leu Ser Leu Lys Glu Arg Phe Pro His Glu Gln Val
 20 25 30

Met Ile Asp Thr Ser Cys Phe Arg Asn Tyr Pro Leu His Asn Gly Phe
 35 40 45

Glu Val Asp Arg Ile Phe Ala Gln Lys Ala Pro Val Ala Ser Trp Arg
 50 55 60

Asn Ile Leu Lys Val Ala Tyr Pro Tyr Pro Asn Tyr Arg Phe Trp Lys
 65 70 75 80

Ile Gly Lys Tyr Ile Leu Pro Lys Arg Lys Thr Met Cys Val Glu Arg
 85 90 95

Lys Asn Phe Ser Phe Asp Ala Ala Val Leu Thr Arg Lys Gly Asp Cys
 100 105 110

Tyr Tyr Asp Gly Tyr Trp Gln His Glu Glu Tyr Phe Cys Asp Met Lys
 115 120 125

Glu Thr Ile Trp Glu Ala Phe Ser Phe Pro Glu Pro Val Asp Gly Arg
 130 135 140

Asn Lys Glu Ile Gly Ala Leu Leu Gln Ala Ser Asp Ser Ala Ser Leu
 145 150 155 160

37847-517 Sequence Listing_ST25 (003)

His Val Arg Arg Gly Asp Tyr Val Asn His Pro Leu Phe Arg Gly Ile
165 170 175

Cys Asp Leu Asp Tyr Tyr Lys Arg Ala Ile His Tyr Met Glu Glu Arg
180 185 190

Val Asn Pro Gln Leu Tyr Cys Val Phe Ser Asn Asp Met Ala Trp Cys
195 200 205

Glu Ser His Leu Arg Ala Leu Leu Pro Gly Lys Glu Val Val Tyr Val
210 215 220

Asp Trp Asn Lys Gly Ala Glu Ser Tyr Val Asp Met Arg Leu Met Ser
225 230 235 240

Leu Cys Arg His Asn Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Gly
245 250 255

Ala Trp Leu Asn Arg Asn Pro Gln Lys Val Val Val Ala Pro Glu Arg
260 265 270

Trp Met Asn Ser Pro Ile Glu Asp Pro Val Ser Asp Lys Trp Ile Lys
275 280 285

Leu

<210> 16

<211> 290

<212> PRT

<213> Butyrivibrio sp. AE2015

<400> 16

Met Ile Ile Ile Gln Leu Lys Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Leu Tyr Lys Ser Leu Lys Lys Arg Gly Lys Glu Val Lys Ile
20 25 30

Asp Asp Lys Thr Gly Phe Val Asn Asp Lys Leu Arg Ile Pro Val Leu

2015259088 28 May 2018

40

45

Asn Leu Thr Asp Ser Lys Met Asp Leu Phe Ser Arg Ile Arg Arg Lys
65 70 75 80

Leu Thr Gly Arg Lys Thr Phe Arg Ile Asp Glu Glu Ser Gly Lys Phe
85 90 95

Asn Pro Glu Ile Leu Glu Lys Glu Asn Ala Tyr Leu Val Gly Tyr Trp
100 105 110

Gln Cys Asp Lys Tyr Phe Asp Asp Lys Asp Val Val Arg Glu Ile Arg
115 120 125

Glu Ala Phe Glu Lys Lys Pro Gln Glu Leu Met Thr Asp Ala Ser Ser
130 135 140

Trp Ser Thr Leu Gln Gln Ile Glu Cys Cys Glu Ser Val Ser Leu His
145 150 155 160

Val Arg Arg Thr Asp Tyr Val Asp Glu Glu His Ile His Ile His Asn
165 170 175

Ile Cys Thr Glu Lys Tyr Tyr Lys Asn Ala Ile Asp Arg Val Arg Lys
180 185 190

Gln Tyr Pro Ser Ala Val Phe Phe Ile Phe Thr Asp Asp Lys Glu Trp
195 200 205

Cys Arg Asp His Phe Lys Gly Pro Asn Phe Ile Val Val Glu Leu Glu
210 215 220

Glu Gly Asp Gly Thr Asp Ile Ala Glu Met Thr Leu Met Ser Arg Cys
225 230 235 240

Lys His His Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Ala Ala Trp

37847-517 Sequence Listing_ST25 (003)

245 250 255

Leu Asn Asp Ser Pro Glu Lys Ile Val Ile Ala Pro Gln Lys Trp Ile
260 265 270

Asn Asn Arg Asp Met Asp Asp Ile Tyr Thr Glu Arg Met Thr Lys Ile
275 280 285

Ala Leu
290

<210> 17
<211> 282
<212> PRT
<213> Prevotella sp. CAG:891

<400> 17

Met Arg Leu Val Lys Met Ile Gly Gly Leu Gly Asn Gln Met Phe Ile
1 5 10 15

Tyr Ala Phe Tyr Leu Gln Met Arg Lys Arg Phe Ser Asn Val Arg Ile
20 25 30

Asp Leu Thr Asp Met Met His Tyr Asn Val His Tyr Gly Tyr Glu Leu
35 40 45

His Lys Val Phe Gly Leu Pro Arg Thr Glu Phe Cys Met Asn Gln Pro
50 55 60

Leu Lys Lys Val Leu Glu Phe Leu Phe Phe Arg Thr Ile Val Glu Arg
65 70 75 80

Lys Gln His Gly Arg Met Glu Pro Tyr Thr Cys Gln Tyr Val Trp Pro
85 90 95

Leu Val Tyr Phe Lys Gly Phe Tyr Gln Ser Glu Arg Tyr Phe Ser Glu
100 105 110

Val Lys Asp Glu Val Arg Glu Cys Phe Thr Phe Asn Pro Ala Leu Ala
115 120 125

37847-517 Sequence Listing_ST25 (003)

Asn Arg Ser Ser Gln Gln Met Met Glu Gln Ile Gln Asn Asp Pro Gln
130 135 140

Ala Val Ser Ile His Ile Arg Arg Gly Asp Tyr Leu Asn Pro Lys His
145 150 155 160

Tyr Asp Thr Ile Gly Cys Ile Cys Gln Leu Pro Tyr Tyr Lys His Ala
165 170 175

Val Ser Glu Ile Lys Lys Tyr Val Ser Asn Pro His Phe Tyr Val Phe
180 185 190

Ser Glu Asp Leu Asp Trp Val Lys Ala Asn Leu Pro Leu Glu Asn Ala
195 200 205

Gln Tyr Ile Asp Trp Asn Lys Gly Ala Asp Ser Trp Gln Asp Met Met
210 215 220

Leu Met Ser Cys Cys Lys His His Ile Ile Cys Asn Ser Thr Phe Ser
225 230 235 240

Trp Trp Ala Ala Trp Leu Asn Pro Ser Val Glu Lys Thr Val Ile Met
245 250 255

Pro Glu Gln Trp Thr Ser Arg Gln Asp Ser Val Asp Phe Val Ala Ser
260 265 270

Cys Gly Arg Trp Val Arg Val Lys Thr Glu
275 280

<210> 18

<211> 281

<212> PRT

<213> Parabacteroides johnsonii

<400> 18

Met Arg Leu Ile Lys Met Ile Gly Gly Leu Gly Asn Gln Met Phe Ile
1 5 10 15

37847-517 Sequence Listing_ST25 (003)

Tyr Ala Phe Tyr Leu Lys Met Lys His His Tyr Pro Asp Thr Asn Ile
 20 25 30
 Asp Leu Ser Asp Met Val His Tyr Lys Val His Asn Gly Tyr Glu Met
 35 40 45
 Asn Arg Ile Phe Asp Leu Ser Gln Thr Glu Phe Cys Ile Asn Arg Thr
 50 55 60
 Leu Lys Lys Ile Leu Glu Phe Leu Phe Phe Lys Lys Ile Tyr Glu Arg
 65 70 75 80
 Arg Gln Asp Pro Ser Thr Leu Tyr Pro Tyr Glu Lys Arg Tyr Phe Trp
 85 90 95
 Pro Leu Leu Tyr Phe Lys Gly Phe Tyr Gln Ser Glu Arg Phe Phe Phe
 100 105 110
 Asp Ile Lys Asp Asp Val Arg Lys Ala Phe Ser Phe Asn Leu Asn Ile
 115 120 125
 Ala Asn Pro Glu Ser Leu Glu Leu Leu Lys Gln Ile Glu Val Asp Asp
 130 135 140
 Gln Ala Val Ser Ile His Ile Arg Arg Gly Asp Tyr Leu Leu Pro Arg
 145 150 155 160
 His Trp Ala Asn Thr Gly Ser Val Cys Gln Leu Pro Tyr Tyr Lys Asn
 165 170 175
 Ala Ile Ala Glu Met Glu Asn Arg Ile Thr Gly Pro Ser Tyr Tyr Val
 180 185 190
 Phe Ser Asp Asp Ile Ser Trp Val Lys Glu Asn Ile Pro Leu Lys Lys
 195 200 205
 Ala Val Tyr Val Thr Trp Asn Lys Gly Glu Asp Ser Trp Gln Asp Met
 210 215 220

37847-517 Sequence Listing_ST25 (003)

Met Leu Met Ser His Cys Arg His His Ile Ile Cys Asn Ser Thr Phe
225 230 235 240

Ser Trp Trp Gly Ala Trp Leu Asn Pro Arg Lys Glu Lys Ile Val Ile
245 250 255

Ala Pro Cys Arg Trp Phe Gln His Lys Glu Thr Pro Asp Met Tyr Pro
260 265 270

Lys Glu Trp Ile Lys Val Pro Ile Asn
275 280

<210> 19

<211> 290

<212> PRT

<213> Akkermansia muciniphila

<400> 19

Met Arg Leu Phe Gly Gly Leu Gly Asn Gln Leu Phe Gln Tyr Ala Phe
1 5 10 15

Leu Phe Ala Leu Ser Arg Gln Gly Gly Lys Ala Arg Leu Glu Thr Ser
20 25 30

Ser Tyr Glu His Asp Asp Lys Arg Val Cys Glu Leu His His Phe Arg
35 40 45

Val Ser Leu Pro Ile Glu Gly Gly Pro Pro Pro Trp Ala Phe Arg Lys
50 55 60

Ser Arg Ile Pro Ala Cys Leu Arg Ser Leu Phe Ala Ala Pro Lys Tyr
65 70 75 80

Pro His Phe Arg Glu Glu Lys Arg His Gly Phe Asp Pro Gly Leu Ala
85 90 95

Ala Pro Pro Arg Arg His Thr Tyr Phe Lys Gly Tyr Phe Gln Thr Glu
100 105 110

37847-517 Sequence Listing_ST25 (003)

Gln Tyr Phe Leu His Cys Arg Glu Gln Leu Cys Arg Glu Phe Arg Leu
115 120 125

Lys Thr Pro Leu Thr Pro Glu Asn Ala Arg Ile Leu Glu Asp Ile Arg
130 135 140

Ser Cys Cys Ser Ile Ser Leu His Ile Arg Arg Thr Asp Tyr Leu Ser
145 150 155 160

Asn Pro Tyr Leu Ser Pro Pro Pro Leu Glu Tyr Tyr Leu Arg Ser Met
165 170 175

Ala Glu Met Glu Gly Arg Leu Arg Ala Ala Gly Ala Pro Gln Glu Ser
180 185 190

Leu Arg Tyr Phe Ile Phe Ser Asp Asp Ile Glu Trp Ala Arg Gln Asn
195 200 205

Leu Arg Pro Ala Leu Pro His Val His Val Asp Ile Asn Asp Gly Gly
210 215 220

Thr Gly Tyr Phe Asp Leu Glu Leu Met Arg Asn Cys Arg His His Ile
225 230 235 240

Ile Ala Asn Ser Thr Phe Ser Trp Trp Ala Ala Trp Leu Asn Glu His
245 250 255

Ala Glu Lys Ile Val Ile Ala Pro Arg Ile Trp Phe Asn Arg Glu Glu
260 265 270

Gly Asp Arg Tyr His Thr Asp Asp Ala Leu Ile Pro Gly Ser Trp Leu
275 280 285

Arg Ile
290

<210> 20
<211> 298
<212> PRT
<213> Salmonella enterica

37847-517 Sequence Listing_ST25 (003)

<400> 20

Met Tyr Ser Cys Leu Ser Gly Gly Leu Gly Asn Gln Met Phe Gln Tyr
 1 5 10 15

Ala Ala Ala Tyr Ile Leu Lys Gln Tyr Phe Gln Ser Thr Thr Leu Val
 20 25 30

Leu Asp Asp Ser Tyr Tyr Tyr Ser Gln Pro Lys Arg Asp Thr Val Arg
 35 40 45

Ser Leu Glu Leu Asn Gln Phe Asn Ile Ser Tyr Asp Arg Phe Ser Phe
 50 55 60

Ala Asp Glu Lys Glu Lys Ile Lys Leu Leu Arg Lys Phe Lys Arg Asn
 65 70 75 80

Pro Phe Pro Lys Gln Ile Ser Glu Ile Leu Ser Ile Ala Leu Phe Gly
 85 90 95

Lys Tyr Ala Leu Ser Asp Arg Ala Phe Tyr Thr Phe Glu Thr Ile Lys
 100 105 110

Asn Ile Asp Lys Ala Cys Leu Phe Ser Phe Tyr Gln Asp Ala Asp Leu
 115 120 125

Leu Asn Lys Tyr Lys Gln Leu Ile Leu Pro Leu Phe Glu Leu Arg Asp
 130 135 140

Asp Leu Leu Asp Ile Cys Lys Asn Leu Glu Leu Tyr Ser Leu Ile Gln
 145 150 155 160

Arg Ser Asn Asn Thr Thr Ala Leu His Ile Arg Arg Gly Asp Tyr Val
 165 170 175

Thr Asn Gln His Ala Ala Lys Tyr His Gly Val Leu Asp Ile Ser Tyr
 180 185 190

Tyr Asn His Ala Met Glu Tyr Val Glu Arg Glu Arg Gly Lys Gln Asn

37847-517 Sequence Listing_ST25 (003)

195

200

205

Phe Ile Ile Phe Ser Asp Asp Val Arg Trp Ala Gln Lys Ala Phe Leu
210 215 220

Glu Asn Asp Asn Cys Tyr Val Ile Asn Asn Ser Asp Tyr Asp Phe Ser
225 230 235 240

Ala Ile Asp Met Tyr Leu Met Ser Leu Cys Lys Asn Asn Ile Ile Ala
245 250 255

Asn Ser Thr Tyr Ser Trp Trp Gly Ala Trp Leu Asn Lys Tyr Glu Asp
260 265 270

Lys Leu Val Ile Ser Pro Lys Gln Trp Phe Leu Gly Asn Asn Glu Thr
275 280 285

Ser Leu Arg Asn Ala Ser Trp Ile Thr Leu
290 295

<210> 21

<211> 281

<212> PRT

<213> Bacteroides sp. CAG:633

<400> 21

Met Arg Leu Ile Lys Met Thr Gly Gly Leu Gly Asn Gln Met Phe Ile
1 5 10 15

Tyr Ala Phe Tyr Leu Arg Met Lys Lys Arg Tyr Pro Lys Val Arg Ile
20 25 30

Asp Leu Ser Asp Met Val His Tyr His Val His His Gly Tyr Glu Met
35 40 45

His Arg Val Phe Asn Leu Pro His Thr Glu Phe Cys Ile Asn Gln Pro
50 55 60

Leu Lys Lys Val Ile Glu Phe Leu Phe Phe Lys Lys Ile Tyr Glu Arg
65 70 75 80

37847-517 Sequence Listing_ST25 (003)

Lys Gln Asp Pro Asn Ser Leu Arg Ala Phe Glu Lys Lys Tyr Leu Trp
85 90 95

Pro Leu Leu Tyr Phe Lys Gly Phe Tyr Gln Ser Glu Arg Phe Phe Ala
100 105 110

Asp Ile Lys Asp Glu Val Arg Lys Ala Phe Thr Phe Asp Ser Ser Lys
115 120 125

Val Asn Ala Arg Ser Ala Glu Leu Leu Arg Arg Leu Asp Ala Asp Ala
130 135 140

Asn Ala Val Ser Leu His Ile Arg Arg Gly Asp Tyr Leu Gln Pro Gln
145 150 155 160

His Trp Ala Thr Thr Gly Ser Val Cys Gln Leu Pro Tyr Tyr Gln Asn
165 170 175

Ala Ile Ala Glu Met Asn Arg Arg Val Ala Ala Pro Ser Tyr Tyr Val
180 185 190

Phe Ser Asp Asp Ile Ala Trp Val Lys Glu Asn Ile Pro Leu Gln Asn
195 200 205

Ala Val Tyr Ile Asp Trp Asn Lys Gly Glu Glu Ser Trp Gln Asp Met
210 215 220

Met Leu Met Ser His Cys Arg His His Ile Ile Cys Asn Ser Thr Phe
225 230 235 240

Ser Trp Trp Gly Ala Trp Leu Asp Pro His Glu Asp Lys Ile Val Ile
245 250 255

Val Pro Asn Arg Trp Phe Gln His Cys Glu Thr Pro Asn Ile Tyr Pro
260 265 270

Ala Gly Trp Val Lys Val Ala Ile Asn
275 280

37847-517 Sequence Listing_ST25 (003)

28 May 2018
2015259088

<210> 22
<211> 318
<212> PRT
<213> Clostridium sp. CAG:306

<400> 22

Met Glu Lys Ile Lys Ile Val Lys Leu Gln Gly Gly Met Gly Asn Gln
1 5 10 15

Met Phe Gln Tyr Ala Phe Gly Lys Gly Leu Glu Ser Lys Phe Gly Cys
20 25 30

Lys Val Leu Phe Asp Lys Ile Asn Tyr Asp Glu Leu Gln Lys Thr Ile
35 40 45

Ile Asn Asn Thr Gly Lys Asn Ala Glu Gly Ile Cys Val Arg Lys Tyr
50 55 60

Glu Leu Gly Ile Phe Asn Leu Asn Ile Asp Phe Ala Thr Ala Glu Gln
65 70 75 80

Ile Gln Glu Cys Ile Gly Glu Lys Leu Asn Lys Ala Cys Tyr Leu Pro
85 90 95

Gly Phe Ile Arg Lys Ile Phe Asn Leu Ser Lys Asn Lys Thr Val Ser
100 105 110

Asn Arg Ile Phe Glu Lys Lys Tyr Gly Glu Tyr Asp Glu Glu Ile Leu
115 120 125

Lys Asp Tyr Ser Leu Ala Tyr Tyr Asp Gly Tyr Phe Gln Asn Pro Lys
130 135 140

Tyr Phe Glu Asp Ile Ser Asp Lys Ile Lys Lys Glu Phe Thr Leu Pro
145 150 155 160

Glu Ile Lys Asn His Asp Ile Tyr Asn Lys Lys Leu Leu Glu Lys Ile
165 170 175

37847-517 Sequence Listing_ST25 (003)

Thr Gln Phe Glu Asn Ser Val Phe Ile His Val Arg Arg Asp Asp Tyr
180 185 190

Leu Asn Ile Asn Cys Glu Ile Asp Leu Asp Tyr Tyr Gln Lys Ala Val
195 200 205

Lys Tyr Ile Leu Lys His Ile Glu Asn Pro Lys Phe Phe Val Phe Cys
210 215 220

Ala Glu Asp Pro Asp Tyr Ile Lys Asn His Phe Asp Ile Gly Tyr Asp
225 230 235 240

Phe Glu Leu Val Gly Glu Asn Asn Lys Thr Gln Asp Thr Tyr Tyr Glu
245 250 255

Asn Met Arg Leu Met Met Ala Cys Lys His Ala Ile Ile Ala Asn Ser
260 265 270

Ser Tyr Ser Trp Trp Ala Ala Trp Leu Ser Asp Tyr Asp Asn Lys Ile
275 280 285

Val Ile Ala Pro Thr Pro Trp Leu Pro Gly Ile Ser Asn Glu Ile Ile
290 295 300

Cys Lys Asn Trp Ile Gln Ile Lys Arg Gly Ile Ser Asn Glu
305 310 315

<210> 23

<211> 288

<212> PRT

<213> Prevotella sp. oral taxon 306

<400> 23

Met Lys Ile Val Lys Ile Leu Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Leu Tyr Leu Ser Leu Gln Glu Ser Phe Pro Lys Glu Arg Val
20 25 30

37847-517 Sequence Listing_ST25 (003)

Ala	Leu	Asp	Leu	Ser	Cys	Phe	Asn	Gly	Tyr	His	Leu	His	Asn	Gly	Phe	35	40	45	
Glu	Leu	Glu	Arg	Ile	Phe	Ser	Leu	Thr	Ala	Gln	Lys	Ala	Ser	Ala	Ala	50	55	60	
Thr	Ile	Met	Arg	Ile	Ala	Tyr	Tyr	Tyr	Pro	Asn	Tyr	Leu	Leu	Trp	Arg	65	70	75	80
Ile	Gly	Lys	Arg	Leu	Leu	Pro	Arg	Arg	Lys	Thr	Met	Cys	Leu	Glu	Ser	85	90	95	
Ser	Thr	Phe	Arg	Tyr	Asp	Glu	Ser	Val	Leu	Thr	Arg	Glu	Gly	Asn	Arg	100	105	110	
Tyr	Phe	Asp	Gly	Tyr	Trp	Gln	Asp	Glu	Arg	Tyr	Phe	Val	Ala	Cys	Arg	115	120	125	
Glu	Lys	Val	Leu	Lys	Ala	Phe	Thr	Phe	Pro	Ala	Phe	Lys	Arg	Thr	Glu	130	135	140	
Asn	Leu	Ser	Leu	Leu	Arg	Lys	Leu	Asp	Lys	Asn	Ser	Val	Ala	Ile	His	145	150	155	160
Val	Arg	Arg	Gly	Asp	Tyr	Ile	Gly	Asn	Gln	Leu	Tyr	Gln	Gly	Ile	Cys	165	170	175	
Asp	Leu	Asp	Tyr	Tyr	Arg	Ala	Ala	Ile	Asp	Lys	Ile	Ser	Thr	Tyr	Val	180	185	190	
Thr	Pro	Ser	Val	Phe	Cys	Ile	Phe	Ser	Asn	Asp	Ile	Ala	Trp	Cys	Gln	195	200	205	
Thr	His	Leu	Gln	Pro	Tyr	Leu	Lys	Ala	Pro	Val	Val	Tyr	Val	Thr	Trp	210	215	220	
Asn	Thr	Gly	Thr	Glu	Ser	Tyr	Arg	Asp	Met	Gln	Leu	Met	Ser	Cys	Cys	225	230	235	240

37847-517 Sequence Listing_ST25 (003)

Ala His Asn Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala Trp
 245 250 255

Leu Asn Gln Asn Asn Glu Lys Val Val Ile Ala Pro Lys Arg Trp Leu
 260 265 270

Asn Met Asp Asp Cys Gln Phe Pro Leu Pro Ala Ser Trp Val Lys Ile
 275 280 285

<210> 24

<211> 268

<212> PRT

<213> Brachyspira sp. CAG:484

<400> 24

Met Gln Leu Val Lys Leu Met Gly Gly Leu Gly Asn Gln Met Phe Gln
 1 5 10 15

Tyr Ala Phe Ala Lys Ala Leu Gly Asp Lys Asn Ile Leu Phe Tyr Gly
 20 25 30

Asp Tyr Lys Lys His Ser Leu Arg Lys Val Glu Leu Asn Arg Phe Lys
 35 40 45

Cys Lys Ala Val Tyr Ile Pro Arg Glu Leu Phe Lys Tyr Leu Lys Phe
 50 55 60

Val Phe Thr Lys Phe Asp Lys Ile Glu Tyr Met Arg Ser Gly Ile Tyr
 65 70 75 80

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

Trp Gln Thr Glu Lys Tyr Phe Lys Gln Ile Arg Pro Arg Leu Leu Lys
 100 105 110

Asp Phe Thr Pro Arg Lys Lys Leu Asp Arg Glu Asn Ala Gly Ile Ile
 115 120 125

Ser Lys Met Gln Gln Ile Asn Ser Val Ser Val His Ile Arg Arg Thr

37847-517 Sequence Listing_ST25 (003)

130 135 140

Asp Tyr Val Asp Glu Ser His Ile Tyr Gly Asp Thr Asn Leu Asp Tyr
145 150 155 160

Tyr Lys Arg Ala Ile Glu Tyr Ile Ser Ser Lys Ile Glu Asn Pro Glu
165 170 175

Phe Phe Phe Phe Ser Asp Asp Met Ala Tyr Val Lys Glu Lys Phe Ala
180 185 190

Gly Leu Lys Phe Pro His Ser Phe Ile Asp Ile Asn Ser Gly Asn Asn
195 200 205

Ser Tyr Lys Asp Leu Ile Leu Met Lys Asn Cys Lys His Asn Ile Ile
210 215 220

Ala Asn Ser Thr Phe Ser Trp Trp Gly Ala Trp Leu Asn Glu Asn Glu
225 230 235 240

Glu Lys Ile Val Ile Ala Pro Ala Lys Trp Phe Val Thr Gly Glu Asn
245 250 255

Asp Lys Asp Ile Val Pro Asp Glu Trp Ile Lys Leu
260 265

<210> 25
<211> 292
<212> PRT
<213> *Thalassospira profundimaris*

<400> 25

Met Val Ile Val Lys Leu Leu Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Thr Gly Arg Ala Val Ala Ser Arg Leu Asp Val Glu Leu Leu
20 25 30

Leu Asp Val Ser Ala Phe Ala His Tyr Asp Leu Arg Arg Tyr Glu Leu
35 40 45

37847-517 Sequence Listing_ST25 (003)

Asp Asp Trp Asn Ile Thr Ala Arg Leu Ala Thr Ser Glu Glu Leu Ala
50 55 60

Arg Ser Gly Val Thr Ala Ala Pro Pro Ser Phe Phe Asp Arg Ile Ala
65 70 75 80

Arg Phe Leu Arg Ile Asp Leu Pro Val Asn Cys Phe Arg Glu Ala Ser
85 90 95

Phe Thr Tyr Asp Pro Arg Ile Leu Glu Val Ser Ser Pro Val Tyr Leu
100 105 110

Asp Gly Tyr Trp Gln Ser Glu Arg Tyr Phe Leu Asp Ile Glu Lys Lys
115 120 125

Leu Arg Gln Glu Phe Gln Leu Lys Ala Ser Ile Asp Ala Asn Asn His
130 135 140

Ser Phe Lys Lys Lys Ile Asp Gly Leu Gly Lys Gln Ala Val Ser Leu
145 150 155 160

His Val Arg Arg Gly Asp Tyr Val Thr Asn Pro Gln Thr Ala Ser Tyr
165 170 175

His Gly Val Cys Ser Leu Asp Tyr Tyr Arg Ala Ala Val Asp Tyr Ile
180 185 190

Ala Glu His Val Ser Asp Pro Cys Phe Phe Val Phe Ser Asp Asp Leu
195 200 205

Glu Trp Val Gln Thr Asn Leu Asn Ile Lys Gln Pro Ile Val Leu Val
210 215 220

Asp Ala Asn Gly Pro Asp Asn Gly Ala Ala Asp Met Ala Leu Met Met
225 230 235 240

Ala Cys Arg His His Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Gly
245 250 255

37847-517 Sequence Listing_ST25 (003)

Ser Trp Leu Asn Pro Leu Asn Asp Lys Ile Ile Val Ala Pro Lys Lys
260 265 270

Trp Phe Gly Arg Ala Asn His Asp Thr Thr Asp Leu Val Pro Asp Ser
275 280 285

Trp Val Arg Leu
290

<210> 26

<211> 308

<212> PRT

<213> Acetobacter sp. CAG:267

<400> 26

Met Ala Val Ser Pro Gln Glu Ser Lys Tyr Ser Ala His Val Ser Pro
1 5 10 15

Asp Lys Pro Leu Arg Ile Val Arg Leu Gly Gly Gly Leu Gly Asn Gln
20 25 30

Met Phe Gln Tyr Ala Phe Gly Leu Ala Ala Gly Asp Val Leu Trp Asp
35 40 45

Asn Thr Ser Phe Leu Thr Asn His Tyr Arg Ser Phe Asp Leu Gly Leu
50 55 60

Tyr Asn Ile Ser Gly Asp Phe Ala Ser Asn Glu Gln Ile Lys Lys Cys
65 70 75 80

Lys Asn Glu Ile Arg Phe Lys Asn Ile Leu Pro Arg Ser Ile Arg Lys
85 90 95

Lys Phe Asn Leu Gly Lys Phe Ile Tyr Leu Lys Thr Asn Arg Val Cys
100 105 110

Glu Arg Gln Ile Asn Arg Tyr Glu Pro Glu Leu Leu Ser Lys Asp Gly
115 120 125

37847-517 Sequence Listing_ST25 (003)

Asp Val Tyr Tyr Asp Gly Val Phe Gln Thr Glu Lys Tyr Phe Lys Pro
130 135 140

Leu Arg Glu Arg Leu Leu His Asp Phe Thr Leu Thr Lys Pro Leu Asp
145 150 155 160

Ala Ala Asn Leu Asp Met Leu Ala Lys Ile Arg Ala Ala Asp Ala Val
165 170 175

Ala Val His Ile Arg Arg Gly Asp Tyr Leu Asn Pro Arg Ser Pro Phe
180 185 190

Thr Tyr Leu Asp Lys Asp Tyr Phe Leu Asn Ala Met Asp Tyr Ile Gly
195 200 205

Lys Arg Val Asp Lys Pro His Phe Phe Ile Phe Ser Ser Asp Thr Asp
210 215 220

Trp Val Arg Thr Asn Ile Gln Thr Ala Tyr Pro Gln Thr Ile Val Glu
225 230 235 240

Ile Asn Asp Glu Lys His Gly Tyr Phe Asp Leu Glu Leu Met Arg Asn
245 250 255

Cys Arg His Asn Ile Ile Ala Asn Ser Thr Phe Ser Trp Trp Gly Ala
260 265 270

Trp Leu Asn Thr Asn Pro Asp Lys Ile Val Val Ala Pro Lys Gln Trp
275 280 285

Phe Arg Pro Asp Ala Ala Glu Tyr Ser Gly Asp Ile Val Pro Asn Asp
290 295 300

Trp Ile Lys Leu
305

<210> 27
<211> 291
<212> PRT

37847-517 Sequence Listing_ST25 (003)

<213> Dysgonomonas mossii

<400> 27

Met Val Thr Val Leu Leu Ser Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Ala Ala Lys Ser Leu Ala Ile Arg Leu Asn Thr Ala Leu Ser
20 25 30

Val Asp Leu Tyr Thr Phe Ser Lys Lys Thr Gln Ala Thr Val Arg Pro
35 40 45

Tyr Glu Leu Gly Ile Phe Asn Ile Glu Asp Val Val Glu Thr Ser Ser
50 55 60

Leu Lys Ala Lys Ala Val Ile Lys Ala Arg Pro Phe Ile Gln Arg His
65 70 75 80

Arg Ser Phe Phe Gln Arg Phe Gly Val Phe Thr Asp Thr Tyr Ala Ile
85 90 95

Leu Tyr Gln Pro Thr Phe Glu Ala Leu Thr Gly Gly Val Ile Met Ser
100 105 110

Gly Tyr Phe Gln Asn Glu Ser Tyr Phe Lys Asn Ile Ser Glu Leu Leu
115 120 125

Arg Lys Asp Phe Ser Phe Lys Tyr Pro Leu Ile Gly Glu Asn Lys Asp
130 135 140

Val Ala Gly Gln Ile Ser Glu Asn Gln Ser Val Ala Val His Ile Arg
145 150 155 160

Arg Gly Asp Tyr Leu Asn Lys Asn Ser Gln Ser Asn Phe Ala Ile Leu
165 170 175

Glu Lys Asp Tyr Tyr Glu Lys Ala Ile Asn Tyr Ile Ser Ala His Val
180 185 190

37847-517 Sequence Listing_ST25 (003)

Lys Asn Pro Glu Phe Tyr Val Phe Ser Glu Asp Phe Asp Trp Ile Lys
195 200 205

Asp Asn Leu Asn Phe Lys Glu Phe Pro Val Thr Phe Ile Asp Trp Asn
210 215 220

Lys Gly Lys Asp Ser Tyr Ile Asp Met Gln Leu Met Ser Leu Cys Lys
225 230 235 240

His Asn Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Ser Ala Trp Leu
245 250 255

Asn Asn Ser Glu Glu Arg Lys Ile Val Ala Pro Glu Arg Trp Phe Val
260 265 270

Asp Glu Gln Lys Asn Glu Leu Leu Asp Cys Phe Tyr Pro Gln Gly Trp
275 280 285

Ile Lys Ile
290

<210> 28
<211> 287
<212> PRT
<213> Clostridium sp. KLE 1755

<400> 28

Met Ile Ile Ile Glu Ile Ser Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Leu Gly Gln Lys Phe Ile Ser Met Gly Lys Glu Val Lys Tyr
20 25 30

Asp Leu Ser Phe Tyr Asn Asp Arg Val Gln Thr Leu Arg Gln Phe Glu
35 40 45

Leu Asp Ile Phe Asp Leu Asp Cys Pro Val Ala Ser Asn Ser Glu Leu
50 55 60

Ser Arg Phe Gly Lys Gly Asn Ser Leu Lys Ser Arg Leu Lys Gln Lys

2015259088 28 May 2018

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

275

280

285

<210> 29
<211> 301
<212> PRT
<213> Gillisia limnaea

<400> 29

Met Phe Ile Ser Lys Asn Thr Val Ile Ile Lys Leu Val Gly Gly Leu
1 5 10 15

Gly Asn Gln Met Phe Gln Phe Ala Ile Ala Lys Ile Ile Ala Glu Lys
20 25 30

Glu Lys Ser Glu Val Leu Val Asp Ile Thr Phe Tyr Thr Glu Leu Thr
35 40 45

Glu Asn Thr Lys Lys Phe Pro Arg His Phe Ser Leu Gly Ile Phe Asn
50 55 60

Ser Ser Phe Ala Ile Ala Ser Lys Lys Glu Ile Asp Tyr Phe Thr Lys
65 70 75 80

Leu Ser Asn Phe Asn Lys Phe Lys Lys Lys Leu Gly Leu Asn Tyr Pro
85 90 95

Thr Ile Phe His Glu Ser Ser Phe Asn Phe Lys Ala Gln Val Leu Glu
100 105 110

Leu Lys Ala Pro Ile Tyr Leu Asn Gly Tyr Phe Gln Ser Phe Arg Tyr
115 120 125

Phe Leu Gly Lys Glu Tyr Val Ile Arg Lys Ile Phe Lys Phe Pro Asp
130 135 140

Glu Ala Leu Asp Lys Asp Asn Asp Asn Ile Lys Arg Lys Ile Ile Gly
145 150 155 160

Lys Thr Ser Val Ser Leu His Ile Arg Arg Gly Asp Tyr Val Asn Asn
165 170 175

37847-517 Sequence Listing_ST25 (003)

28 May 2018
2015259088

Lys Lys Thr Gln Gln Phe His Gly Asn Cys Thr Ile Asp Tyr Tyr Gln
180 185 190

Ser Ala Ile Ala Tyr Leu Ser Ser Lys Leu Thr Asp Phe Asn Leu Ile
195 200 205

Phe Phe Ser Asp Asp Ile His Trp Val Arg Gln Gln Phe Lys Asn Ile
210 215 220

Ser Asn Gln Lys Ile Tyr Val Ser Gly Asn Leu Asn His Asn Ser Trp
225 230 235 240

Lys Asp Met Tyr Leu Met Ser Leu Cys Asp His Asn Ile Ile Ala Asn
245 250 255

Ser Ser Phe Ser Trp Trp Gly Ala Trp Leu Asn Lys Asn Pro Glu Lys
260 265 270

Ile Ile Ile Ala Pro Lys Arg Trp Phe Ala Asp Thr Glu Gln Asp Lys
275 280 285

Asn Ser Ile Asp Leu Ile Pro Ser Glu Trp Tyr Arg Ile
290 295 300

<210> 30
<211> 288
<212> PRT
<213> Methylobacterium mobilis

<400> 30

Met Leu Val Ser Arg Ile Ile Gly Gly Leu Gly Asn Gln Met Phe Glu
1 5 10 15

Tyr Ala Ala Ala Arg Ala Ala Ser Leu Arg Ile Ser Val Gln Leu Lys
20 25 30

Leu Asp Leu Ser Gly Phe Glu Thr Tyr Asp Leu His Ala Tyr Gly Leu
35 40 45

37847-517 Sequence Listing_ST25 (003)

Asn Asn Phe Asn Ile Val Glu Asp Val Ala Lys Lys Asp Asp Tyr Phe
 50 55 60

Ile Gly Ala Pro Glu Ser Leu Leu Lys Lys Ile Lys Lys Tyr Leu Arg
 65 70 75 80

Gly Leu Ile Gln Leu Glu Ser Phe Arg Glu Ser Asp Leu Ser Phe Asp
 85 90 95

Ser Lys Val Leu Glu Leu Asn Asp Asn Thr Tyr Leu Asp Gly Tyr Trp
 100 105 110

Gln Cys Glu Arg Tyr Phe Ile Asp Phe Asp Lys Gln Ile Arg Gln Asp
 115 120 125

Phe Ser Phe Lys Phe Ala Pro Asp Ala Leu Asn Gln Arg Tyr Leu Glu
 130 135 140

Leu Ile Asp Ser Val Asn Ala Val Ser Val His Ile Arg Arg Gly Asp
 145 150 155 160

Tyr Val Ser Asn Ser Thr Thr Asn Glu Ile His Gly Val Cys Asp Leu
 165 170 175

Asp Tyr Tyr Gln Arg Ala Ala Glu Phe Met Arg Ala Arg Ile Gly Pro
 180 185 190

Glu Asn Leu His Phe Phe Val Phe Ser Asp Asp Thr Asp Trp Val Lys
 195 200 205

Glu Asn Ile Ser Phe Gly Ser Asp Thr Thr Phe Ile Ser His Asn Asp
 210 215 220

Ala Ala Lys Asn Tyr Glu Asp Met Arg Leu Met Ser Ala Cys Lys His
 225 230 235 240

His Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Ala Ala Trp Leu Asn
 245 250 255

37847-517 Sequence Listing_ST25 (003)

Pro Ser Lys Gln Lys Val Val Ile Ala Pro Arg Gln Trp Phe Lys Ser
260 265 270

Thr Leu Leu Asn Ser Asp Asp Ile Val Pro Ala Ser Trp Val Arg Leu
275 280 285

<210> 31

<211> 287

<212> PRT

<213> Runella slithyformis

<400> 31

Met Ile Ile Val Lys Leu Ser Gly Gly Leu Gly Asn Gln Leu Phe Gln
1 5 10 15

Tyr Ala Phe Gly Arg His Leu Ala Thr Val Asn Gln Lys Glu Leu Lys
20 25 30

Leu Asp Thr Ser Ala Leu Thr Lys Thr Ser Asp Trp Thr Asn Arg Ser
35 40 45

Tyr Ala Leu Asp Ala Phe Asn Ile Arg Ala Gln Glu Ala Thr Pro Glu
50 55 60

Glu Ile Lys Ala Leu Ala Gly Lys Pro Asn Arg Leu Leu Gln Arg Val
65 70 75 80

Gly Arg Lys Val Gly Ile Thr Pro Ile Gln Tyr Phe Gln Glu Pro His
85 90 95

Phe His Phe Tyr Ser Ser Ala Leu Ser Ile Lys Ser Ser His Tyr Leu
100 105 110

Glu Gly Tyr Trp Gln Ser Glu Lys Tyr Phe Glu Ala Ile Thr Pro Ile
115 120 125

Leu Arg Glu Glu Phe Ala Phe Thr Ile Ser Pro Ser Thr His Ala Gln
130 135 140

37847-517 Sequence Listing_ST25 (003)

Thr Ile Lys Glu Lys Ile Ser Asn Gly Thr Ser Val Ser Ile His Leu
145 150 155 160

Arg Arg Gly Asp Tyr Val Lys Thr Ser Lys Ala Asn Arg Tyr Leu Arg
165 170 175

Pro Leu Thr Met Asp Tyr Tyr Gln Lys Ala Ile Asp Tyr Ile Asn Gln
180 185 190

Arg Val Lys Asn Pro Asn Phe Phe Leu Phe Ser Asp Asp Ile Lys Trp
195 200 205

Ala Lys Ser Gln Val Thr Phe Pro Pro Thr Thr His Phe Ser Thr Gly
210 215 220

Thr Ser Ala His Glu Asp Leu Trp Leu Met Thr His Cys Arg His His
225 230 235 240

Ile Ile Ala Asn Ser Thr Phe Ser Trp Trp Gly Ala Trp Leu Asn Gln
245 250 255

Gln Pro Asp Lys Ile Val Ile Ala Pro Gln Lys Trp Phe Ser Thr Glu
260 265 270

Arg Phe Asp Thr Lys Asp Leu Leu Pro Glu Pro Trp Ile Gln Leu
275 280 285

<210> 32

<211> 288

<212> PRT

<213> Pseudoalteromonas haloplanktis

<400> 32

Met Ile Lys Val Lys Ala Ile Gly Gly Leu Gly Asn Gln Leu Phe Gln
1 5 10 15

Tyr Ala Thr Ala Arg Ala Ile Ala Glu Lys Arg Gly Asp Gly Val Val
20 25 30

Val Asp Met Ser Asp Phe Ser Ser Tyr Lys Thr His Pro Phe Cys Leu

37847-517 Sequence Listing_ST25 (003)

35	40	45																
Asn	Lys	Phe	Arg	Cys	Lys	Ala	Thr	Tyr	Glu	Ser	Lys	Pro	Lys	Leu	Ile			
50						55					60							
Asn	Lys	Leu	Leu	Ser	Asn	Glu	Lys	Ile	Arg	Asn	Leu	Leu	Gln	Lys	Leu			
65					70					75					80			
Gly	Phe	Ile	Lys	Lys	Tyr	Tyr	Phe	Glu	Thr	Gln	Leu	Pro	Phe	Asn	Glu			
				85					90					95				
Asp	Val	Leu	Leu	Asn	Asn	Ser	Ile	Asn	Tyr	Leu	Thr	Gly	Tyr	Phe	Gln			
		100						105					110					
Ser	Glu	Lys	Tyr	Phe	Leu	Ser	Ile	Arg	Glu	Cys	Leu	Leu	Asp	Glu	Leu			
		115					120					125						
Thr	Leu	Ile	Glu	Asp	Leu	Asn	Ile	Ala	Glu	Thr	Ala	Val	Ser	Lys	Ala			
	130					135					140							
Ile	Lys	Asn	Ala	Lys	Asn	Ser	Ile	Ser	Ile	His	Ile	Arg	Arg	Gly	Asp			
145					150					155					160			
Tyr	Val	Ser	Asn	Glu	Gly	Ala	Asn	Lys	Thr	His	Gly	Val	Cys	Asp	Ser			
				165					170					175				
Asp	Tyr	Phe	Lys	Lys	Ala	Leu	Asn	Tyr	Phe	Ser	Glu	Arg	Lys	Leu	Leu			
		180						185					190					
Asp	Glu	His	Thr	Glu	Leu	Phe	Ile	Phe	Ser	Asp	Asp	Ile	Glu	Trp	Cys			
	195						200					205						
Arg	Asn	Asn	Leu	Ser	Phe	Asp	Tyr	Lys	Met	Asn	Phe	Val	Asp	Gly	Ser			
	210					215					220							
Ser	Glu	Arg	Pro	Glu	Val	Asp	Met	Val	Leu	Met	Ser	Gln	Cys	Lys	His			
225					230					235					240			
Gln	Val	Ile	Ser	Asn	Ser	Thr	Phe	Ser	Trp	Trp	Gly	Ala	Trp	Leu	Asn			

37847-517 Sequence Listing_ST25 (003)

245 250 255

Lys Asn Asp Glu Lys Val Val Val Ala Pro Lys Glu Trp Phe Lys Ser
260 265 270

Thr Asp Leu Asp Ser Thr Asp Ile Val Pro Asn Gln Trp Ile Lys Leu
275 280 285

<210> 33

<211> 277

<212> PRT

<213> Unknown

<220>

<223> uncultured bacterium

<400> 33

Met Leu Thr Leu Lys Leu Lys Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Ala Ser His Asn Leu Ala Lys Asn Lys Lys Thr Lys Ile Asn
20 25 30

Phe Asp Leu Ser Phe Phe Ser Asp Ile Glu Val Arg Asp Ile Lys Arg
35 40 45

Asp Tyr Leu Leu Asp Lys Phe Asn Ile Ser Ala Asp Ile Ser Phe Asp
50 55 60

Gln Lys Asn Ser Ile Ser Gly Phe Arg Lys Phe Leu Val Lys Val Ile
65 70 75 80

Ser Lys Phe Phe Gly Glu Val Phe Tyr Tyr Arg Leu Lys Phe Leu Ser
85 90 95

Ser Lys Tyr Leu Asp Gly Tyr Phe Gln Ser Glu Lys Tyr Phe Lys Asn
100 105 110

Val Glu Glu Asp Ile Arg Lys Asp Phe Thr Leu Lys Asp Glu Met Gly
115 120 125

37847-517 Sequence Listing_ST25 (003)

Val Glu Ala Lys Lys Ile Glu Gln Gln Ile Val Asn Ser Lys Asn Ser
130 135 140

Val Ser Leu His Ile Arg Arg Gly Asp Tyr Val Asp Asp Leu Lys Thr
145 150 155 160

Asn Ile Tyr His Gly Val Cys Asn Leu Asp Tyr Tyr Lys Arg Ser Ile
165 170 175

Lys Tyr Leu Lys Glu Asn Phe Gly Glu Ile Asn Ile Phe Val Phe Ser
180 185 190

Asp Asp Ile Ala Trp Val Lys Glu Asn Leu Ala Phe Glu Asn Leu Gln
195 200 205

Phe Val Ser Arg Pro Asp Ile Lys Asp Tyr Glu Glu Leu Met Leu Met
210 215 220

Ser Lys Cys Glu His Asn Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp
225 230 235 240

Gly Ala Trp Leu Asn Glu Asn Lys Asn Lys Ile Ile Ile Ala Pro Lys
245 250 255

Glu Trp Phe Gln Lys Phe Asn Ile Asn Glu Lys His Ile Val Pro Lys
260 265 270

Ser Trp Ile Arg Leu
275

<210> 34

<211> 289

<212> PRT

<213> Clostridium sp. KLE 1755

<400> 34

Met Val Ile Val Gln Leu Ser Gly Gly Leu Gly Asn Gln Met Phe Glu
1 5 10 15

37847-517 Sequence Listing_ST25 (003)

Tyr Ala Leu Tyr Leu Ser Leu Lys Ala Lys Gly Lys Val Val Lys Ile
 20 25 30
 Asp Asp Ile Thr Cys Tyr Glu Gly Pro Gly Thr Arg Pro Lys Gln Leu
 35 40 45
 Asp Val Phe Gly Val Ser Tyr Glu Arg Ala Thr Lys Gln Glu Leu Thr
 50 55 60
 Glu Met Thr Asp Ser Ser Leu Asp Pro Val Ser Arg Ile Arg Arg Lys
 65 70 75 80
 Leu Thr Gly Arg Lys Thr Lys Ala Tyr Arg Glu Lys Asp Ile Asn Phe
 85 90 95
 Asp Pro Gln Val Met Glu Arg Asp Pro Ala Leu Leu Glu Gly Cys Phe
 100 105 110
 Gln Ser Glu Lys Tyr Phe Gln Asp Cys Arg Glu Gln Val Arg Glu Ala
 115 120 125
 Tyr Arg Phe Arg Gly Ile Glu Ser Gly Ala Tyr Pro Leu Pro Glu Ala
 130 135 140
 Tyr Arg Arg Leu Glu Lys Glu Ile Ala Asp Cys Lys Ser Val Ser Val
 145 150 155 160
 His Ile Arg Arg Gly Asp Tyr Leu Glu Glu Ser His Gly Gly Leu Tyr
 165 170 175
 Thr Gly Ile Cys Thr Glu Gln Tyr Tyr Gln Glu Ala Phe Ala Arg Met
 180 185 190
 Glu Lys Glu Val Pro Gly Ala Lys Phe Phe Leu Phe Ser Asn Asp Pro
 195 200 205
 Asp Trp Thr Arg Glu His Phe Lys Gly Glu Asn Arg Ile Leu Val Glu
 210 215 220

37847-517 Sequence Listing_ST25 (003)

Gly Ser Thr Glu Asp Thr Gly Tyr Leu Asp Leu Tyr Leu Met Ser Lys
225 230 235 240

Cys Lys His Asn Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala
245 250 255

Trp Leu Asn Asp Asn Pro Glu Lys Lys Val Thr Ala Pro Ala Arg Trp
260 265 270

Leu Asn Gly Arg Glu Cys Arg Asp Ile Tyr Thr Glu Arg Met Ile Arg
275 280 285

Ile

<210> 35

<211> 281

<212> PRT

<213> Francisella philomiragia

<400> 35

Met Lys Ile Ile Lys Ile Gln Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Phe Tyr Lys Ser Leu Lys Asn Asn Cys Ile Asp Cys Tyr Val
20 25 30

Asp Ile Lys Asn Tyr Asp Thr Tyr Lys Leu His Tyr Gly Phe Glu Leu
35 40 45

Asn Arg Ile Phe Lys Asn Ile Asp Leu Ser Phe Ala Arg Lys Tyr His
50 55 60

Lys Lys Glu Val Leu Gly Lys Leu Phe Ser Ile Ile Pro Ser Lys Phe
65 70 75 80

Ile Val Lys Phe Asn Lys Asn Tyr Ile Leu Gln Lys Asn Phe Ala Phe
85 90 95

Asp Lys Ala Tyr Phe Glu Ile Asp Asn Cys Tyr Leu Asp Gly Tyr Trp

37847-517 Sequence Listing_ST25 (003)

100

105

110

Gln Ser Glu Lys Tyr Phe Lys Lys Ile Thr Lys Asp Ile Tyr Asp Ala
115 120 125

Phe Thr Phe Glu Pro Leu Asp Ser Ile Asn Phe Glu Phe Leu Lys Asn
130 135 140

Ile Gln Asp Tyr Asn Leu Val Ser Ile His Val Arg Arg Gly Asp Tyr
145 150 155 160

Val Asn His Pro Leu His Gly Gly Ile Cys Asp Leu Glu Tyr Tyr Asn
165 170 175

Lys Ala Ile Ser Phe Ile Arg Ser Lys Val Ala Asn Val His Phe Leu
180 185 190

Val Phe Ser Asn Asp Ile Leu Trp Cys Lys Asp Asn Leu Lys Leu Asp
195 200 205

Arg Val Thr Tyr Ile Asp His Asn Arg Trp Met Asp Ser Tyr Lys Asp
210 215 220

Met His Leu Met Ser Leu Cys Lys His Asn Ile Ile Ala Asn Ser Ser
225 230 235 240

Phe Ser Trp Trp Gly Ala Trp Leu Asn Gln Asn Asp Asp Lys Ile Val
245 250 255

Ile Ala Pro Ser Lys Trp Phe Asn Asp Asp Lys Ile Asn Gln Lys Asp
260 265 270

Ile Cys Pro Asn Ser Trp Val Arg Ile
275 280

<210> 36

<211> 288

<212> PRT

<213> Pseudomonas fluorescens

37847-517 Sequence Listing_ST25 (003)

<400> 36

Met Val Ile Ala His Leu Ile Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Ala Ala Arg Ala Leu Ser Ser Ala Lys Lys Glu Pro Leu Leu
20 25 30

Leu Asp Thr Ser Ser Phe Glu Ser Tyr Thr Leu His Gln Gly Phe Glu
35 40 45

Leu Ser Lys Leu Phe Ala Gly Glu Met Cys Ile Ala Arg Asp Lys Asp
50 55 60

Ile Asn His Val Leu Ser Trp Gln Ala Phe Pro Arg Ile Arg Asn Phe
65 70 75 80

Leu His Arg Pro Lys Leu Ala Phe Leu Arg Lys Ala Ser Leu Ile Ile
85 90 95

Glu Pro Ser Phe His Tyr Trp Asn Gly Ile Gln Lys Ala Pro Ala Asp
100 105 110

Cys Tyr Leu Met Gly Tyr Trp Gln Ser Glu Arg Tyr Phe Gln Asp Ala
115 120 125

Ala Glu Glu Ile Arg Lys Asp Phe Thr Phe Lys Leu Asn Met Ser Pro
130 135 140

Gln Asn Ile Ala Thr Ala Asp Gln Ile Leu Asn Thr Asn Ala Ile Ser
145 150 155 160

Leu His Val Arg Arg Gly Asp Tyr Val Asn Asn Ser Val Tyr Ala Ala
165 170 175

Cys Thr Val Glu Tyr Tyr Gln Ala Ala Ile Gln Leu Leu Ser Lys Arg
180 185 190

Val Asp Ala Pro Thr Phe Phe Val Phe Ser Asp Asp Ile Asp Trp Val
195 200 205

37847-517 Sequence Listing_ST25 (003)

Lys Asn Asn Leu Asn Ile Gly Phe Pro His Cys Tyr Val Asn His Asn
210 215 220

Lys Gly Ser Glu Ser Tyr Asn Asp Met Arg Leu Met Ser Met Cys Gln
225 230 235 240

His Asn Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala Trp Leu
245 250 255

Asn Ser Asn Ala Asp Lys Ile Val Val Ala Pro Lys Gln Trp Phe Ile
260 265 270

Asn Asn Thr Asn Val Asn Asp Leu Phe Pro Pro Ala Trp Val Thr Leu
275 280 285

<210> 37

<211> 293

<212> PRT

<213> Herbaspirillum sp. YR522

<400> 37

Met Ile Ala Thr Arg Leu Ile Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Ala Gly Arg Ala Leu Ala Leu Arg Val Gly Ser Pro Leu Leu
20 25 30

Leu Asp Val Ser Gly Phe Ala Asn Tyr Glu Leu Arg Arg Tyr Glu Leu
35 40 45

Asp Gly Phe Arg Ile Asp Ala Thr Ala Ala Ser Ala Gln Gln Leu Ala
50 55 60

Arg Leu Gly Val Asn Ala Thr Pro Gly Thr Ser Leu Leu Ala Arg Val
65 70 75 80

Leu Arg Lys Val Trp Pro Gln Pro Ala Asp Arg Ile Leu Arg Glu Ala
85 90 95

37847-517 Sequence Listing_ST25 (003)

Ser Phe Thr Tyr Asp Ala Arg Ile Glu Gln Ala Ser Ala Pro Val Tyr
 100 105 110

Leu Asp Gly Tyr Trp Gln Ser Glu Arg Tyr Phe Ala Arg Ile Arg Gln
 115 120 125

His Leu Leu Asp Glu Phe Thr Leu Lys Gly Asp Trp Gly Ser Asp Asn
 130 135 140

Ala Ala Met Ala Ala Gln Ile Ala Thr Ala Gly Ala Gly Ala Val Ser
 145 150 155 160

Leu His Val Arg Arg Gly Asp Tyr Val Ser Asn Ala His Thr Ala Gln
 165 170 175

Tyr His Gly Val Cys Ser Leu Asp Tyr Tyr Arg Asp Ala Val Ala His
 180 185 190

Ile Gly Gly Arg Val Glu Ala Pro His Phe Phe Val Phe Ser Asp Asp
 195 200 205

His Glu Trp Val Arg Glu Asn Leu Gln Ile Gly His Pro Ala Thr Phe
 210 215 220

Val Gln Ile Asn Ser Ala Asp His Gly Ile Tyr Asp Met Met Leu Met
 225 230 235 240

Lys Ser Cys Arg His His Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp
 245 250 255

Gly Ala Trp Leu Asn Pro Ala Glu Asp Lys Ile Val Val Ala Pro Gln
 260 265 270

Arg Trp Phe Lys Asp Ala Thr Asn Asp Thr Arg Asp Leu Ile Pro Ala
 275 280 285

Ala Trp Val Arg Leu
 290

37847-517 Sequence Listing_ST25 (003)

<210> 38
 <211> 288
 <212> PRT
 <213> Prevotella histicola

<400> 38

Met Lys Ile Val Lys Ile Leu Gly Gly Leu Gly Asn Gln Met Phe Gln
 1 5 10 15

Tyr Ala Leu Tyr Leu Ser Leu Lys Glu Thr Phe Pro Gln Glu Asn Val
 20 25 30

Thr Val Asp Leu Ser Cys Phe His Gly Tyr His Leu His Asn Gly Phe
 35 40 45

Glu Ile Ala Arg Ile Phe Ser Leu His Pro Asp Lys Ala Thr Val Met
 50 55 60

Glu Ile Leu Arg Ile Ala Tyr Tyr Tyr Pro Asn Tyr Phe Phe Trp Gln
 65 70 75 80

Ile Gly Lys Arg Val Leu Pro Gln Arg Lys Thr Met Cys Thr Glu Ser
 85 90 95

Thr Lys Leu Leu Phe Asp Lys Ser Val Leu Gln Arg Glu Gly Asp Arg
 100 105 110

Tyr Phe Asp Gly Tyr Trp Gln Asp Glu Arg Tyr Phe Ile Asp Cys Arg
 115 120 125

Arg Thr Ile Leu Asn Thr Phe Lys Phe Pro Pro Phe Thr Asp Asp Asn
 130 135 140

Asn Leu Ala Leu Leu Lys Lys Met Asp Thr Asn Ser Val Ser Ile His
 145 150 155 160

Val Arg Arg Gly Asp Tyr Val Gly Asn Lys Leu Tyr Gln Gly Ile Cys
 165 170 175

37847-517 Sequence Listing_ST25 (003)

Asp Leu Asn Tyr Tyr Arg Glu Ala Ile Met Lys Ile Ser Ser Tyr Ile
 180 185 190

Ser Pro Ser Met Phe Cys Val Phe Ser Asn Asp Ile Glu Trp Cys Arg
 195 200 205

Asp Asn Leu Glu Ser Phe Ile Lys Ala Pro Ile Tyr Tyr Val Asp Trp
 210 215 220

Asn Ser Gly Thr Glu Ser Tyr Arg Asp Met Gln Leu Met Ser Cys Cys
 225 230 235 240

Gly His Asn Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala Trp
 245 250 255

Leu Asn Gln Asn Ser Ser Lys Ile Val Ile Ala Pro Lys Arg Trp Ile
 260 265 270

Asn Leu Lys Asn Cys Gly Phe Met Leu Pro Ser Arg Trp Val Lys Ile
 275 280 285

<210> 39

<211> 273

<212> PRT

<213> Flavobacterium sp. WG21

<400> 39

Met Ile Val Val Gln Leu Ile Gly Gly Leu Gly Asn Gln Leu Phe Gln
 1 5 10 15

Tyr Ala Ala Ala Lys Ala Leu Ala Leu Gln Thr Lys Gln Lys Phe Ser
 20 25 30

Leu Asp Val Ser Gln Phe Glu Ser Tyr Lys Leu His Asn Tyr Ala Leu
 35 40 45

Asn His Phe Asn Val Ile Ser Lys Asn Tyr Lys Lys Pro Asn Arg Tyr
 50 55 60

Leu Arg Lys Ile Lys Ser Phe Tyr Gln Lys Asn Val Phe Tyr Lys Glu

37847-517 Sequence Listing_ST25 (003)

65		70		75		80									
Val	Asp	Phe	Gly	Tyr	Asn	Pro	Asp	Leu	Ile	His	Leu	Lys	Gly	Gly	Ile
			85					90					95		
Ile	Phe	Leu	Glu	Gly	Tyr	Phe	Gln	Ser	Glu	Lys	Tyr	Phe	Ile	Lys	Tyr
		100					105					110			
Glu	Lys	Glu	Ile	Arg	Glu	Asp	Phe	Glu	Leu	Arg	Thr	Pro	Leu	Lys	Lys
		115					120					125			
Glu	Thr	Lys	Ala	Ala	Ile	Ala	Lys	Ile	Glu	Ser	Val	Asn	Ser	Val	Ser
	130					135					140				
Ile	His	Ile	Arg	Arg	Gly	Asp	Tyr	Ile	Asn	Asn	Pro	Leu	His	Asn	Thr
145					150				155					160	
Ser	Lys	Glu	Glu	Tyr	Tyr	Asn	Lys	Ala	Leu	Glu	Ile	Val	Glu	Asn	Lys
			165					170					175		
Ile	Asn	Asn	Pro	Val	Tyr	Phe	Val	Phe	Ser	Asp	Asp	Met	Glu	Trp	Val
			180					185					190		
Lys	Ala	Asn	Phe	Ser	Thr	Lys	Gln	Glu	Thr	Ile	Phe	Ile	Asp	Phe	Asn
	195						200					205			
Asp	Ala	Ser	Thr	Asn	Phe	Glu	Asp	Leu	Lys	Leu	Met	Thr	Ser	Cys	Lys
	210					215					220				
His	Asn	Ile	Ile	Ala	Asn	Ser	Ser	Phe	Ser	Trp	Trp	Gly	Gly	Trp	Leu
225					230					235				240	
Asn	Lys	Asn	Pro	Asp	Lys	Ile	Val	Ile	Ala	Pro	Lys	Arg	Trp	Phe	Asn
			245						250					255	
Asp	Asp	Ser	Ile	Asn	Thr	Asn	Asp	Ile	Ile	Pro	Thr	Asn	Trp	Val	Lys
			260					265					270		
Ile															

37847-517 Sequence Listing_ST25 (003)

<210> 40
 <211> 267
 <212> PRT
 <213> Polaribacter franzmannii

<400> 40

Met Ile Ile Val Arg Ile Val Gly Gly Leu Gly Asn Gln Met Phe Gln
 1 5 10 15

Tyr Ala Tyr Ala Lys Ala Leu Gln Gln Lys Gly Tyr Gln Val Lys Ile
 20 25 30

Asp Ile Thr Lys Phe Lys Lys Tyr Asn Leu His Gly Gly Tyr Gln Leu
 35 40 45

Asp Gln Phe Lys Ile Asp Leu Glu Thr Ser Ser Pro Ile Ala Asn Val
 50 55 60

Leu Cys Arg Ile Gly Leu Arg Arg Ser Val Lys Glu Lys Ser Leu Leu
 65 70 75 80

Phe Asp Glu Lys Phe Leu Glu Ile Pro Gln Arg Glu Tyr Ile Lys Gly
 85 90 95

Tyr Phe Gln Thr Glu Lys Tyr Phe Ser Ser Ile Thr Pro Ile Leu Arg
 100 105 110

Lys Gln Phe Ile Val Gln Lys Glu Leu Cys Asn Thr Thr Leu Arg Tyr
 115 120 125

Leu Lys Glu Ile Thr Ile Gln Lys Asn Ala Cys Ser Leu His Ile Arg
 130 135 140

Arg Gly Asp Tyr Ile Ser Asp Glu Lys Ala Asn Ser Val His Gly Thr
 145 150 155 160

Cys Asp Leu Pro Tyr Tyr Lys Lys Ser Ile Lys Arg Ile Gln Asp Glu
 165 170 175

37847-517 Sequence Listing_ST25 (003)

28 May 2018
2015259088

Tyr Lys Asp Ala His Phe Phe Ile Phe Ser Asp Asp Ile Ser Trp Ala
180 185 190

Lys Lys Asn Leu Thr Asn Lys Asn Thr Thr Phe Ile Glu His Ile Val
195 200 205

Met Pro His Glu Asp Met His Leu Met Ser Leu Cys Lys His Asn Ile
210 215 220

Thr Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala Trp Leu Asn Gln His
225 230 235 240

Glu Asn Lys Thr Val Ile Ala Pro Lys Asn Trp Phe Val Asn Arg Glu
245 250 255

Asn Glu Val Ala Cys Ala Asn Trp Ile Gln Leu
260 265

<210> 41

<211> 267

<212> PRT

<213> Polaribacter sp. MED152

<400> 41

Met Val Val Val Arg Ile Leu Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Tyr Ala Lys Ser Leu Ala Glu Lys Gly Tyr Glu Val Gln Ile
20 25 30

Asp Ile Ser Lys Phe Lys Ser Tyr Lys Leu His Gly Gly Tyr His Leu
35 40 45

Asp Lys Phe Arg Ile Asp Leu Glu Thr Ala Asn Ser Ser Ser Ala Phe
50 55 60

Leu Ser Lys Ile Gly Leu Lys Lys Thr Ile Lys Glu Pro Asn Leu Leu
65 70 75 80

37847-517 Sequence Listing_ST25 (003)

Phe His Lys Asp Leu Leu Lys Val Asn Asn Asn Ala Phe Ile Lys Gly
85 90 95

Tyr Phe Gln Ala Glu Gln Tyr Phe Ser Asp Ile Arg Glu Ile Leu Ile
100 105 110

Asn Gln Phe Lys Ile Lys Lys Glu Leu Ala Lys Ser Thr Leu Ala Ile
115 120 125

Lys Asn Gln Ile Glu Leu Leu Lys Thr Thr Cys Ser Leu His Val Arg
130 135 140

Arg Gly Asp Tyr Ile Ser Asp Lys Lys Ala Asn Lys Val His Gly Thr
145 150 155 160

Cys Asp Leu Asp Tyr Tyr Ser Ser Ala Ile Glu His Ile Ser Lys Gln
165 170 175

Asn Ser Asn Val His Phe Phe Val Phe Ser Asp Asp Ile Ala Trp Val
180 185 190

Lys Asp Asn Leu Asn Ile Thr Asn Ala Thr Tyr Ile Asp His Asn Val
195 200 205

Ile Pro His Glu Asp Met Tyr Leu Met Thr Leu Cys Asn His Asn Ile
210 215 220

Thr Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala Trp Leu Asn Gln Asn
225 230 235 240

Pro Asp Lys Ile Val Ile Ala Pro Lys Asn Trp Phe Val Asp Lys Glu
245 250 255

Asn Glu Val Ala Cys Lys Ser Trp Ile Thr Leu
260 265

<210> 42
<211> 285
<212> PRT

37847-517 Sequence Listing_ST25 (003)

<213> Methanococcus maripaludis

<400> 42

Met Lys Ile Ile Gln Leu Lys Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Leu Tyr Lys Ser Leu Lys Lys Arg Gly Gln Glu Val Leu Leu
20 25 30

Asp Ile Ser Trp Tyr Leu Lys Asn Asn Ala His Asn Gly Tyr Glu Leu
35 40 45

Glu Trp Val Phe Gly Leu Ser Pro Glu Tyr Ala Ser Ile Arg Gln Cys
50 55 60

Phe Lys Leu Gly Asp Ile Pro Ile Asn Leu Ile Tyr Asn Val Lys Arg
65 70 75 80

Lys Val Phe Pro Lys Lys Thr His Phe Phe Glu Lys Ser Asn Phe Asn
85 90 95

Tyr Asp Asn Asn Val Phe Glu Val Thr Asn Gly Tyr Phe Glu Gly Tyr
100 105 110

Trp Gln Asn Glu Asn Tyr Phe Lys Asn Phe Arg Ser Glu Ile Leu Asn
115 120 125

Asp Phe Ser Phe Lys Asn Ile Asp Lys Arg Asn Ala Glu Phe Ser Glu
130 135 140

Tyr Leu Lys Ser Ile Asn Ser Val Ser Val His Val Arg Arg Gly Asp
145 150 155 160

Tyr Val Thr Asn Gln Lys Ala Leu Asn Val His Gly Asn Ile Cys Asn
165 170 175

Leu Glu Tyr Tyr Asn Lys Ala Ile Asn Leu Ala Asn Asn Asn Leu Lys
180 185 190

37847-517 Sequence Listing_ST25 (003)

Asn Pro Lys Phe Val Ile Phe Ser Asp Asp Ile Thr Trp Cys Lys Ser
 195 200 205

Asn Leu Gly Ile Asp Asp Pro Val Tyr Val Asp Trp Asn Thr Gly Pro
 210 215 220

Tyr Ser Tyr Gln Asp Met Tyr Leu Met Ser Asn Cys Lys Asn Asn Ile
 225 230 235 240

Ile Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala Trp Leu Asn Gln Asn
 245 250 255

Thr Glu Lys Lys Val Phe Ser Pro Lys Lys Trp Val Asn Asp Arg Asn
 260 265 270

Asn Val Asn Ile Val Pro Asn Gly Trp Ile Lys Ile Lys
 275 280 285

<210> 43

<211> 292

<212> PRT

<213> Gallionella sp. SCGC AAA018-N21

<400> 43

Met Ile Ile Ala His Ile Ile Gly Gly Leu Gly Asn Gln Met Phe Gln
 1 5 10 15

Tyr Ala Ala Gly Arg Ala Leu Ser Leu Ala Arg Gly Val Pro Phe Lys
 20 25 30

Leu Asp Ile Ser Gly Phe Glu Gly Tyr Asp Leu His Gln Gly Phe Glu
 35 40 45

Leu Gln Arg Val Phe Asn Cys Ala Ala Gly Ile Ala Ser Glu Ala Glu
 50 55 60

Val Arg Asp Ser Leu Gly Trp Gln Phe Ser Ser Pro Ile Arg Arg Ile
 65 70 75 80

Val Ala Arg Pro Ser Leu Ala Val Leu Arg Arg Ser Thr Phe Val Val

37847-517 Sequence Listing_ST25 (003)

85 90 95

Glu Pro His Phe His Tyr Trp Ala Gly Ile Lys Gln Val Pro Asp Asn
100 105 110

Cys Tyr Leu Ala Gly Tyr Trp Gln Ser Glu Gln Tyr Phe Gln Ser His
115 120 125

Ala Ala Val Ile Arg Thr Asp Phe Ala Phe Lys Pro Pro Leu Ser Gly
130 135 140

Gln Asn Ser Lys Leu Ala Met Gln Ile Ala Gln Gly Asn Ala Val Ser
145 150 155 160

Leu His Ile Arg Arg Gly Asp Tyr Ala Asn Asn Pro Lys Thr Thr Ala
165 170 175

Thr His Gly Leu Cys Ser Leu Asp Tyr Tyr Arg Ala Ala Ile Gln His
180 185 190

Ile Ala Glu Arg Val Gln Ser Pro His Phe Phe Ile Phe Ser Asp Asp
195 200 205

Ile Ala Trp Val Lys Ser Asn Leu Ala Ile Asn Phe Pro His Gln Tyr
210 215 220

Val Asp His Asn Gln Gly Thr Glu Ser Tyr Asn Asp Met Arg Leu Met
225 230 235 240

Ser Leu Cys Gln His Asn Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp
245 250 255

Gly Ala Trp Leu Asn Thr Asn Ala His Lys Ile Val Ile Ala Pro Lys
260 265 270

Gln Trp Phe Ala Asn Thr Thr His Val Ala Asp Leu Ile Pro Ser Ser
275 280 285

Trp Glu Arg Leu

37847-517 Sequence Listing_ST25 (003)

290

<210> 44
 <211> 311
 <212> PRT
 <213> *Azospira oryzae*

<400> 44

Met Gln Ser Pro Ala Cys Ile Ala Gly Ala Arg Ala Trp Trp Val Gly
 1 5 10 15

Tyr Gly Met Ala Glu Ala Met Gln Pro Val Val Val Gly Leu Ser Gly
 20 25 30

Gly Leu Gly Asn Gln Met Phe Gln Tyr Ala Ala Gly Arg Ala Leu Ala
 35 40 45

His Arg Leu Gly His Pro Leu Ser Leu Asp Leu Ser Trp Phe Gln Gly
 50 55 60

Arg Gly Asp Arg His Phe Ala Leu Ala Pro Phe His Ile Ala Ala Ser
 65 70 75 80

Leu Glu Arg Ala Trp Pro Arg Leu Pro Pro Ala Met Gln Ala Gln Leu
 85 90 95

Ser Arg Leu Ser Arg Arg Trp Ala Pro Arg Ile Met Gly Ala Pro Val
 100 105 110

Phe Arg Glu Pro His Phe His Tyr Val Pro Ala Phe Ala Ala Leu Ala
 115 120 125

Ala Pro Val Phe Leu Glu Gly Tyr Trp Gln Ser Glu Arg Tyr Phe Arg
 130 135 140

Glu Leu Arg Glu Pro Leu Leu Gln Asp Phe Ser Leu Arg Gln Pro Leu
 145 150 155 160

Pro Ala Ser Cys Gln Pro Ile Leu Ala Ala Ile Gly Asn Ser Asp Ala
 165 170 175

37847-517 Sequence Listing_ST25 (003)

Ile Cys Val His Val Arg Arg Gly Asp Tyr Leu Ser Asn Pro Val Ala
180 185 190

Ala Lys Val His Gly Val Cys Pro Val Asp Tyr Tyr Gln Gln Gly Val
195 200 205

Ala Glu Leu Ser Ala Ser Leu Ala Arg Pro His Cys Phe Val Phe Ser
210 215 220

Asp Asp Pro Glu Trp Val Arg Gly Ser Leu Ala Phe Pro Cys Pro Met
225 230 235 240

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

Leu Met Ala Ala Cys Gln His Phe Val Ile Ala Asn Ser Ser Leu Ser
260 265 270

Trp Trp Gly Ala Trp Leu Gly Gln Ala Ala Gly Lys Arg Val Ile Ala
275 280 285

Pro Ser Arg Trp Phe Leu Thr Ser Asp Lys Asp Ala Arg Asp Leu Leu
290 295 300

Pro Pro Ser Trp Glu Arg Arg
305 310

<210> 45

<211> 290

<212> PRT

<213> Prevotella paludivivens

<400> 45

Met Lys Ile Val Lys Ile Ile Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Leu Ala Met Ala Leu Asn Lys Asn Phe Thr Asp Glu Glu Val
20 25 30

37847-517 Sequence Listing_ST25 (003)

Lys Leu Asp Ile His Cys Phe Asn Gly Tyr Thr Lys His Gln Gly Phe
 35 40 45
 Glu Ile Asp Arg Val Phe Gly Asn Glu Phe Glu Leu Ala Ser Tyr Arg
 50 55 60
 Asp Val Ala Lys Val Ala Tyr Pro Tyr Phe Asn Phe Gln Leu Trp Arg
 65 70 75 80
 Ile Gly Ser Arg Ile Phe Pro Asp Arg Arg His Met Ile Ser Glu Asp
 85 90 95
 Thr Ser Phe Lys Ile Met Pro Glu Val Ile Thr Ser His Asn Tyr Lys
 100 105 110
 Tyr Tyr Asp Gly Tyr Trp Gln His Glu Glu Tyr Phe Lys Asn Ile His
 115 120 125
 Asp Glu Ile Leu Asp Ala Phe Lys Phe Pro Lys Phe Gln Asp Glu Arg
 130 135 140
 Asn Lys Ala Leu Ala Glu Arg Leu Ser Asp Ser Asn Ser Ile Ser Ile
 145 150 155 160
 His Ile Arg Arg Gly Asp Tyr Leu Asn Asp Glu Leu Phe Lys Gly Thr
 165 170 175
 Cys Gly Ile Glu Tyr Tyr Lys Lys Ala Ile Glu Glu Ile Asn Glu Arg
 180 185 190
 Thr Val Pro Thr Leu Phe Cys Val Phe Ser Asn Asp Ile His Trp Cys
 195 200 205
 Lys Glu Asn Ile Glu Pro Leu Leu Asn Gly Lys Glu Thr Ile Tyr Val
 210 215 220
 Asp Trp Asn Thr Gly Ser Asp Asn Tyr Arg Asp Met Gln Leu Met Thr
 225 230 235 240

37847-517 Sequence Listing_ST25 (003)

Lys Cys Lys His Asn Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Gly
245 250 255

Ala Trp Leu Asn Asn Thr Lys Asp Lys Ile Val Ile Ala Pro Arg Ile
260 265 270

Trp Tyr Asn Thr Lys Glu Lys Val Ser Pro Val Ala Asn Ser Trp Ile
275 280 285

Lys Leu
290

<210> 46

<211> 300

<212> PRT

<213> Gramella forsetii

<400> 46

Met Ser Asn Lys Asn Pro Val Ile Val Glu Ile Met Gly Gly Leu Gly
1 5 10 15

Asn Gln Met Phe Gln Phe Ala Val Ala Lys Leu Leu Ala Glu Lys Asn
20 25 30

Ser Ser Val Leu Leu Val Asp Thr Asn Phe Tyr Lys Glu Ile Ser Gln
35 40 45

Asn Leu Lys Asp Phe Pro Arg Tyr Phe Ser Leu Gly Ile Phe Asp Ile
50 55 60

Ser Tyr Lys Met Gly Thr Glu Asn Gly Met Val Asn Phe Lys Asn Leu
65 70 75 80

Ser Phe Lys Asn Arg Val Ser Arg Lys Leu Gly Leu Asn Tyr Pro Lys
85 90 95

Ile Phe Lys Glu Lys Ser Tyr Arg Phe Asp Ala Asp Leu Phe Asn Lys
100 105 110

37847-517 Sequence Listing_ST25 (003)

Lys Thr Pro Ile Tyr Leu Lys Gly Tyr Phe Gln Ser Tyr Lys Tyr Phe
 115 120 125

Ile Gly Val Glu Ser Lys Ile Arg Gln Trp Phe Glu Phe Pro Tyr Glu
 130 135 140

Asn Leu Gly Val Gly Asn Glu Glu Ile Lys Ser Lys Ile Leu Glu Lys
 145 150 155 160

Thr Ser Val Ser Val His Ile Arg Arg Gly Asp Tyr Val Glu Asn Lys
 165 170 175

Lys Thr Lys Glu Phe His Gly Asn Cys Ser Leu Glu Tyr Tyr Lys Asn
 180 185 190

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

Phe Ser Asp Asp Ile Ser Trp Val Arg Asp Glu Phe Lys Asp Leu Pro
 210 215 220

Asn Glu Lys Val Phe Val Thr Gly Asn Leu His Glu Asn Ser Trp Lys
 225 230 235 240

Asp Met Tyr Leu Met Ser Leu Cys Asp His Asn Ile Ile Ala Asn Ser
 245 250 255

Ser Phe Ser Trp Trp Ala Ala Trp Leu Asn Asn Asn Ser Glu Lys Asn
 260 265 270

Val Ile Ala Pro Lys Lys Trp Phe Ala Asp Ile Asp Gln Glu Gln Lys
 275 280 285

Ser Leu Asp Leu Leu Pro Pro Ser Trp Ile Arg Met
 290 295 300

<210> 47

<211> 298

<212> PRT

<213> Mariprofundus ferrooxydans

37847-517 Sequence Listing_ST25 (003)

<400> 47

Met Ile Ile Val Gln Phe Thr Gly Gly Leu Gly Asn Gln Met Phe Gln
 1 5 10 15

Tyr Ala Leu Gly Arg Arg Leu Ser Leu Leu His Asp Val Glu Leu Lys
 20 25 30

Phe Asp Leu Ser Phe Tyr Gln His Asp Ile Leu Arg Asp Phe Met Leu
 35 40 45

Asp Arg Phe Gln Val Asn Gly Gln Val Ala Thr Glu Lys Glu Ile Glu
 50 55 60

Ala Tyr Thr Asn Thr Pro Ile Phe Ala Leu Asp Arg Pro Leu Leu Asp
 65 70 75 80

Arg Leu Val Arg Trp Gly Leu Tyr Arg Gly Ile Val Ser Val Ser Asp
 85 90 95

Glu Pro Pro Gly Lys Gln Ala Leu Met Val Tyr Asn Ser Arg Val Leu
 100 105 110

Gln Ala Pro Arg Asn Thr Tyr Val Gln Gly Tyr Trp Gln Ser Glu Lys
 115 120 125

Tyr Phe Met Pro Ile Arg Gln Lys Leu Leu Asp Asp Phe Ser Leu Val
 130 135 140

Asp Lys Ala Asp Gln Ala Asn Gly Ala Met Leu Glu Lys Ile Arg Gln
 145 150 155 160

Cys His Ser Val Ser Leu His Val Arg Arg Gly Asp Tyr Val Ser Asn
 165 170 175

Pro Leu Thr Asn His Ser His Gly Thr Cys Gly Leu Glu Tyr Tyr Glu
 180 185 190

Lys Ala Ile Ala Leu Ile Gly Ser Lys Val Asp Asp Pro His Phe Phe

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

195

200

205

Val Phe Ser Asp Asp Pro Glu Trp Thr Arg Asp His Leu Lys Cys Arg
210 215 220

Phe Pro Met Thr Tyr Val Thr Cys Asn Ser Ala Asp Ser Cys Glu Trp
225 230 235 240

Asp Met Glu Leu Met Arg His Cys Arg Asp His Ile Ile Ala Asn Ser
245 250 255

Ser Phe Ser Trp Trp Gly Ala Trp Leu Asn Met Asn Pro Asp Lys Val
260 265 270

Val Val Ala Pro Ala Ala Trp Phe Asn Asn Phe Ser Ala Asp Thr Ser
275 280 285

Asp Leu Ile Pro Asp Ser Trp Val Arg Ile
290 295

<210> 48

<211> 291

<212> PRT

<213> Bacillus cereus

<400> 48

Met Lys Ile Ile Gln Val Ser Ser Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Leu Tyr Lys Lys Ile Ser Leu Asn Asp Asn Asp Val Phe Leu
20 25 30

Asp Ser Ser Thr Ser Tyr Met Met Tyr Lys Asn Gln His Asn Gly Tyr
35 40 45

Glu Leu Glu Arg Ile Phe His Ile Lys Pro Arg His Ala Gly Lys Glu
50 55 60

Ile Ile Asp Asn Leu Ser Asp Leu Asp Ser Glu Leu Ile Ser Arg Ile
65 70 75 80

37847-517 Sequence Listing_ST25 (003)

Arg Arg Lys Leu Phe Gly Ala Lys Lys Ser Met Tyr Val Glu Leu Lys
85 90 95

Glu Phe Glu Tyr Asp Pro Ile Ile Phe Glu Lys Lys Glu Thr Tyr Phe
100 105 110

Lys Gly Tyr Trp Gln Asn Tyr Asn Tyr Phe Lys Asp Ile Glu Gln Glu
115 120 125

Leu Arg Lys Asp Phe Val Phe Thr Glu Lys Leu Asp Lys Arg Asn Glu
130 135 140

Lys Leu Ala Asn Glu Ile Arg Asn Lys Asn Ser Val Ser Ile His Ile
145 150 155 160

Arg Arg Gly Asp Tyr Tyr Leu Asn Lys Val Tyr Glu Glu Lys Phe Gly
165 170 175

Asn Ile Ala Asn Leu Glu Tyr Tyr Leu Lys Ala Ile Asn Leu Val Lys
180 185 190

Lys Lys Ile Glu Asp Pro Lys Phe Tyr Ile Phe Ser Asp Asp Ile Asp
195 200 205

Trp Ala Gln Lys Asn Ile Asn Leu Thr Asn Asp Val Val Tyr Ile Ser
210 215 220

His Asn Gln Gly Asn Glu Ser Tyr Lys Asp Met Gln Leu Met Ser Leu
225 230 235 240

Cys Lys His Asn Ile Ile Ala Asn Ser Thr Phe Ser Trp Trp Gly Ala
245 250 255

Phe Leu Asn Asn Asn Asp Asp Lys Ile Val Val Ala Pro Lys Lys Trp
260 265 270

Ile Asn Ile Lys Gly Leu Glu Lys Val Glu Leu Phe Pro Glu Asn Trp
275 280 285

2015259088 28 May 2018

Ile Thr Tyr
290

<210> 49

<211> 300

<212> PRT

<213> Firmicutes bacterium CAG:534

<400> 49

Met Ile Ile Ile Arg Met Thr Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Leu Tyr Leu Lys Leu Arg Ala Met Gly Lys Glu Val Lys Met
20 25 30

Asp Asp Phe Thr Glu Tyr Glu Gly Arg Glu Ala Arg Pro Leu Ser Leu
35 40 45

Trp Ala Phe Gly Ile Glu Tyr Asp Arg Ala Ser Arg Glu Glu Leu Cys
50 55 60

Arg Met Thr Asp Gly Phe Leu Asp Pro Val Ser Arg Ile Arg Arg Lys
65 70 75 80

Leu Phe Gly Arg Lys Ser Leu Glu Tyr Met Glu Lys Asp Cys Asn Phe
85 90 95

Asp Pro Glu Ile Leu Asn Arg Asp Pro Ala Tyr Leu Thr Gly Tyr Phe
100 105 110

Gln Ser Glu Lys Tyr Phe Ala Asp Ile Glu Glu Glu Val Arg Gln Ala
115 120 125

Phe Arg Phe Ser Glu Arg Ile Trp Glu Gly Ile Pro Ser Gln Leu Leu
130 135 140

Glu Arg Ile Arg Ser Tyr Glu Gln Gln Ile Lys Thr Thr Met Ala Val
145 150 155 160

37847-517 Sequence Listing_ST25 (003)

Ser Val His Ile Arg Arg Gly Asp Tyr Leu Gln Asn Glu Glu Ala Tyr
 165 170 175

Gly Gly Ile Cys Thr Glu Arg Tyr Tyr Lys Thr Ala Ile Glu Tyr Val
 180 185 190

Lys Lys Arg Gln Gln Asp Ala Ser Phe Phe Val Phe Thr Asn Asp Pro
 195 200 205

Asp Tyr Ala Gly Glu Trp Ile Leu Lys Asn Phe Gly Gln Glu Lys Glu
 210 215 220

Arg Phe Val Leu Ile Glu Gly Thr Gln Glu Glu Asn Gly Tyr Leu Asp
 225 230 235 240

Leu Tyr Leu Met Ser Leu Cys Arg His His Ile Leu Ala Asn Ser Ser
 245 250 255

Phe Ser Trp Trp Gly Ala Tyr Leu Asn Pro Ser Arg Glu Lys Met Val
 260 265 270

Ile Val Pro His Lys Trp Phe Gly Asn Gln Glu Cys Arg Asp Ile Tyr
 275 280 285

Met Glu Asn Met Ile Arg Ile Ala Lys Glu Gln Ser
 290 295 300

<210> 50

<211> 292

<212> PRT

<213> Sideroxydans lithotrophicus

<400> 50

Met Val Ile Ser Asn Ile Ile Gly Gly Leu Gly Asn Gln Met Phe Gln
 1 5 10 15

Tyr Ala Ala Ala Arg Ala Leu Ser Leu Lys Leu Glu Val Pro Leu Lys
 20 25 30

37847-517 Sequence Listing_ST25 (003)

```

Leu Asp Ile Ser Gly Phe Thr Asn Tyr Ala Leu His Gln Gly Phe Glu
   35                               40                               45

Leu Asp Arg Ile Phe Gly Cys Lys Ile Glu Ile Ala Ser Glu Ala Asp
   50                               55                               60

Val His Glu Ile Leu Gly Trp Gln Ser Ala Ser Gly Ile Arg Arg Val
   65                               70                               75

Val Ser Arg Pro Gly Met Ser Ile Phe Arg Arg Lys Gly Phe Val Val
   85                               90                               95

Glu Pro His Phe Ser Tyr Trp Asn Gly Ile Arg Lys Ile Thr Gly Asp
  100                               105                               110

Cys Tyr Leu Ala Gly Tyr Trp Gln Ser Glu Lys Tyr Phe Leu Asp Ala
  115                               120                               125

Ala Val Glu Ile Arg Lys Asp Phe Ser Phe Lys Leu Pro Leu Asp Ser
  130                               135                               140

His Asn Ala Glu Leu Ala Glu Lys Ile Asp Gln Glu Asn Ala Val Ser
  145                               150                               155                               160

Leu His Ile Arg Arg Gly Asp Tyr Ala Asn Asn Pro Leu Thr Ala Ala
  165                               170                               175

Thr His Gly Leu Cys Ser Leu Asp Tyr Tyr Arg Lys Ser Ile Lys His
  180                               185                               190

Ile Ala Gly Gln Val Arg Asn Pro Tyr Phe Phe Val Phe Ser Asp Asp
  195                               200                               205

Ile Ala Trp Val Lys Asp Asn Leu Glu Ile Glu Phe Pro Ser Gln Tyr
  210                               215                               220

Val Asp Tyr Asn His Gly Ser Met Ser Phe Asn Asp Met Arg Leu Met
  225                               230                               235                               240

```

37847-517 Sequence Listing_ST25 (003)

Ser Leu Cys Lys His His Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp
 245 250 255

Gly Ala Trp Leu Asn Pro Asn Pro Glu Lys Val Val Ile Ala Pro Glu
 260 265 270

Arg Trp Phe Ala Asn Arg Thr Asp Val Gln Asp Leu Leu Pro Pro Gly
 275 280 285

Trp Val Lys Leu
 290

<210> 51
 <211> 292
 <212> PRT
 <213> zeta proteobacterium SCGC AB-137-C09

<400> 51

Met Ile Val Ser Gln Ile Ile Gly Gly Leu Gly Asn Gln Met Phe Gln
 1 5 10 15

Tyr Ala Thr Gly Arg Ala Leu Ser His Arg Leu His Asp Thr Phe Phe
 20 25 30

Leu Asp Leu Asp Gly Phe Ser Gly Tyr Gln Leu His Gln Gly Phe Glu
 35 40 45

Leu Ser Asn Val Phe Gln Cys Glu Val Asn Val Ala Thr Arg Ser Gln
 50 55 60

Met Gln Ala Leu Leu Gly Trp Arg Ser Phe Ser Ser Val Arg Arg Leu
 65 70 75 80

Leu Met Lys Arg Ser Leu Lys Trp Ala Arg Gly His Arg Val Met Ile
 85 90 95

Glu Pro His Phe His Tyr Trp Ser Arg Phe Ala Glu Ile Asn Glu Gly
 100 105 110

Cys Tyr Leu Ser Gly Tyr Trp Gln Ser Glu Arg Tyr Phe Lys Pro Ile

2015259088 28 May 2018

120

125

Val	Asn	Leu	Asp	Leu	Ala	Gln	Gln	Met	Thr	Glu	Val	Asn	Ser	Val	Ser
145					150					155					160

Leu His Val Arg Arg Gly Asp Tyr Ala Ser Asp Ala Asn Thr Asn His
165 170 175

Thr His Gly Leu Cys Pro Leu Asp Tyr Tyr Arg Asp Ala Ile Leu Tyr
180 185 190

Ile Ala Gln Asn Thr Val Ala Pro Ser Phe Phe Ile Phe Ser Asp Asp
195 200 205

Ile Glu Trp Cys Arg Glu His Leu Lys Leu Ser Phe Pro Ala Thr Tyr
210 215 220

Ile Asp His Asn Lys Gly Ser Asn Ser Tyr Cys Asp Met Gln Leu Met
225 230 235 240

Ser Leu Cys His His His Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp
245 250 255

Gly Ala Trp Leu Asn Thr Arg Leu Asp Lys Ile Val Ile Ala Pro Lys
260 265 270

Gln Trp Phe Ala Asn Gly Asn Arg Thr Asp Asp Leu Ile Pro Ala Glu
275 280 285

Trp Leu Val Met
290

<210> 52

<211> 283

<212> PRT

<213> Pedobacter heparinus

37847-517 Sequence Listing_ST25 (003)

<400> 52

Met Lys Ile Ile Arg Phe Leu Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Phe Tyr Lys Ser Leu Gln His Arg Phe Pro His Val Lys Ala
20 25 30

Asp Leu Gln Gly Tyr Gln Glu Tyr Thr Leu His Asn Gly Phe Glu Leu
35 40 45

Glu His Ile Phe Asn Ile Lys Val Asn Ser Val Ser Ser Phe Thr Ser
50 55 60

Asp Leu Phe Tyr Asn Lys Lys Trp Leu Tyr Arg Lys Leu Arg Arg Ile
65 70 75 80

Leu Asn Leu Arg Asn Thr Tyr Ile Glu Glu Lys Lys Leu Phe Ser Phe
85 90 95

Asp Pro Ser Leu Leu Asn Asn Pro Lys Ser Ala Tyr Tyr Trp Gly Tyr
100 105 110

Trp Gln Asn Phe Gln Tyr Phe Glu His Ile Ala Asp Asp Leu Arg Lys
115 120 125

Asp Phe Gln Phe Arg Ala Pro Leu Ser Ala Gln Asn Gln Glu Val Leu
130 135 140

Asp Gln Thr Lys Leu Ser Asn Ser Ile Ser Leu His Ile Arg Arg Gly
145 150 155 160

Asp Tyr Ile Lys Asp Pro Leu Leu Gly Gly Leu Cys Gly Pro Glu Tyr
165 170 175

Tyr Gln Thr Ala Ile Asn Tyr Ile Thr Ser Lys Val Asn Ala Ala Arg
180 185 190

Phe Phe Ile Phe Ser Asp Asp Ile Asp Trp Cys Ile Ala Asn Leu Lys
195 200 205

37847-517 Sequence Listing_ST25 (003)

Leu Gln Asp Cys Ser Phe Ile Ser Trp Asn Lys Gly Thr Ser Ser Tyr
 210 215 220

Ile Asp Met Gln Leu Met Ser Ser Cys Lys His His Ile Val Ala Asn
 225 230 235 240

Ser Ser Phe Ser Trp Trp Ala Ala Trp Leu Asn Pro Asn Pro Asp Lys
 245 250 255

Ile Val Ile Ala Pro Glu Lys Trp Thr Asn Asp Lys Asp Ile Asn Val
 260 265 270

Arg Met Ser Phe Pro Gln Gly Trp Ile Ser Leu
 275 280

<210> 53

<211> 278

<212> PRT

<213> Methylophilus methylotrophus

<400> 53

Met Phe Gln Tyr Ala Met Gly Leu Ser Leu Ala Glu Asn Asn Gln Thr
 1 5 10 15

Pro Leu Lys Leu Asp Leu Ser Gln Phe Thr Asp Tyr Lys Leu His Asn
 20 25 30

Gly Phe Glu Leu Ser Lys Val Phe Asn Cys Ser Ala Glu Thr Ala Ser
 35 40 45

Val Thr Gln Ile Glu Thr Leu Leu Gly Ile Cys Lys Tyr Ser Phe Ile
 50 55 60

Arg Arg Ile Leu Lys Asn Thr Tyr Leu Lys Asn Leu Arg Pro Ala Gln
 65 70 75 80

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

37847-517 Sequence Listing_ST25 (003)

Gly Asp Asn Val Tyr Leu Glu Gly Tyr Trp Gln Ser Gln Lys Tyr Phe
100 105 110

Ile Asp Tyr Glu Ser Thr Ile Arg Thr His Phe Thr Phe Lys Asn Ile
115 120 125

Leu Ser Gly Glu Asn Leu Lys Leu Ser Asp Arg Ile Lys Gly Ser Asn
130 135 140

Ser Val Ser Leu His Ile Arg Arg Gly Asp Tyr Val Thr Asn Lys Asn
145 150 155 160

Asn Ala Phe Ile Gly Thr Cys Ser Leu Ile Tyr Tyr Gln Asn Ala Ile
165 170 175

Glu Tyr Phe Ser Thr Lys Ile Ala Asp Pro Ile Phe Phe Ile Phe Ser
180 185 190

Asp Asp Ile Thr Trp Ala Lys Ser Asn Leu Arg Leu Ala Asn Glu His
195 200 205

Tyr Phe Val Gly His Asn Gln Gly Glu Asp Ser His Phe Asp Met Gln
210 215 220

Leu Met Ser Leu Cys Lys His His Ile Ile Ala Asn Ser Ser Phe Ser
225 230 235 240

Trp Trp Gly Ala Trp Leu Asn Pro Ser Lys Asp Lys Ile Ile Ile Ala
245 250 255

Pro Lys Lys Trp Phe Ala Ser Gly Leu Asn Asp Gln Asp Leu Val Pro
260 265 270

Lys Asp Trp Leu Arg Ile
275

<210> 54
<211> 285
<212> PRT

37847-517 Sequence Listing_ST25 (003)

<213> Rhodobacterales bacterium HTCC2255

<400> 54

Met Ile Tyr Thr Arg Ile Arg Gly Gly Leu Gly Asn Gln Leu Phe Gln
 1 5 10 15

Tyr Ser Ala Ala Arg Ser Leu Ala Asp Tyr Leu Asn Val Ser Leu Gly
 20 25 30

Leu Asp Thr Arg Glu Phe Asp Glu Asn Ser Pro Tyr Lys Met Ser Leu
 35 40 45

Asn His Phe Asn Ile Arg Ala Asp Leu Asn Pro Pro Asp Leu Ile Lys
 50 55 60

His Lys Lys Asp Gly Lys Ile Ala Tyr Ile Ile Asp His Ile Lys Gly
 65 70 75 80

Asn Gln Lys Lys Val Tyr Lys Glu Pro Phe Leu Ser Phe Asp Lys Asn
 85 90 95

Leu Phe Ser Asn Val Asp Gly Thr Tyr Leu Lys Gly Tyr Trp Gln Ser
 100 105 110

Glu Lys Tyr Phe Leu Arg Asn Arg Lys Asn Ile Leu Ser Asp Ile Asn
 115 120 125

Leu Ile Lys Lys Thr Asp Lys Phe Asn Thr Ile Asn Leu Lys Glu Ile
 130 135 140

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

Ser Asn Glu Ser Tyr Asn Glu Thr His Gly Ile Cys Ser Leu Ser Tyr
 165 170 175

Tyr Thr Asp Ala Val Glu Tyr Ile Lys Asn Arg Leu Gly Glu Asn Ile
 180 185 190

37847-517 Sequence Listing_ST25 (003)

Lys Val Phe Ala Phe Ser Asp Asp Pro Asp Trp Val Leu Glu Asn Leu
195 200 205

Lys Leu Ser Val Asp Ile Lys Ile Ile Asn Asn Asn Thr Ser Ala Asn
210 215 220

Ser Phe Glu Asp Leu Arg Leu Met Leu Asn Cys Asp His Asn Ile Ile
225 230 235 240

Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala Trp Leu Asn Gln Asn Pro
245 250 255

Glu Lys Ile Val Ile Ser Pro Lys Lys Trp Tyr Asn Lys Lys Gln Leu
260 265 270

Gln Asn Ala Asp Ile Val Pro Ser Ser Trp Leu Lys Tyr
275 280 285

<210> 55

<211> 295

<212> PRT

<213> Spirulina subsalsa

<400> 55

Met Ala Lys Ile Ile Ala Arg Ile Arg Gly Gly Ile Gly Asn Gln Ile
1 5 10 15

Phe Ile Tyr Ala Ala Ala Arg Arg Leu Glu Leu Ile Asn Asn Ala Glu
20 25 30

Leu Val Leu Asp Ser Val Ser Gly Phe Val His Asp Leu Gln Tyr Arg
35 40 45

Gln His Tyr Gln Leu Asp His Phe His Ile Pro Cys Arg Lys Ala Thr
50 55 60

Pro Ala Glu Arg Phe Glu Pro Phe Ser Arg Val Arg Arg Tyr Leu Lys
65 70 75 80

Arg Gln Leu Asn Gln Arg Leu Pro Phe Glu Gln Arg Arg Tyr Val Ile

37847-517 Sequence Listing_ST25 (003)

85

90

95

Gln Glu Ser Ile Asp Phe Asp Pro Arg Leu Ile Glu Phe Lys Pro Arg
100 105 110

Gly Thr Val His Leu Glu Gly Tyr Trp Gln Ser Glu Asp Tyr Phe Lys
115 120 125

Asp Ile Glu Ala Thr Ile Arg Gln Asp Leu Gln Ile Gln Pro Pro Thr
130 135 140

Asp Pro Thr Asn Leu Ala Ile Val Gln His Ile His Gln His Thr Ser
145 150 155 160

Val Ala Val His Ile Arg Phe Phe Asp Gln Pro Asn Ala Asp Thr Met
165 170 175

Asn Asn Ala Pro Ser Asp Tyr Tyr His Arg Ala Val Glu Ala Met Glu
180 185 190

Thr Phe Val Pro Gly Ala His Tyr Tyr Leu Phe Ser Asp Gln Pro Glu
195 200 205

Ala Ala Lys Ser Arg Ile Pro Leu Pro Asp Glu Arg Val Thr Leu Val
210 215 220

Asn His Asn Arg Gly Asn Lys Leu Ala Tyr Ala Asp Leu Trp Leu Met
225 230 235 240

Thr Gln Cys Gln His Phe Ile Ile Ala Asn Ser Thr Phe Ser Trp Trp
245 250 255

Gly Ala Trp Leu Ala Glu Asn Gln Lys Lys Gln Val Ile Ala Pro Gly
260 265 270

Phe Glu Lys Arg Glu Gly Val Ser Trp Trp Gly Phe Lys Gly Leu Leu
275 280 285

Pro Lys Gln Trp Ile Lys Leu

37847-517 Sequence Listing_ST25 (003)

290

295

<210> 56
 <211> 287
 <212> PRT
 <213> *Vibrio cyclitrophicus*

<400> 56

Met Val Ile Val Lys Ile Thr Gly Gly Leu Gly Asn Gln Leu Phe Gln
 1 5 10 15

Tyr Ala Thr Gly Ser Ala Leu Ala Asn Lys Leu Ser Cys Glu Leu Val
 20 25 30

Leu Asp Leu Ser Phe Tyr Pro Thr Gln Thr Leu Arg Lys Tyr Glu Leu
 35 40 45

Ala Lys Phe Asn Ile Asn Ala Arg Val Ala Thr Asp Arg Glu Ile Phe
 50 55 60

Leu Ala Gly Gly Gly Asn Asp Phe Phe Ser Lys Ala Leu Lys Lys Leu
 65 70 75 80

Gly Leu Thr Ser Ile Ile Phe Pro Glu Tyr Ile Lys Glu Gln Glu Ser
 85 90 95

Ile Lys Tyr Val Gly Lys Ile Asp Leu Cys Lys Ser Gly Ala Tyr Leu
 100 105 110

Asp Gly Tyr Trp Gln Asn Pro Leu Tyr Phe Ser Gln Asn Lys Ile Glu
 115 120 125

Leu Thr Arg Glu Phe Leu Pro Arg Ala Gln Leu Ser Pro Ser Ala Leu
 130 135 140

Ala Trp Lys Asp His Ile Ser Gln Ala Ser Asn Ser Val Ser Leu His
 145 150 155 160

Val Arg Arg Gly Asp Tyr Val Glu Asn Ala His Thr Asn Asn Ile His
 165 170 175

37847-517 Sequence Listing_ST25 (003)

28 May 2018
2015259088

Gly Thr Cys Ser Leu Glu Tyr Tyr Gln His Ala Ile Glu Lys Ile Arg
180 185 190

Ser Glu Val His Asn Pro Val Phe Phe Val Phe Ser Asp Asp Ile Glu
195 200 205

Trp Cys Lys Leu Asn Leu Ser Ser Leu Ala Glu Val Glu Phe Val Asp
210 215 220

Asn Thr Thr Ser Ala Ile Asp Asp Leu Met Leu Met Arg Gln Cys Lys
225 230 235 240

His Ser Ile Ile Ala Asn Ser Thr Phe Ser Trp Trp Gly Ala Trp Leu
245 250 255

Lys Leu Asp Gly Leu Val Ile Ala Pro Arg Asn Trp Phe Ser Ser Ala
260 265 270

Ser Arg Asn Leu Lys Gly Ile Tyr Pro Lys Glu Trp His Ile Leu
275 280 285

<210> 57
<211> 288
<212> PRT
<213> Lachnospiraceae bacterium NK4A179

<400> 57

Met Arg Ser Val Val Asp Ile Lys Gly Gly Tyr Gly Asn Gln Leu Phe
1 5 10 15

Cys Tyr Ser Phe Gly Tyr Ala Val Ser Lys Glu Thr Gly Ser Glu Leu
20 25 30

Ile Ile Asp Thr Ser Met Leu Asp Met Asn Asn Val Lys Asp Arg Asn
35 40 45

Tyr Gln Leu Gly Val Leu Gly Ile Thr Tyr Asp Ser His Ile Ser Tyr
50 55 60

37847-517 Sequence Listing_ST25 (003)

Lys Tyr Gly Lys Asp Phe Leu Ser Arg Lys Thr Gly Leu Asn Arg Leu
 65 70 75 80
 Arg Lys Lys Ser Ala Ile Gly Phe Gly Thr Val Val Phe Lys Glu Lys
 85 90 95
 Glu Gln Tyr Val Tyr Asp Pro Ser Val Phe Glu Ile Lys Arg Asp Thr
 100 105 110
 Tyr Phe Asp Gly Phe Trp Gln Ser Ser Arg Tyr Phe Glu Lys Tyr Ser
 115 120 125
 Asp Asp Leu Arg Lys Met Leu Lys Pro Lys Lys Ile Ser Asn Ala Ala
 130 135 140
 Glu Lys Leu Ala Glu Asp Ala Arg Asp Cys Leu Ser Val Ser Val His
 145 150 155 160
 Ile Arg Arg Gly Asp Tyr Val Ser Leu Gly Trp Thr Leu Lys Asp Asp
 165 170 175
 Tyr Tyr Ile Lys Ala Leu Asp Ile Ile Lys Glu Arg Tyr Gly Ser Glu
 180 185 190
 Pro Val Phe Phe Val Phe Ser Asp Asn Lys Lys Tyr Ala Asp Asp Phe
 195 200 205
 Phe Ser Ala Ala Gly Leu Lys Tyr Arg Leu Met Asp Tyr Glu Thr Asp
 210 215 220
 Asp Ala Val Arg Asp Asp Met Phe Leu Met Ser Arg Cys Ser His Asn
 225 230 235 240
 Ile Met Ala Asn Ser Ser Tyr Ser Trp Trp Gly Ala Phe Leu Asn Asp
 245 250 255
 Asn Lys Asp Lys Thr Val Ile Cys Pro Glu Thr Gly Val Trp Gly Gly
 260 265 270

37847-517 Sequence Listing_ST25 (003)

Asp Phe Tyr Pro Glu Gly Trp Met Lys Val Thr Ala Ser Ser Gly Lys
 275 280 285

<210> 58
 <211> 291
 <212> PRT
 <213> Unknown

<220>
 <223> uncultured bacterium

<400> 58

Met Ile Ile Val Asn Leu Tyr Gly Gly Leu Gly Asn Gln Met Phe Gln
 1 5 10 15

Tyr Ala Leu Gly Arg His Leu Ala Glu Lys Asn Asn Thr Glu Leu Lys
 20 25 30

Leu Asp Ile Ser Ala Phe Glu Ser Tyr Lys Leu Arg Lys Tyr Glu Leu
 35 40 45

Gly Asn Leu Asn Ile Ile Glu Lys Phe Ala Leu Pro Glu Glu Ile Ser
 50 55 60

Arg Leu Ser Thr Leu Pro Thr Gly Lys Ile Glu Arg Phe Ile Arg Lys
 65 70 75 80

Thr Leu Arg Lys Pro Val Lys Lys Pro Glu Ser Tyr Ile Lys Glu Asn
 85 90 95

Ile Thr Gly Gly Phe Asn Pro Lys Ile Leu Asp Leu Gln Asn Asn Ile
 100 105 110

Tyr Leu Glu Gly Tyr Trp Gln Ser Glu Lys Tyr Phe Ile Glu Ile Glu
 115 120 125

Asp Ile Ile Arg Lys Glu Phe Ser Phe Lys Phe Pro Ala Thr Gly Lys
 130 135 140

Asn Lys Glu Ile Leu Glu Asn Ile Leu Asn Ile Asn Ser Val Ser Leu

37847-517 Sequence Listing_ST25 (003)

145 150 155 160

His Ile Arg Arg Gly Asp Tyr Val Thr Asn Pro Glu Val Asn Gln Val
165 170 175

His Gly Val Cys Ser Leu Asp Tyr Tyr Lys Ser Cys Val Asp Phe Ile
180 185 190

Glu Lys Lys Leu Glu Ser Pro Tyr Phe Tyr Ile Phe Ser Asp Asp Ile
195 200 205

Glu Trp Val Lys Asn Asn Leu Gln Ile Gln Ser Gln Val Tyr Tyr Val
210 215 220

Asp His Asn Thr Val Asp Asn Ala Ile Glu Asp Met Arg Leu Met Phe
225 230 235 240

Ser Cys Lys His Asn Ile Leu Ala Asn Ser Ser Phe Ser Trp Trp Gly
245 250 255

Ala Trp Leu Asn Ser Asn Pro Asp Lys Met Val Ile Thr Pro Arg Lys
260 265 270

Trp Phe Asn Thr Thr Tyr Asp Ser Asn Asp Leu Ile Pro Glu Arg Trp
275 280 285

Ile Lys Leu
290

<210> 59
<211> 760
<212> PRT
<213> Bacteroides fragilis

<400> 59

Met Lys Ile Gly Ile Ile Tyr Ile Val Thr Gly Pro Tyr Ile Lys Phe
1 5 10 15

Trp Asn Glu Phe Tyr Ser Ser Ser Gln Leu Tyr Phe Cys Val Glu Ala
20 25 30

37847-517 Sequence Listing_ST25 (003)

Glu Lys Asn Tyr Glu Val Phe Thr Asp Ser Ser Glu Leu Ala Ser Gln
35 40 45

Arg Leu Pro Asn Val His Met His Leu Ile Glu Asp Lys Gly Trp Ile
50 55 60

Val Asn Val Ser Ser Lys Ser Lys Phe Ile Cys Glu Ile Arg Asn Gln
65 70 75 80

Leu Thr Ser Tyr Asp Tyr Ile Phe Tyr Leu Asn Gly Asn Phe Lys Phe
85 90 95

Ile Ser Pro Ile Tyr Cys Asp Glu Ile Leu Pro Gln Ala Glu His Asn
100 105 110

Tyr Leu Thr Ala Leu Ser Phe Ser His Tyr Leu Thr Ile His Pro Asp
115 120 125

His Tyr Pro Tyr Asp Arg Asn Lys Asn Cys Asn Ala Phe Ile Pro Tyr
130 135 140

Gly Gln Gly Lys Tyr Tyr Phe Gln Gly Gly Phe Tyr Gly Gly Arg Thr
145 150 155 160

Gln Glu Val Leu Ser Leu Ser Glu Trp Cys Arg Asp Ala Ile Glu Ala
165 170 175

Asp Phe Asn Lys Lys Val Ile Ala Arg Phe His Asp Glu Ser Tyr Ile
180 185 190

Asn Arg Tyr Leu Leu Thr Gln His Pro Lys Val Leu Asn Asp Lys Tyr
195 200 205

Ala Phe Gln Asp Ile Trp Pro Tyr Glu Gly Glu Tyr Lys Ala Ile Val
210 215 220

Leu Asn Lys Glu Glu Val Pro Glu Asp Asn Asn Leu Gln Glu Met Lys
225 230 235 240

37847-517 Sequence Listing_ST25 (003)

Gln Asn Tyr Ile Asp Pro Ser Leu Ser Phe Leu Leu Asn Asp Glu Leu
245 250 255

Lys Phe Ile Pro Ile Ser Ile Val Gln Leu Tyr Gly Gly Leu Gly Asn
260 265 270

Gln Met Phe Gly Tyr Ala Phe Tyr Leu Tyr Ile Arg His Ile Ser Thr
275 280 285

Gln Glu Arg Lys Leu Leu Ile Asp Pro Ala Pro Cys Lys Arg Tyr Gly
290 295 300

Asn His Asn Gly Tyr Glu Leu Pro Ser Ile Phe Ser Lys Ile Cys Gln
305 310 315 320

Asp Ile His Ile Ser Asp Glu Thr Lys Asn Asn Ile Arg Lys Leu Arg
325 330 335

Lys Gly Thr Ser Leu Ser Ile Glu Glu Val Arg Ala Ser Met Pro Gln
340 345 350

Ser Phe Lys Glu Lys Lys Gln Pro Ile Ile Phe Tyr Ser Gly Cys Trp
355 360 365

Gln Cys Val Thr Tyr Val Glu Thr Val Lys Asp Glu Ile Lys Lys Asp
370 375 380

Phe Ile Phe Asp Glu Ser Lys Leu Asn Glu Pro Ser Ala Gln Met Leu
385 390 395 400

Arg Ile Ile Arg Arg Ser Asn Ser Val Ser Val His Ile Arg Arg Asn
405 410 415

Asp Tyr Leu Ile Gly Asn Asn Glu Phe Leu Tyr Gly Gly Ile Cys Thr
420 425 430

Lys Ser Tyr Tyr Glu Lys Ala Ile Ser Gln Met Tyr Thr Leu Leu Lys
435 440 445

37847-517 Sequence Listing_ST25 (003)

Asp Glu Pro Ile Phe Ile Tyr Phe Thr Asp Asp Pro Glu Trp Val Arg
450 455 460

Ser Asn Phe Ala Leu Asp Lys Ser Tyr Leu Val Asp Trp Asn Lys Asn
465 470 475 480

Lys Asp Asn Trp Gln Asp Met Tyr Leu Met Ser Ala Cys Arg His His
485 490 495

Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Ala Ala Trp Leu Gly Gly
500 505 510

Phe Pro Glu Lys Lys Val Ile Ala Pro Ser Thr Trp Leu Asn Gly Met
515 520 525

Gln Thr Pro Asp Ile Leu Pro Thr Glu Trp Ile Lys Ile Pro Ile Thr
530 535 540

Pro Asp Lys Lys Ile Leu Asp Arg Ile Cys Asn His Leu Ile Leu His
545 550 555 560

Ser Ser Tyr Met Lys Gln Leu Gly Leu Asn Ser Gly Lys Met Gly Val
565 570 575

Val Ile Phe Phe Phe His Tyr Ala Arg Tyr Thr Gln Asn Pro Leu Tyr
580 585 590

Glu Asn Tyr Ala Gly Asp Leu Phe Asp Glu Leu Tyr Glu Glu Ile His
595 600 605

Lys Gly Ile Ser Phe Ser Phe Leu Asp Gly Leu Cys Gly Ile Ala Trp
610 615 620

Ala Val Glu Tyr Leu Val His Glu Gln Phe Ile Glu Gly Asn Thr Asp
625 630 635 640

Asp Ser Leu Ala Glu Ile Asp Phe Lys Val Met Gln Ile Asp Pro Arg
645 650 655

37847-517 Sequence Listing_ST25 (003)

28 May 2018
2015259088

Arg Phe Thr Asp Tyr Ser Phe Glu Thr Gly Leu Glu Gly Ile Ala Cys
660 665 670

Tyr Val Leu Ser Arg Leu Leu Ser Pro Arg Val Cys Ser Ser Ser Leu
675 680 685

Thr Leu Asp Ser Val Tyr Leu Lys Asp Leu Thr Glu Ala Cys Arg Lys
690 695 700

Val Pro Val Asp Lys Ala Asn Tyr Thr Arg Leu Phe Leu Asn Tyr Ile
705 710 715 720

Glu Ser Lys Glu Val Gly Tyr Ser Phe Lys Asp Val Leu Met Gln Val
725 730 735

Leu Asn His Ser Glu Lys Ala Phe Gly Ser Asp Gly Leu Thr Trp Gln
740 745 750

Thr Gly Leu Thr Met Ile Met Arg
755 760

<210> 60

<211> 290

<212> PRT

<213> Butyrivibrio sp. AE3009

<400> 60

Met Ile Ile Ile Gln Leu Lys Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Leu Tyr Lys Glu Leu Arg Ser Arg Gly Lys Glu Val Lys Ile
20 25 30

Asp Asp Val Thr Gly Phe Val Asp Asp Glu Leu Arg Thr Pro Val Leu
35 40 45

Gln Arg Phe Gly Ile Glu Tyr Asp Arg Ala Thr Arg Glu Glu Val Val
50 55 60

37847-517 Sequence Listing_ST25 (003)

Lys Leu Thr Asp Ser Lys Met Asp Ile Phe Ser Arg Ile Arg Arg Lys
65 70 75 80

Leu Thr Gly Arg Lys Thr Cys Arg Ile Asp Glu Glu Ser Gly Thr Phe
85 90 95

Asn Pro Asp Ile Leu Glu Leu Asp Glu Ala Tyr Leu Val Gly Tyr Trp
100 105 110

Gln Ser Asp Lys Tyr Phe Arg Asn Glu Asp Val Ile Ala Gln Leu Arg
115 120 125

Gln Glu Phe Gln Lys Arg Pro Gln Glu Ile Met Thr Asp Ser Ala Ser
130 135 140

Trp Ala Thr Leu Gln Gln Ile Glu Cys Cys Gln Ser Val Ser Leu His
145 150 155 160

Ile Arg Arg Thr Asp Tyr Ile Asp Glu Glu His Asn His Ile His Asn
165 170 175

Leu Cys Thr Glu Lys Tyr Tyr Lys Gly Ala Ile Asp Arg Ile Arg Ser
180 185 190

Gln Tyr Pro Ser Ala Val Phe Phe Ile Phe Thr Asp Asp Lys Glu Trp
195 200 205

Cys Arg Asn His Phe Arg Gly Pro Asn Phe Phe Val Val Glu Leu Ala
210 215 220

Glu Lys Glu Asn Thr Asp Ile Ala Glu Met Leu Leu Met Ser Ser Cys
225 230 235 240

Lys His His Ile Cys Ala Asn Ser Ser Phe Ser Trp Trp Ser Ala Trp
245 250 255

Leu Asn Asp Ser Pro Glu Lys Met Val Ile Val Pro Asn Lys Trp Ile
260 265 270

37847-517 Sequence Listing_ST25 (003)

Asn Asn Arg Asp Met Asp Asp Ile Tyr Thr Asp Arg Met Thr Lys Met
275 280 285

Ala Ile
290

<210> 61
<211> 277
<212> PRT
<213> Bacteroides ovatus

<400> 61

Met Lys Gln Thr Ile Ile Leu Ser Gly Gly Leu Gly Asn Gln Met Phe
1 5 10 15

Gln Tyr Ala Phe Phe Leu Ser Met Lys Ala Lys Gly Lys Ser Cys Ser
20 25 30

Leu Asp Thr Thr Leu Phe Gln Thr Asn Lys Met His Asn Gly Phe Glu
35 40 45

Leu Lys Ser Val Phe Asp Ile Pro Asp Ser Pro Asn Gln Ala Ser Ala
50 55 60

Leu His Ser Leu Leu Ile Lys Met Leu Arg Arg Tyr Lys Pro Lys Ser
65 70 75 80

Ile Leu Thr Ile Asp Glu Pro Tyr Thr Phe Cys Pro Asp Ala Leu Glu
85 90 95

Ser Lys Lys Ser Phe Leu Met Gly Asp Trp Leu Ser Pro Lys Tyr Phe
100 105 110

Glu Ser Ile Lys Asp Val Val Val Asn Ala Tyr Arg Phe His Asn Ile
115 120 125

Gly Asn Lys Asn Val Asp Thr Ala Asn Glu Met His Gly Asn Asn Ser
130 135 140

37847-517 Sequence Listing_ST25 (003)

Val Ser Ile His Ile Arg Arg Gly Asp Tyr Leu Lys Leu Pro Tyr Tyr
145 150 155 160

Cys Val Cys Asn Glu Asn Tyr Tyr Arg Gln Ala Ile Glu Gln Ile Lys
165 170 175

Asp Arg Val Asp Asn Pro Ile Phe Tyr Val Phe Ser Asn Glu Pro Ser
180 185 190

Trp Cys Asp Ser Phe Met Lys Glu Phe Arg Val Asn Phe Lys Ile Val
195 200 205

Asn Trp Asn Gln Gly Lys Asp Ser Tyr Gln Asp Met Tyr Leu Met Thr
210 215 220

Gln Cys Lys His Asn Ile Ile Ala Asn Ser Thr Phe Ser Trp Trp Gly
225 230 235 240

Ala Trp Leu Asn Asn Asn Thr Asp Lys Ile Ile Val Ala Pro Ser Lys
245 250 255

Trp Phe Lys Asn Ser Glu His Asn Ile Asn Cys Lys Glu Trp Leu Leu
260 265 270

Ile Asp Thr Ser Lys
275

<210> 62
<211> 291
<212> PRT
<213> Desulfospira joergensenii

<400> 62

Met Gly Lys Lys Tyr Val Glu Thr Val Val Asn Gly Gly Leu Gly Asn
1 5 10 15

Gln Ile Phe Gln Phe Ser Ala Gly Phe Ala Leu Ser Lys Arg Leu Asn
20 25 30

Leu Asp Leu Val Leu Asn Ile Ser Thr Phe Asp Ser Cys Gln Lys Arg

37847-517 Sequence Listing_ST25 (003)

35	40	45																	
Asn	Phe	Glu	Leu	Tyr	Thr	Phe	Pro	Lys	Ile	Lys	Asn	Ser	Phe	Ala	Cys				
50						55					60								
Ile	Lys	Asp	Asp	Asp	Pro	Gly	Val	Phe	Ser	Arg	Leu	Arg	Ile	Pro	Phe				
65					70					75					80				
Leu	Asn	Phe	Lys	Glu	Lys	Ile	Lys	Gln	Phe	His	Glu	Ser	His	Phe	Phe				
				85					90					95					
Phe	Asp	Pro	Ala	Phe	Phe	Asp	Ile	Arg	Glu	Pro	Val	Arg	Ile	Glu	Gly				
			100					105					110						
Tyr	Phe	Gln	Ser	Tyr	Lys	Tyr	Phe	Glu	Lys	Tyr	Ser	Asp	Gln	Leu	Lys				
		115					120					125							
Asp	Ile	Leu	Leu	Asp	Ile	Pro	Leu	Thr	Ser	Arg	Leu	Lys	Thr	Val	Leu				
	130					135					140								
Lys	Val	Ile	Ser	Ser	Lys	Lys	Glu	Ser	Val	Ser	Val	His	Ile	Arg	Arg				
145					150					155					160				
Gly	Asp	Tyr	Ile	Ser	Asp	Gln	Gly	Ile	Asn	Glu	Val	His	Gly	Thr	Leu				
				165					170					175					
Asn	Glu	Ala	Tyr	Tyr	Leu	Asn	Ser	Ile	Lys	Leu	Met	Glu	Lys	Met	Phe				
			180					185					190						
Pro	Glu	Ser	Phe	Phe	Phe	Leu	Phe	Thr	Asp	Asp	Pro	His	Tyr	Val	Glu				
		195					200					205							
Glu	Asn	Phe	Lys	Phe	Leu	Glu	Asp	Thr	Ser	Cys	Ile	Ile	Ser	Asp	Asn				
	210					215					220								
Asp	Cys	Leu	Pro	Tyr	Glu	Asp	Met	Tyr	Leu	Met	Ala	Asn	Cys	His	His				
225					230					235					240				
Asn	Ile	Ile	Ala	Asn	Ser	Ser	Phe	Ser	Trp	Trp	Gly	Ala	Trp	Leu	Asn				

37847-517 Sequence Listing_ST25 (003)

245 250 255

Gln Asn Pro Glu Lys Ile Val Ile Ala Pro Arg Lys Trp Phe Ser Arg
260 265 270

Lys Ile Leu Met Glu Lys Pro Val Met Asp Leu Leu Pro Asp Asp Trp
275 280 285

Ile Leu Leu
290

<210> 63

<211> 303

<212> PRT

<213> Lachnospiraceae bacterium 10-1

<400> 63

Met Asn Ile Ile Arg Met Thr Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Leu Phe Leu Arg Leu Lys Ala Gln Gly Lys Glu Val Lys Phe
20 25 30

Asp Asp Arg Thr Glu Tyr Lys Gly Glu Glu Ala Arg Pro Ile Leu Leu
35 40 45

Trp Ala Phe Gly Ile Asp Tyr Pro Ala Ala Gly Glu Glu Glu Val Asn
50 55 60

Glu Leu Thr Asp Gly Val Met Lys Phe Ser His Arg Leu Arg Arg Lys
65 70 75 80

Leu Phe Gly Arg Lys Ser Lys Glu Tyr Arg Glu Lys Ser Cys Asn Phe
85 90 95

Asp Gln Gln Ile Leu Glu Lys Glu Pro Ala Tyr Phe Thr Gly Tyr Phe
100 105 110

Gln Ser Glu Arg Tyr Phe Glu Glu Val Lys Glu Gln Val Arg Lys Ala
115 120 125

37847-517 Sequence Listing_ST25 (003)

Phe Gln Phe Ser Gly Lys Ile Trp Gly Ser Val Ser Lys Glu Leu Glu
130 135 140

Glu Arg Ile Arg Glu Tyr Gln Thr Lys Ile Glu Asn Lys Ser Gln Met
145 150 155 160

Pro Val Ser Val His Ile Arg Arg Gly Asp Tyr Leu Glu Asn Asp Glu
165 170 175

Ala Tyr Gly Gly Ile Cys Thr Asp Ala Tyr Tyr Arg Lys Ala Ile Glu
180 185 190

Met Met Glu Glu Lys Phe Pro Asn Thr Val Phe Tyr Ile Phe Ser Asn
195 200 205

Asp Thr Gly Trp Ala Lys Gln Trp Ile Asp His Phe Tyr Lys Glu Lys
210 215 220

Ser Arg Phe Ile Val Ile Glu Gly Thr Thr Glu Asp Thr Gly Tyr Leu
225 230 235 240

Asp Leu Phe Leu Met Ser Lys Cys Arg Ala His Ile Ile Ala Asn Ser
245 250 255

Ser Phe Ser Trp Trp Gly Ala Trp Leu Asp Pro Asp Gln Glu Lys Ile
260 265 270

Val Ile Ala Pro Ser Lys Trp Val Asn Asn Gln Asp Met Lys Asp Ile
275 280 285

Tyr Thr Arg Glu Met Ile Lys Ile Ser Pro Lys Gly Glu Val Arg
290 295 300

<210> 64

<211> 289

<212> PRT

<213> Bacteroides dorei

<400> 64

37847-517 Sequence Listing_ST25 (003)

Met Val Val Val Tyr Val Gly Ala Gly Leu Ala Asn Arg Met Phe Gln
 1 5 10 15
 Tyr Ala Phe Ala Leu Ser Leu Arg Glu Lys Gly Leu Asp Val Phe Ile
 20 25 30
 Asp Glu Asp Ser Phe Ile Pro Arg Phe Asp Phe Glu Arg Thr Lys Leu
 35 40 45
 Asp Ser Val Phe Val Asn Val Asn Ile Gln Arg Cys Asp Lys Asn Ser
 50 55 60
 Phe Pro Leu Val Leu Arg Glu Asp Arg Phe Tyr Lys Leu Leu Lys Arg
 65 70 75 80
 Ile Ser Glu Tyr Met Ser Asp Asn Arg Tyr Ile Glu Arg Trp Asn Leu
 85 90 95
 Asp Tyr Leu Pro Tyr Ile His Lys Lys Ala Ser Thr Asn Cys Ile Phe
 100 105 110
 Ile Gly Phe Trp Ile Ser Tyr Lys Tyr Phe Gln Ser Ser Glu Asp Ala
 115 120 125
 Val Arg Lys Ala Phe Thr Phe Lys Pro Leu Asp Ser Ile Arg Asn Val
 130 135 140
 Glu Leu Ala Thr Lys Leu Val Thr Glu Asn Ser Val Ala Val His Phe
 145 150 155 160
 Arg Lys Asn Ile Asp Tyr Leu Lys Asn Leu Pro Asn Thr Cys Pro Pro
 165 170 175
 Ser Tyr Tyr Tyr Glu Ala Ile Asn Tyr Ile Lys Lys Tyr Val Pro Asn
 180 185 190
 Pro Lys Phe Tyr Phe Phe Ser Asp Asn Trp Asp Trp Val Arg Glu Asn
 195 200 205

37847-517 Sequence Listing_ST25 (003)

Ile Arg Gly Val Glu Phe Thr Ala Val Asp Trp Asn Pro Ser Ser Gly
210 215 220

Ile His Ser His Cys Asp Met Gln Leu Met Ser Leu Cys Lys His Asn
225 230 235 240

Ile Ile Ala Asn Ser Thr Tyr Ser Trp Trp Ser Ala Tyr Leu Asn Glu
245 250 255

Asn Asn Asn Lys Ile Val Val Cys Pro Lys Asp Trp Tyr Gly Gly Met
260 265 270

Val Lys Lys Leu Asp Thr Ile Ile Pro Glu Ser Trp Ile Ile Ile Asn
275 280 285

Gly

<210> 65

<211> 292

<212> PRT

<213> Firmicutes bacterium CAG:24

<400> 65

Met Val Ile Val Lys Met Ser Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Leu Tyr Arg Lys Ile Gln Gln Thr Gly Lys Asp Val Lys Leu
20 25 30

Asp Leu Phe Ser Phe Gln Asp Lys Asn Ala Phe Arg Arg Phe Ser Leu
35 40 45

Asp Ile Phe Pro Ile Glu Tyr Gln Thr Ala Asn Leu Glu Glu Cys Arg
50 55 60

Lys Leu Gly Glu Cys Ser Tyr Arg Pro Val Asp Lys Ile Arg Arg Lys
65 70 75 80

37847-517 Sequence Listing_ST25 (003)

Met	Phe	Gly	Leu	Lys	Glu	Ser	Tyr	Tyr	Gln	Glu	Asp	Leu	Asp	Lys	Gly	85	90	95
Tyr	Gln	Pro	Glu	Ile	Leu	Glu	Met	Asn	Pro	Val	Tyr	Leu	Asp	Gly	Tyr	100	105	110
Trp	Gln	Cys	Glu	Arg	Tyr	Phe	Gln	Asp	Ile	Arg	Glu	Lys	Ile	Leu	Glu	115	120	125
Asp	Tyr	Thr	Phe	Pro	Lys	Lys	Ile	Ser	Ile	Glu	Ser	Ser	Arg	Leu	Gln	130	135	140
Glu	Arg	Ile	Lys	Asn	Thr	Glu	Ser	Val	Ser	Ile	His	Ile	Arg	Arg	Gly	145	150	155
Asp	Tyr	Leu	Asp	Ala	Ala	Asn	Tyr	Lys	Ile	Tyr	Gly	Asn	Ile	Cys	Thr	165	170	175
Ile	Glu	Tyr	Tyr	Gln	Ser	Ala	Ile	Ser	Arg	Met	Arg	Lys	Leu	Cys	Glu	180	185	190
Lys	Pro	Asn	Phe	Tyr	Leu	Phe	Ser	Asn	Asp	Pro	Glu	Trp	Ala	Lys	Glu	195	200	205
Ile	Phe	Gly	Asp	Thr	Glu	Asp	Ile	Thr	Ile	Val	Glu	Glu	Asp	Lys	Glu	210	215	220
Arg	Pro	Asp	Tyr	Glu	Asp	Met	Phe	Leu	Met	Ser	Arg	Cys	Lys	His	Asn	225	230	235
Ile	Ile	Ala	Asn	Ser	Ser	Phe	Ser	Trp	Trp	Ala	Ala	Trp	Leu	Asn	Gln	245	250	255
Asn	Glu	Asn	Lys	Arg	Val	Ile	Ala	Pro	Val	Lys	Trp	Phe	Asn	Asn	His	260	265	270
Ser	Val	Thr	Asp	Val	Ile	Cys	Asp	Asp	Trp	Ile	Arg	Ile	Asp	Gly	Asp	275	280	285

37847-517 Sequence Listing_ST25 (003)

His Lys Gly Ala
290

<210> 66

<211> 294

<212> PRT

<213> Clostridium hathewayi CAG:224

<400> 66

Met Ile Tyr Val Asn Ile Arg Gly Arg Leu Gly Asn Gln Leu Phe Ile
1 5 10 15

Tyr Ala Phe Ala Arg Ala Leu Gln Lys Ser Thr Asn Gln Gln Ile Thr
20 25 30

Leu Asn Tyr Thr Ser Phe Arg Lys His Tyr Asn Asn Thr Ala Met Asp
35 40 45

Leu Glu Gln Phe Asn Ile Pro Glu Asp Ile Met Phe Glu Asn Ser Lys
50 55 60

Glu Leu Pro Trp Phe Ala Asn Thr Asp Gly Lys Val Ile Arg Ile Leu
65 70 75 80

Arg His Tyr Phe Pro Lys Leu Ile Arg Ser Ile Leu Gln Lys Met Asn
85 90 95

Val Leu Met Trp Leu Gly Asp Glu Tyr Val Glu Val Lys Val Asn Lys
100 105 110

Arg Arg Asp Ile Tyr Ile Asp Gly Phe Trp Gln Ser Ser Arg Tyr Phe
115 120 125

Lys Ser Val Tyr Lys Glu Leu Lys Asn Glu Leu Ile Pro Lys Met Glu
130 135 140

Met Ser Lys Glu Ile Lys Thr Met Gly Asp Leu Ile Asn Gln Lys Glu
145 150 155 160

Ser Val Cys Val Ser Val Arg Arg Gly Asp Tyr Val Thr Val Lys Lys

37847-517 Sequence Listing_ST25 (003)

165 170 175

Asn Arg Asp Val Tyr Tyr Ile Cys Asp Glu Lys Tyr Leu Asn Thr Ser
180 185 190

Ile Met Arg Met Val Glu Leu Val Pro Asn Val Thr Trp Phe Ile Phe
195 200 205

Ser Asp Asp Ala Asp Trp Val Lys Asp Asn Ile Val Phe Pro Gly Glu
210 215 220

Val Phe Tyr Gln Pro Pro Arg Val Thr Pro Leu Glu Thr Leu Tyr Leu
225 230 235 240

Met Lys Ala Cys Lys His Phe Ile Ile Ser Asn Ser Ser Phe Ser Trp
245 250 255

Trp Gly Gln Tyr Leu Ser Asn Asn Asp Asn Lys Ile Val Ile Gly Pro
260 265 270

Ala Lys Trp Tyr Val Asp Gly Arg Lys Thr Asp Ile Ile Glu Glu Glu
275 280 285

Trp Ile Lys Ile Glu Val
290

<210> 67

<211> 294

<212> PRT

<213> Syntrophus acidotrophicus

<400> 67

Met Val Ile Val Arg Leu Thr Gly Gly Ile Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Ala Ala Arg Arg Val Ser Leu Val Asn Asn Ala Pro Leu Phe
20 25 30

Leu Asp Leu Gly Trp Phe Gln Glu Thr Gly Ser Trp Thr Pro Arg Lys
35 40 45

37847-517 Sequence Listing_ST25 (003)

Tyr Glu Leu Asp Ala Phe Arg Ile Ala Gly Glu Ser Ala Ser Val Gly
 50 55 60
 Asp Ile Lys Asp Phe Lys Ser Arg Arg Gln Asn Ala Phe Phe Arg Arg
 65 70 75 80
 Leu Pro Leu Phe Leu Lys Lys Arg Ile Phe His Thr Arg Gln Thr His
 85 90 95
 Ile Ile Glu Lys Ser Tyr Asn Phe Asp Pro Glu Ile Leu Asn Leu Gln
 100 105 110
 Gly Asn Val Tyr Leu Asp Gly Tyr Trp Gln Ser Glu Lys Tyr Phe Ser
 115 120 125
 Asp Val Asp Ser Glu Ile Arg Arg Glu Phe Ser Phe Gln Thr Asp Pro
 130 135 140
 Ala Glu Arg Asn Arg Lys Ile Leu Glu Arg Ile Ala Ser Cys Glu Ser
 145 150 155 160
 Val Ser Ile His Ile Arg Arg Gly Asp Tyr Val Thr Leu Pro Asp Ala
 165 170 175
 Asn Ala Phe His Gly Leu Cys Thr Pro Ala Tyr Tyr Arg Leu Ala Val
 180 185 190
 Glu Gln Ile Ser Arg Lys Val Val Glu Pro Val Phe Phe Val Phe Ser
 195 200 205
 Asp Asp Ile Ala Trp Ala Arg Gly Asn Leu Lys Leu Gly Phe Glu Thr
 210 215 220
 Cys Phe Met Asp Gln Asn Gly Pro Asp Arg Gly Asp Glu Asp Leu Arg
 225 230 235 240
 Leu Met Ile Ala Cys Arg His His Ile Ile Ala Asn Ser Ser Phe Ser
 245 250 255

37847-517 Sequence Listing_ST25 (003)

Trp Trp Gly Ala Trp Leu Cys Ser Asn Pro Glu Lys Ile Val Tyr Ala
260 265 270

Pro Arg Lys Trp Phe Asn Asn Gly Leu Asp Thr Pro Asp Asn Ile Pro
275 280 285

Ala Ser Trp Ile Arg Ile
290

<210> 68

<211> 290

<212> PRT

<213> Bacteroides caccae

<400> 68

Met Lys Ile Val Lys Ile Ile Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Leu Tyr Leu Ser Leu Lys Lys Lys Tyr Pro Lys Glu Lys Ile
20 25 30

Lys Ile Asp Ile Ser Met Phe Glu Thr Tyr Gly Leu His Asn Gly Phe
35 40 45

Glu Leu Lys Arg Ile Phe Asp Ile Asp Ala Glu Tyr Ala Ser Arg Glu
50 55 60

Glu Ile Arg Glu Leu Ser Phe Tyr Ile Lys Ile Tyr Lys Leu Gln Arg
65 70 75 80

Ile Phe Arg Lys Ile Phe Pro Val Arg Lys Thr Glu Cys Val Glu Lys
85 90 95

Tyr Asp Phe Lys Phe Met Ser Glu Val Trp Ser Asn Cys Asp Arg Tyr
100 105 110

Tyr Glu Gly Tyr Trp Gln Asn Trp Glu Tyr Phe Ile Glu Ala Gln Thr
115 120 125

37847-517 Sequence Listing_ST25 (003)

Glu Val Arg Ser Thr Phe Thr Phe Lys Lys Glu Leu Val Gly Arg Asn
130 135 140

Ala Lys Val Ile Arg Glu Ile Gln Tyr Ala Lys Met Pro Val Ser Leu
145 150 155 160

His Ile Arg Arg Gly Asp Tyr Leu His His Lys Leu Phe Gly Gly Leu
165 170 175

Cys Asp Leu Asn Tyr Tyr Lys Lys Ala Ile Asp Tyr Val Leu Asn Asn
180 185 190

Tyr Asp Thr Pro Gln Phe Tyr Leu Phe Ser Asn Asp Ile Glu Trp Cys
195 200 205

Lys Thr Tyr Ile Leu Pro Leu Val Gln Gly Tyr Pro Phe Ile Leu Val
210 215 220

Asp Trp Asn Ser Gly Val Glu Ser Tyr Ile Asp Met Gln Leu Met Ser
225 230 235 240

Cys Cys Arg Ile Asn Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Ala
245 250 255

Ala Trp Leu Asn Asp Ser Ser Glu Lys Ile Val Ile Ala Pro Lys Leu
260 265 270

Trp Ala His Ser Pro Tyr Gly Lys Glu Ile Gln Leu Lys Ser Trp Leu
275 280 285

Leu Phe
290

<210> 69

<211> 291

<212> PRT

<213> Butyrivibrio fibrisolvens

<400> 69

37847-517 Sequence Listing_ST25 (003)

Met	Ile	Ile	Ile	Glu	Met	Ser	Gly	Gly	Leu	Gly	Asn	Gln	Met	Phe	Gln	1	5	10	15
Tyr	Ala	Leu	Tyr	Lys	Ser	Met	Leu	His	Lys	Gly	Leu	Asp	Val	Thr	Ile	20	25	30	
Asp	Lys	Ser	Ile	Tyr	Arg	Asp	Val	Asp	His	Lys	Glu	Gln	Val	Asp	Leu	35	40	45	
Asp	Arg	Phe	Pro	Asn	Val	Ser	Tyr	Ile	Glu	Ala	Asp	Arg	Lys	Leu	Ser	50	55	60	
Ser	Thr	Leu	Arg	Gly	Tyr	Gly	Tyr	Asn	Asp	Ser	Ile	Ile	Asp	Lys	Ile	65	70	75	80
Arg	Asn	Lys	Leu	Asn	Lys	Ser	Lys	Arg	Asn	Leu	Tyr	His	Glu	Asp	Leu	85	90	95	
Asp	Lys	Gly	Tyr	Gln	Pro	Glu	Ile	Phe	Glu	Phe	Asp	Asn	Val	Tyr	Leu	100	105	110	
Asn	Gly	Tyr	Trp	Gln	Cys	Glu	Arg	Tyr	Phe	Lys	Asp	Ile	Lys	Asn	Glu	115	120	125	
Ile	Lys	Lys	Asp	Phe	Ile	Phe	Pro	Cys	Thr	Gln	Ser	Gly	Asp	Asp	Lys	130	135	140	
Ile	Lys	Ala	Leu	Thr	Ile	Glu	Met	Glu	Ser	Cys	Asn	Ser	Val	Ser	Leu	145	150	155	160
His	Val	Arg	Arg	Gly	Asp	Tyr	Leu	Lys	Pro	Gly	Leu	Ile	Glu	Ile	Tyr	165	170	175	
Gly	Asn	Ile	Cys	Thr	Glu	Glu	Tyr	Tyr	Lys	Lys	Ser	Ile	Glu	Tyr	Ile	180	185	190	
Lys	Glu	Arg	Val	Asp	Asn	Pro	Val	Phe	Tyr	Ile	Phe	Ser	Asn	Asp	Met	195	200	205	

37847-517 Sequence Listing_ST25 (003)

Ala Trp Val Arg Asp Asn Phe Lys Ser Asp Asp Phe Arg Tyr Val Asn
210 215 220

Glu Asp Gly Ala Phe Asp Gly Met Thr Asp Met Tyr Leu Met Thr Arg
225 230 235 240

Cys Arg His Asn Ile Val Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala
245 250 255

Trp Leu Asn Lys His Asp Asp Asn Ile Val Ile Cys Pro Asn Arg Trp
260 265 270

Val Asn Thr His Thr Val Thr Asp Ile Ile Cys Glu Asp Trp Ile Arg
275 280 285

Ile Asp Val
290

<210> 70
<211> 309
<212> PRT
<213> Parabacteroides distasonis

<400> 70

Met Ile Val Gly Gly Asn Asp Tyr Cys Lys Val Lys Val Val Asn Ile
1 5 10 15

Ile Gly Gly Leu Gly Asn Gln Met Phe Gln Tyr Ala Phe Ala Leu Ser
20 25 30

Leu Lys Glu His Phe Pro Lys Glu Glu Ile Arg Ile Asp Ile Ser His
35 40 45

Phe Asn Tyr Leu Phe Val Asn Lys Val Gly Ala Ala Asn Leu His Asn
50 55 60

Gly Tyr Glu Leu Asp Lys Ile Phe Phe Asn Ile Glu Leu Lys Lys Ala
65 70 75 80

Asn Ala Trp Gln Leu Met Lys Leu Thr Trp Phe Ile Pro Asn Tyr Leu

28 May 2018
2015259088

37847-517 Sequence Listing_ST25 (003)

85

90

95

Ile Ser Arg Ile Ala Arg Lys Ile Leu Pro Val Arg Asn Ser Glu Tyr
100 105 110

Ile Gln Asn Ser Ser Asp Cys Phe Phe Tyr Asp Pro Met Val Tyr Asn
115 120 125

Lys Gln Gly Ser Cys Tyr Tyr Glu Gly Tyr Trp Gln Ala Ile Gly Tyr
130 135 140

Tyr Glu Ser Met Arg Asp Lys Leu Cys Lys Ile Phe Gln His Pro Ser
145 150 155 160

Pro Glu Gly Lys Asn Lys Gln Tyr Ile Glu Asn Met Glu Ser Ser Asn
165 170 175

Ser Val Gly Ile His Ile Arg Arg Gly Asp Tyr Leu Leu Ser Asp Asn
180 185 190

Phe Arg Gly Ile Cys Glu Val Asp Tyr Tyr Lys Arg Ala Ile Asp Lys
195 200 205

Ile Leu Gln Asp Gly Glu Lys His Val Phe Tyr Leu Phe Ser Asn Asp
210 215 220

Gln Lys Trp Cys Glu Glu Tyr Ile Leu Pro Leu Leu Gly Asn Tyr Glu
225 230 235 240

Ile Ile Phe Val Thr Gly Asn Ile Gly Arg Asp Ser Cys Trp Asp Met
245 250 255

Phe Leu Met Thr His Cys Lys Asp Leu Ile Ile Ala Asn Ser Ser Phe
260 265 270

Ser Trp Trp Gly Ala Phe Leu Asn Lys Arg Gly Gly Arg Val Val Thr
275 280 285

Pro Lys Arg Trp Met Asn Arg Asn Ile Arg Tyr Asp Leu Trp Met Pro

37847-517 Sequence Listing_ST25 (003)

290

295

300

Glu Trp Ile Arg Ile
305

<210> 71

<211> 295

<212> PRT

<213> Geobacter uraniireducens

<400> 71

Met Ile Ile Ala Arg Leu Gln Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Val Gly Leu His Leu Ala Leu Thr His Asn Val Glu Leu Lys
20 25 30

Ile Asp Ile Thr Met Phe Ser Asp Tyr Lys Trp His Thr Tyr Ser Leu
35 40 45

Arg Pro Phe Asn Ile Arg Glu Ser Ile Ala Thr Glu Glu Glu Ile Lys
50 55 60

Ala Leu Thr Asp Val Lys Met Asp Arg Pro Tyr Lys Lys Ile Asp Asn
65 70 75 80

Phe Leu Cys Arg Leu Leu Arg Lys Ser Gln Lys Ile Ser Ala Thr His
85 90 95

Val Lys Glu Lys His Phe His Tyr Asp Pro Asp Ile Leu Lys Leu Pro
100 105 110

Asp Asn Val Tyr Leu Asp Gly Tyr Trp Gln Ser Glu Lys Tyr Phe Lys
115 120 125

Glu Ile Glu Asn Ile Ile Arg Gln Thr Phe Ile Ile Lys Asn Pro Gln
130 135 140

Leu Gly Arg Asp Lys Glu Leu Ala Cys Lys Ile Leu Ser Thr Glu Ser
145 150 155 160

37847-517 Sequence Listing_ST25 (003)

Val Cys Leu His Ile Arg Arg Gly Asn Tyr Val Thr Asp Lys Thr Thr
165 170 175

Asn Ser Val Leu Gly Pro Cys Asp Leu Ser Tyr Tyr Ser Asn Cys Ile
180 185 190

Lys Ser Leu Ala Gly Asn Asn Lys Asp Pro His Phe Phe Val Phe Ser
195 200 205

Asn Asp His Glu Trp Val Ser Lys Asn Leu Lys Leu Asp Tyr Pro Thr
210 215 220

Ile Tyr Val Asp His Asn Asn Glu Asp Lys Asp Tyr Glu Asp Leu Arg
225 230 235 240

Leu Met Ser Gln Cys Lys His His Ile Ile Ala Asn Ser Thr Phe Ser
245 250 255

Trp Trp Ser Ala Trp Leu Cys Ser Asn Pro Asp Lys Val Ile Tyr Ala
260 265 270

Pro Gln Lys Trp Phe Arg Val Asp Glu Tyr Asn Thr Lys Asp Leu Leu
275 280 285

Pro Ser Asn Trp Leu Ile Leu
290 295

<210> 72

<211> 293

<212> PRT

<213> Lachnospiraceae bacterium A4

<400> 72

Met Ile Ile Val Lys Ile Tyr Glu Gly Leu Gly Asn Gln Leu Phe Gln
1 5 10 15

Tyr Ala Phe Ala Arg Ser Ile Gln Val Asn Gly Lys Lys Val Phe Leu
20 25 30

37847-517 Sequence Listing_ST25 (003)

Asp Thr Ser Gly Tyr Thr Asp Gln Leu Phe Pro Leu Cys Arg Thr Ser
 35 40 45
 Thr Arg Arg Arg Tyr Gln Leu Asn Cys Phe Asn Ile Arg Ile Lys Glu
 50 55 60
 Val Glu Lys Lys Asn Ile Glu Lys Tyr Ser Phe Leu Ile Gln Glu Asp
 65 70 75 80
 Met Phe Gly Lys Leu Ile Ser Lys Leu Ala Lys Leu His Leu Trp Met
 85 90 95
 Tyr Lys Val Thr Ile Gln Gln Asn Ala Gln Glu Tyr Lys Glu Ser Tyr
 100 105 110
 Leu Asn Thr Arg Gly Asn Val Tyr Tyr Lys Gly Trp Phe Gln Asn Pro
 115 120 125
 Lys Tyr Phe Ser Ser Ile Arg Arg Leu Leu Leu Lys Glu Ile Thr Pro
 130 135 140
 Lys Tyr Lys Ile Arg Ile Pro Ala Glu Leu Arg Glu Leu Leu Gln Glu
 145 150 155 160
 Asp Asn Ile Val Ala Val His Cys Arg Arg Gly Asp Tyr Gln Tyr Ile
 165 170 175
 Arg Asn Cys Leu Pro Val Asn Tyr Tyr Lys Lys Ala Met Ala Tyr Met
 180 185 190
 Glu Lys Lys Leu Gly Val Pro Arg Tyr Leu Phe Phe Ser Asp Asp Leu
 195 200 205
 Ser Trp Val Lys Arg Gln Phe Gly Asn Lys Asp Asn Asn Tyr Tyr Ile
 210 215 220
 Glu Asp Tyr Gly Lys Phe Glu Asp Tyr Gln Glu Leu Met Ile Met Ser
 225 230 235 240

37847-517 Sequence Listing_ST25 (003)

Arg Cys Arg Asn Phe Ile Ile Ala Asn Ser Thr Phe Ser Trp Trp Ala
245 250 255

Ala Trp Leu Cys Ser Tyr Glu Asn Lys Val Val Ile Met Pro Arg Val
260 265 270

Trp Thr Tyr Val Gly Gly Gln Gly Val Glu Met Ser Asp Phe Pro Ala
275 280 285

Asp Trp Ile Arg Ile
290

<210> 73

<211> 294

<212> PRT

<213> Colwellia psychroerythrus

<400> 73

Met Lys Val Val Arg Val Cys Gly Gly Phe Gly Asn Gln Leu Phe Gln
1 5 10 15

Tyr Ala Phe Tyr Leu Ala Val Lys His Lys Phe Asn Glu Thr Thr Lys
20 25 30

Leu Asp Ile His Asp Met Ala Ser Tyr Glu Leu His Asn Gly Tyr Glu
35 40 45

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

Lys Leu Ala Val Gln Ser Thr Lys Asn Ile Phe Thr Lys Leu Leu Lys
65 70 75 80

Glu Ile Lys Lys Tyr Thr Pro Phe Ile Pro Arg Thr Tyr Ile Lys Glu
85 90 95

Lys Lys His Leu His Phe Ser Tyr Gln Glu Val Asp Leu Gly Thr Lys
100 105 110

37847-517 Sequence Listing_ST25 (003)

Asp Thr Ser Ile Tyr Tyr Arg Gly Ser Trp Gln Asn Pro Gln Tyr Phe
115 120 125

Asn Ser Ile Ala Ser Glu Ile Arg Glu Lys Leu Thr Phe Pro Glu Phe
130 135 140

Thr Glu Pro Lys Ser Leu Ala Leu His Gln Glu Ile Ser Glu His Glu
145 150 155 160

Thr Val Ala Val His Ile Arg Arg Gly Asp Tyr Leu Lys His Lys Ala
165 170 175

Leu Gly Gly Ile Cys Asp Leu Pro Tyr Tyr Gln Asn Ala Ile Lys Glu
180 185 190

Ile Glu Gly Leu Val Glu Lys Pro Leu Phe Val Ile Phe Ser Asp Asp
195 200 205

Ile Thr Trp Cys Arg Ala Asn Ile Asn Val Glu Lys Val Arg Phe Val
210 215 220

Asp Trp Asn Ser Gly Glu Gln Ser Phe Gln Asp Met His Leu Met Ser
225 230 235 240

Leu Cys Thr His Asn Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Gly
245 250 255

Ala Trp Leu Asn Ala Asn Pro Asn Lys Ile Val Ile Ser Pro Asn Lys
260 265 270

Trp Ile His Tyr Thr Asp Ser Met Gly Ile Val Pro Ser Glu Trp Ile
275 280 285

Lys Val Glu Thr Ser Ile
290

<210> 74
<211> 282
<212> PRT
<213> Roseobacter sp. MED193

37847-517 Sequence Listing_ST25 (003)

<400> 74

Met Ile Thr Ser Arg Leu His Gly Arg Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Ala Ala Arg Ala Leu Ala His Arg Leu Gly Cys Gly Val Ala
20 25 30

Leu Asp Gly Arg Gly Ala Glu Leu Arg Gly Glu Gly Val Leu Thr Arg
35 40 45

Val Phe Asp Leu Pro Leu Ser Ala Ala Pro Lys Leu Pro Pro Leu Lys
50 55 60

Gln His Ala Pro Leu Arg Tyr Gly Leu Trp Arg Gly Leu Gly Leu Ala
65 70 75 80

Pro Arg Phe Arg Arg Glu Arg Gly Leu Gly Tyr Asn Thr Ala Phe Glu
85 90 95

Thr Trp Glu Asp Gly Cys Tyr Leu His Gly Tyr Trp Gln Ser Glu Arg
100 105 110

Tyr Phe Glu Glu Ile Ser Asp Leu Ile Arg Ala Asp Phe Thr Phe Pro
115 120 125

Asp Phe Ser Asn Arg Gln Asn Ala Glu Met Ala Ala Arg Ile Met Glu
130 135 140

Asp Asn Ala Ile Ser Leu His Val Arg Arg Gly Asp Tyr Val Ala Leu
145 150 155 160

Ser Ala His Val Leu Cys Asp Gln Ala Tyr Tyr Glu Ala Ala Leu Thr
165 170 175

Arg Leu Leu Glu Gly Leu Ser Gln Asp Ala Pro Thr Val Tyr Val Phe
180 185 190

Ser Asp Asp Pro Asp Trp Ala Lys Ala Asn Leu Pro Leu Pro Cys Lys

37847-517 Sequence Listing_ST25 (003)

195

200

205

Lys Val Val Val Asp Phe Asn Gly Pro Glu Thr Asp Phe Glu Asp Met
210 215 220

Arg Leu Met Ser Leu Cys Lys His Asn Ile Ile Gly Asn Ser Ser Phe
225 230 235 240

Ser Trp Trp Ala Ala Trp Leu Asn Ala Asn Pro Gln Lys Arg Val Ala
245 250 255

Gly Pro Ala Asn Trp Phe Gly Asp Pro Lys Leu Ser Asn Pro Asp Ile
260 265 270

Leu Pro Ser Gln Trp Leu Lys Val Ala Pro
275 280

<210> 75

<211> 296

<212> PRT

<213> Cesiribacter andamanensis

<400> 75

Met Met Ile Val Arg Leu Cys Gly Gly Leu Gly Asn Gln Leu Phe Gln
1 5 10 15

Tyr Ala Val Gly Lys Gln Leu Ser Val Lys Asn Asn Ile Pro Leu Lys
20 25 30

Ile Asp Asp Ser Trp Leu Arg Leu Pro Asp Ala Arg Lys Tyr Arg Leu
35 40 45

Gln Phe Phe Gln Ile Glu Glu Pro Leu Ala Ser Pro Gln Glu Val Glu
50 55 60

Arg Phe Val Gly Pro Tyr Glu Ser Gln Ser Leu Tyr Ala Arg Leu Tyr
65 70 75 80

Arg Lys Val Gln Asn Met Leu Pro Arg His Arg Arg Arg Tyr Phe Gln
85 90 95

37847-517 Sequence Listing_ST25 (003)

Glu Ser Gly Phe Trp Ala Tyr Glu Pro Glu Leu Met Arg Ile Arg Ser
100 105 110

Gln Val Phe Leu Glu Gly Phe Trp Gln His His Ala Tyr Phe Thr Arg
115 120 125

Leu His Pro Gln Val Leu Glu Ala Leu Gln Leu Arg Glu Glu Tyr Arg
130 135 140

Gln Glu Pro Tyr Ala Val Leu Asp Gln Ile Arg Glu Asp Ala Ala Ser
145 150 155 160

Val Ser Leu His Ile Arg Arg Gly Asp Tyr Val Ser Asp Pro Tyr Asn
165 170 175

Leu Gln Phe Phe Gly Val Met Pro Leu Ser Tyr Tyr Gln Gln Ala Val
180 185 190

Ala Tyr Met Gln Glu Gln Leu His Ala Pro Thr Phe Tyr Ile Phe Ser
195 200 205

Asp Asp Leu Asp Trp Ala Arg Ala His Leu Lys Leu Gln Ala Pro Met
210 215 220

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

Ala Met Arg Leu Cys Arg His Asn Ile Leu Ala Asn Ser Ser Phe Ser
245 250 255

Trp Trp Gly Ala Tyr Leu Asn Thr Asn Pro His Lys Arg Val Ile Ala
260 265 270

Pro Arg Gln Trp Val Ala Asp Pro Glu Leu Lys Asp Lys Val Gln Ile
275 280 285

Gln Met Pro Asp Trp Ile Leu Leu
290 295

37847-517 Sequence Listing_ST25 (003)

28 May 2018
2015259088

<210> 76
 <211> 309
 <212> PRT
 <213> Rhodopirellula salentina
 <400> 76

Met Ile Ala Thr Arg Leu Ile Gly Gly Leu Gly Asn Gln Met Phe Gln
 1 5 10 15

Tyr Ala Tyr Gly Phe Ser Leu Ala Arg Arg Arg Ser Glu Arg Leu Val
 20 25 30

Leu Asp Val Ser Ala Phe Glu Ser Tyr Asp Leu His Ala Leu Ala Ile
 35 40 45

Asp Gln Phe Asp Ile Ser Ala Ala Arg Met Thr Gln Ala Glu Phe Ala
 50 55 60

Arg Ile Pro Gly Arg Tyr Arg Gly Lys Ser Arg Trp Ala Glu Arg Val
 65 70 75 80

Ala Asn Phe Ala Gly Gly Leu Gln Ser Cys Asp Lys Arg Pro Leu Arg
 85 90 95

Leu Arg Arg Glu Lys Pro Phe Gly Phe Ala Glu Lys Tyr Leu Ala Glu
 100 105 110

Gly Ser Asp Leu Tyr Leu Asp Gly Tyr Trp Gln Ser Glu Arg Tyr Phe
 115 120 125

Pro Gly Leu Gln Ala Glu Leu Lys Lys Glu Phe Gln Leu Lys Arg Gly
 130 135 140

Leu Ser Asp Glu Ser Ser Arg Val Leu Asp Glu Ile Gln Ser Ser Met
 145 150 155 160

Ser Val Ala Met His Val Arg Arg Gly Asp Tyr Val Thr Asn Ala Glu
 165 170 175

37847-517 Sequence Listing_ST25 (003)

Thr Leu Arg Ile Tyr Arg Arg Leu Asp Ala Glu Tyr Tyr Arg Lys Cys
180 185 190

Leu Asn Asp Leu Arg Gln Arg Phe Ser Asn Leu Asn Val Phe Val Phe
195 200 205

Ser Asn Asp Ile Gln Trp Cys Gln Asp His Leu Asp Val Gly Leu Lys
210 215 220

Gln Arg Pro Val Thr His Asn Asp Ala Thr Thr Ala Ile Glu Asp Met
225 230 235 240

Phe Leu Met Ser Gln Cys Asp His Ser Ile Ile Ala Asn Ser Ser Phe
245 250 255

Ser Trp Trp Ala Ala Phe Leu Gly Arg Ser Asp Ala Gln Arg Arg Val
260 265 270

Tyr Tyr Pro Asp Pro Trp Phe Asn Pro Gly Thr Leu Asn Gly Asp Ser
275 280 285

Leu Gly Cys Ala Asn Trp Val Ser Glu Ser Ser Ile Ser Val Ser Arg
290 295 300

Pro Ser Arg Ala Ala
305

<210> 77

<211> 289

<212> PRT

<213> Butyrivibrio sp. AD3002

<400> 77

Met Ile Ile Ile Arg Met Met Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Leu Tyr Leu Gln Leu Lys Ala Leu Gly Lys Glu Val Lys Ile
20 25 30

37847-517 Sequence Listing_ST25 (003)

Asp	Asp	Val	Tyr	Gly	Phe	Arg	Asp	Asp	Pro	Gln	Arg	Asp	Pro	Val	Leu
		35					40					45			
Glu	Lys	Met	Tyr	Gly	Ile	Thr	Tyr	Thr	Lys	Ala	Ser	Asp	Ala	Glu	Val
	50					55					60				
Val	Asp	Ile	Thr	Asp	Ser	His	Leu	Asp	Ile	Phe	Ser	Arg	Ile	Arg	Arg
65					70					75					80
Lys	Leu	Phe	Gly	Arg	Lys	Ser	His	Glu	Tyr	Ile	Glu	Glu	Thr	Gly	Leu
				85					90					95	
Phe	Asp	Pro	Lys	Val	Phe	Glu	Phe	Glu	Thr	Ala	Tyr	Leu	Asn	Gly	Tyr
			100					105					110		
Phe	Gln	Ser	Asp	Lys	Tyr	Phe	Pro	Asp	Lys	Glu	Val	Leu	Ala	Gln	Leu
		115					120					125			
Arg	Arg	Glu	Phe	Val	Ile	Lys	Pro	Asp	Asp	Val	Phe	Thr	Ser	Ala	Asp
	130					135					140				
Ser	Trp	Glu	Leu	Tyr	Arg	Gln	Ile	Arg	Glu	Thr	Glu	Ser	Val	Ser	Ile
145					150					155					160
His	Val	Arg	Arg	Gly	Asp	Tyr	Leu	Leu	Pro	Gly	Thr	Val	Glu	Thr	Phe
				165					170					175	
Gly	Gly	Ile	Cys	Asp	Asn	Asp	Tyr	Tyr	Lys	Arg	Ala	Ile	Asp	Arg	Met
			180					185					190		
Val	Ser	Glu	His	Pro	Asp	Ala	Ile	Phe	Phe	Val	Phe	Thr	Ser	Asp	Lys
		195					200					205			
Glu	Trp	Cys	Glu	Gln	Asn	Val	Ser	Gly	Lys	Lys	Phe	Arg	Ile	Val	Asp
	210					215					220				
Thr	Lys	Glu	Glu	Asn	Asp	Asp	Ala	Ala	Asp	Leu	Leu	Leu	Met	Ser	Leu
225					230					235					240

37847-517 Sequence Listing_ST25 (003)

Cys Lys His His Ile Leu Ala Asn Ser Ser Tyr Ser Trp Trp Ser Ala
 245 250 255

Trp Met Asn Asp Ser Pro Glu Lys Thr Val Ile Val Pro Ser Lys Trp
 260 265 270

Leu Asn Thr Lys Pro Met Asp Asp Ile Tyr Thr Ser Arg Met Thr Lys
 275 280 285

Ile

<210> 78

<211> 279

<212> PRT

<213> Segetibacter koreensis

<400> 78

Met Val Val Val Lys Leu Ile Gly Gly Met Gly Asn Gln Met Phe Gln
 1 5 10 15

Tyr Ala Ile Gly Arg His Leu Ala Ile Lys Asn Lys Cys Pro Leu Tyr
 20 25 30

Phe Asp His Ile Glu Leu Glu Asn Lys Asn Thr Ala Asn Thr Pro Arg
 35 40 45

Asn Tyr Glu Leu Asp Ile Phe Asn Val Gln Tyr Gln Lys Asn Pro Phe
 50 55 60

Leu Gln Ser Asn Arg Phe Val Ala Lys Val Tyr His Lys Leu Phe Ser
 65 70 75 80

Val Gln Arg Ile Lys Glu Pro Asp Phe Thr Phe His Pro His Ile Leu
 85 90 95

Asn Val Gln Gly Asn Ile His Leu Asn Gly Tyr Trp Gln Asn Glu Asn
 100 105 110

Tyr Phe Lys Glu Ile Glu Glu Ile Ile Arg Gln Asp Phe Thr Phe Lys

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

115

120

125

Thr Pro Ala Asn Glu Lys Ile Glu Ser Ile Leu Gln Gln Ile Ala Ala
130 135 140

Thr Asn Ser Val Ser Leu His Val Arg Arg Gly Asp Tyr Ile Thr Leu
145 150 155 160

Thr Glu Ala Asn Gln Phe His Gly Val Cys Ser Asp Thr Tyr Tyr Gln
165 170 175

Lys Ala Ile Ala Lys Ile Lys Glu Ala Ile Pro Ala Pro His Leu Phe
180 185 190

Val Phe Ser Asp Asp Ile His Trp Val Lys Gln Asn Met Pro Phe Thr
195 200 205

Glu Glu His Thr Phe Val Asp Gly Asn Thr Gly Lys Asn Ser Phe Glu
210 215 220

Asp Leu Arg Leu Met Ala Ala Cys Arg His Asn Ile Leu Ala Asn Ser
225 230 235 240

Ser Phe Ser Trp Trp Ala Gly Trp Leu Asn Lys Asn Pro Glu Lys Met
245 250 255

Val Ile Ala Pro Glu Lys Trp Phe Arg Ala Val His Thr Asp Ile Val
260 265 270

Pro Pro Ser Trp Ile Lys Met
275

<210> 79

<211> 288

<212> PRT

<213> Amphritea japonica

<400> 79

Met Val Ile Val Arg Leu Ile Gly Gly Leu Gly Asn Gln Leu Phe Gln
1 5 10 15

37847-517 Sequence Listing_ST25 (003)

Tyr Ala Tyr Ala Leu Ser Leu Leu Glu Gln Gly Tyr Asp Val Lys Leu
 20 25 30
 Asp Ala Ser Ala Phe Glu Ser Tyr Thr Leu His Gly Gly Phe Gly Leu
 35 40 45
 Gly Glu Tyr Ala Glu Arg Leu Glu Val Ala Thr Thr Glu Glu Val Asp
 50 55 60
 Met Val Ser Arg Val Gly Arg Ile Ser Thr Leu Leu Arg Lys Leu Gln
 65 70 75 80
 Gly Lys Lys Ser Arg Arg Val Ile Lys Glu Ser Asn Phe Ser Tyr Asp
 85 90 95
 Glu Lys Met Leu Thr Pro Glu Asp Ser His Tyr Leu Val Gly Tyr Phe
 100 105 110
 Gln Ser Glu Leu Tyr Phe Asn Lys Ile Arg Gly Glu Leu Leu Ser Ala
 115 120 125
 Leu Asp Leu Lys His Lys Leu Ser Pro Tyr Thr Glu Ala Ser Tyr Leu
 130 135 140
 Ala Ile Ala Asp Ala Ser Val Ser Val Ser Met His Ile Arg Arg Gly
 145 150 155 160
 Asp Tyr Val Ser Asp Lys Ala Ala His Asn Thr His Gly Val Cys Ser
 165 170 175
 Leu Asp Tyr Tyr Tyr Ala Ala Val Thr Phe Phe Glu Glu Arg Tyr Pro
 180 185 190
 Asp Val Asp Phe Tyr Ile Phe Ser Asp Asp Ile Glu Trp Val Lys Glu
 195 200 205
 Asn Leu Asn Val Gln Arg Ala His Tyr Ile Ser Ser Glu Glu Lys Arg
 210 215 220

37847-517 Sequence Listing_ST25 (003)

Phe Ala Gly Glu Asp Ile Tyr Leu Met Ser Gln Cys Asp His Asn Ile
225 230 235 240

Val Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala Trp Leu Asn Ala Asn
245 250 255

Glu Asp Lys Ile Val Val Ala Pro Arg Gln Trp Tyr Ala Asp Ser Asn
260 265 270

Met Gln Arg Leu Ser Lys Thr Leu Val Pro Asp Thr Trp Ile Arg Leu
275 280 285

<210> 80

<211> 280

<212> PRT

<213> Desulfovibrio vulgaris

<400> 80

Met Arg Pro Val Val Val Asp Ile Phe Gly Gly Leu Gly Asn Gln Met
1 5 10 15

Phe Gln Tyr Ala Ala Ala Lys Ser Leu Ala Glu Arg Leu Gly Val Arg
20 25 30

Leu Glu Leu Asp Val Ser Met Phe Ser Gly Asp Pro Leu Arg Ala Phe
35 40 45

Ser Leu Gly Glu Phe Ala Ile Thr Asp His Val Arg Gly Lys Ser Arg
50 55 60

Ser Ser Leu Leu Val Arg Phe Ala Arg Ser Leu Gly Phe Gly Ser Ser
65 70 75 80

Ser Lys Cys Val Glu Pro Phe Phe His Tyr Trp Glu Gly Ile Asn Glu
85 90 95

Ile Glu Ala Pro Val His Met His Gly Tyr Trp Gln Ser Glu Lys Tyr
100 105 110

37847-517 Sequence Listing_ST25 (003)

Phe Lys Ala Tyr Glu Asp Leu Ile Arg Arg Thr Phe Ser Phe Ser Ala
115 120 125

Cys Glu Gly Val Ala Ser Ser Gly Lys Tyr Ala Gly Val Ser Ser Pro
130 135 140

Met Ser Val Ser Val His Leu Arg Arg Gly Asp Tyr Lys Glu Gln Lys
145 150 155 160

Asn Val Val Val His Gly Ile Leu Gly Arg Glu Tyr Tyr Asp Ala Ala
165 170 175

Tyr Ser Ile Ile Lys Gln Gly Cys Pro Ser Ala Cys Phe Phe Val Phe
180 185 190

Thr Asp Ala Ile Asn Glu Ala Val Asp Phe Phe Ser His Trp Asn Asp
195 200 205

Val Leu Phe Val Asp Gly Asn Asn Gln Tyr Gln Asp Met Tyr Leu Met
210 215 220

Ser Gln Cys Arg His His Ile Ile Ala Asn Ser Ser Tyr Ser Trp Trp
225 230 235 240

Gly Ala Trp Leu Gly Ala Phe Ser Asp Gly Met Thr Val Ala Pro Lys
245 250 255

Met Trp Phe Ala Tyr Asp Val Leu Lys Glu Lys Ser Ile Lys Asp Leu
260 265 270

Phe Pro Glu Asp Trp Ile Val Leu
275 280

<210> 81

<211> 290

<212> PRT

<213> Spirosoma spitsbergense

<400> 81

37847-517 Sequence Listing_ST25 (003)

Met	Ile	Ile	Ser	Arg	Ile	Thr	Ser	Gly	Leu	Gly	Asn	Gln	Leu	Phe	Gln	1	5	10	15
Tyr	Ala	Val	Ala	Arg	His	Leu	Ser	Leu	Lys	Asn	Lys	Thr	Ser	Leu	Tyr	20	25	30	
Val	Asp	Leu	Ser	Tyr	Tyr	Leu	Tyr	Gln	Tyr	His	Asp	Asp	Thr	Ser	Arg	35	40	45	
Asn	Phe	Lys	Leu	Gly	Asn	Phe	Ser	Val	Pro	Tyr	His	Thr	Leu	Gln	Gln	50	55	60	
Ser	Pro	Val	Glu	Tyr	Val	Ser	Lys	Ala	Thr	Lys	Leu	Leu	Pro	Asn	Arg	65	70	75	80
Ser	Leu	Arg	Pro	Phe	Phe	Leu	Phe	Gln	Lys	Glu	Arg	Gln	Phe	His	Phe	85	90	95	
Asp	Glu	Gln	Ile	Leu	Gln	Ser	Arg	Ala	Gly	Cys	Val	Ile	Leu	Glu	Gly	100	105	110	
Phe	Trp	Gln	Ser	Glu	Ala	Tyr	Phe	Arg	Asp	Asn	Ala	Asp	Thr	Ile	Arg	115	120	125	
Arg	Asp	Leu	Gln	Leu	Ser	Gly	Thr	Pro	Ser	Pro	Glu	Phe	Asn	Gln	Tyr	130	135	140	
Arg	Glu	Leu	Ile	Arg	Glu	Thr	Pro	Met	Ser	Val	Ser	Ile	His	Val	Arg	145	150	155	160
Arg	Ser	Asp	Tyr	Val	Asn	His	Pro	Glu	Phe	Ser	Gln	Thr	Phe	Gly	Phe	165	170	175	
Val	Gly	Ile	Asp	Tyr	Tyr	Lys	Arg	Ala	Ile	Glu	Leu	Ala	Arg	Lys	Glu	180	185	190	
Leu	Ala	Asn	Pro	Arg	Phe	Phe	Val	Phe	Ser	Asp	Asp	Lys	Glu	Trp	Ser	195	200	205	

37847-517 Sequence Listing_ST25 (003)

Lys Thr Asn Leu Pro Leu Gly Glu Asp Ser Val Phe Val Gln Asn Thr
210 215 220

Gly Leu Asn Gly Asp Val Ala Asp Leu Val Leu Met Ser His Cys Gln
225 230 235 240

His His Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala Trp Leu
245 250 255

Asn Pro Asn Ala Gly Lys Leu Val Ile Thr Pro Lys Asn Trp Tyr Lys
260 265 270

Asn Lys Pro Ala Trp Asn Thr Lys Asp Leu Leu Pro Pro Thr Trp Leu
275 280 285

Ser Ile
290

<210> 82
<211> 306
<212> PRT
<213> Lachnospiraceae bacterium 28-4

<400> 82

Met Asn Ile Ile Arg Met Ser Gly Gly Ile Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Leu Tyr Leu Lys Leu Val Ser Leu Gly Lys Glu Val Lys Phe
20 25 30

Asp Asp Val Thr Glu Tyr Glu Leu Asp Asn Ala Arg Pro Ile Met Leu
35 40 45

Ser Val Phe Gly Ile Asp Tyr Pro Lys Ala Ser Arg Glu Glu Leu Val
50 55 60

Glu Leu Thr Asp Ala Ser Met Asp Phe Leu Ser Arg Val Arg Arg Lys
65 70 75 80

Ile Phe Gly Arg Lys Ser Gly Glu Tyr His Glu Ala Ser Ala Asp Tyr

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

85 90 95

Asp Glu Thr Val Leu Glu Lys Glu His Ala Tyr Leu Cys Gly Cys Phe
100 105 110

Gln Ser Glu Arg Tyr Phe Lys Asp Ile Glu Tyr Glu Val Arg Glu Ala
115 120 125

Tyr Arg Phe Arg Asn Val Val Val Pro Glu Glu Ile Arg Gly Gly Ile
130 135 140

Glu Thr Tyr Glu Arg Gln Ile Gly Glu Ser Leu Ser Val Ser Ile His
145 150 155 160

Ile Arg Arg Gly Asp Tyr Leu Asp Ala Ala Asp Val Tyr Gly Gly Ile
165 170 175

Cys Thr Asp Ala Tyr Tyr Asn Gln Ala Ile Arg Tyr Met Ile Lys Lys
180 185 190

Tyr Glu Asn Pro Ser Phe Phe Val Phe Thr Asn Asp Thr Phe Trp Ala
195 200 205

Glu Lys Trp Cys Glu Val Arg Glu Arg Glu Thr Gly Lys Arg Phe Thr
210 215 220

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

Met Ser Arg Cys Lys Ala His Ile Ile Ala Asn Ser Ser Phe Ser Trp
245 250 255

Trp Gly Ala Trp Leu Asp Ala Ser Pro Asp Lys Cys Val Val Ala Pro
260 265 270

Val Lys Trp Ile Asn Thr Arg Glu Cys Arg Asp Ile Tyr Thr Glu Asp
275 280 285

Met Val Arg Ile Gly Ser Asn Gly Lys Ile Ser Phe Ser Asn Cys Ser

37847-517 Sequence Listing_ST25 (003)

290

295

300

Ser Leu
305

<210> 83

<211> 281

<212> PRT

<213> Lachnospiraceae bacterium COE1

<400> 83

Met Val Val Val Arg Ile Trp Glu Gly Leu Gly Asn Gln Leu Phe Gln
1 5 10 15Tyr Ala Tyr Ala Arg Ala Leu Ser Leu Arg Thr Lys Asp Arg Val Tyr
20 25 30Leu Asp Ile Ser Glu Tyr Glu Met Ser Pro Lys Pro Val Arg Lys Tyr
35 40 45Glu Leu Cys His Phe Lys Ile Lys Gln Pro Val Ile Asn Cys Gly Arg
50 55 60Ile Phe Pro Phe Val Asn Lys Asp Ser Phe Tyr Thr Lys Asn Asn Gln
65 70 75 80Tyr Leu Arg Tyr Phe Pro Ala Gly Leu Ile Lys Glu Glu Asp Cys Tyr
85 90 95Phe Lys Arg Asp Phe Cys Glu Leu Lys Gly Leu Leu Tyr Leu Lys Gly
100 105 110Trp Phe Gln Ser Glu Lys Tyr Phe Lys Glu Phe Glu Ser His Ile Arg
115 120 125Glu Glu Ile Tyr Pro Arg Asn Lys Ile Lys Ile Thr Arg Gly Leu Arg
130 135 140Lys Ile Leu Asn Ser Asp Asn Thr Val Ser Val His Ile Arg Arg Gly
145 150 155 160

37847-517 Sequence Listing_ST25 (003)

Asp Phe Gly Lys Asp His Asn Ile Leu Pro Ile Glu Tyr Tyr Glu Asn
165 170 175

Ser Lys Arg Val Ile Leu Glu Arg Val Asp Asn Pro Tyr Phe Ile Ile
180 185 190

Phe Ser Asp Asp Ile Leu Trp Val Lys Glu Asn Met Asn Phe Gly Leu
195 200 205

Asn Cys Phe Tyr Met Asp Lys Glu Tyr Ser Tyr Lys Asp Tyr Glu Glu
210 215 220

Leu Met Ile Met Ser Arg Cys Lys His Asn Ile Ile Ala Asn Ser Thr
225 230 235 240

Phe Ser Trp Trp Gly Ala Trp Leu Asn Pro Ser Lys Asp Lys Ile Val
245 250 255

Ile Ala Pro Lys Lys Trp Phe Leu Tyr Asn Pro Lys Lys Asp Phe Asp
260 265 270

Ile Val Pro Asn Asp Trp Ile Arg Val
275 280

<210> 84
<211> 301
<212> PRT
<213> Parabacteroides distasonis CL09T03C24

<400> 84

Met Lys Ile Val Asn Ile Ile Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Phe Ala Val Ala Leu Lys Ala Lys Tyr Pro Asn Glu Glu Val
20 25 30

Phe Ile Asp Thr Gln His Tyr Lys Asn Ala Phe Ile Lys Val Tyr His
35 40 45

37847-517 Sequence Listing_ST25 (003)

Gly Asn Asn Phe Tyr His Asn Gly Tyr Glu Ile Asp Lys Val Phe Pro
50 55 60

Asn Ala Thr Leu Glu Pro Ala Arg Pro Lys Asp Leu Met Lys Val Ser
65 70 75 80

Phe Tyr Ile Pro Asn Gln Val Leu Ala Arg Ala Val Arg Arg Ile Phe
85 90 95

Pro Lys Arg Lys Thr Glu Phe Val Thr Asp Gln Gln Pro Tyr Val Phe
100 105 110

Ile Pro Glu Ala Leu Ser Val Ile Asp Asp Cys Tyr Phe Asp Gly Tyr
115 120 125

Trp Met Thr Pro Leu Tyr Phe Asp Lys Tyr Arg Asp Arg Ile Leu Lys
130 135 140

Glu Phe Thr Phe Arg Pro Phe Asp Thr Lys Glu Asn Leu Glu Leu Glu
145 150 155 160

Pro Leu Leu Lys Gln Asp Asn Ser Val Thr Val His Ile Arg Arg Gly
165 170 175

Asp Tyr Val Gly Ser Ser Ser Phe Gly Gly Ile Cys Thr Leu Asp Tyr
180 185 190

Tyr Arg Asn Ala Ile Arg Glu Ala Tyr Asn Leu Ile Thr Ser Pro Glu
195 200 205

Phe Phe Ile Phe Ser Asn Asp Gln Lys Trp Cys Met Glu Asn Met Arg
210 215 220

Asn Glu Phe Gly Asp Ala Lys Val His Phe Ile Ala His Asn Arg Gly
225 230 235 240

Ala Asp Ser Tyr Arg Asp Met Gln Leu Leu Ser Ile Ala Arg Cys Asn
245 250 255

37847-517 Sequence Listing_ST25 (003)

Ile Leu Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala Tyr Leu Asn Gln
260 265 270

Arg Lys Asn Cys Phe Ile Ile Cys Pro His Lys Trp His Asn Thr Leu
275 280 285

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

<210> 85

<211> 283

<212> PRT

<213> Bacteroides sp. HPS0048

<400> 85

Met Phe Val Ile Arg Leu Ile Gly Gly Val Gly Asn Gln Leu Phe Gln
1 5 10 15

Tyr Thr Phe Gly Gln Phe Leu Arg His Lys Phe Gly Val Glu Val Cys
20 25 30

Tyr Asp Ile Val Ala Phe Asp Thr Val Asp Lys Gly Arg Asn Leu Glu
35 40 45

Leu Gln Leu Leu Asp Glu Ser Leu Pro Leu Phe Glu Thr Ser Asn Phe
50 55 60

Phe Phe Ser Lys Tyr Lys Ser Trp Lys Lys Arg Leu Phe Leu Tyr Gly
65 70 75 80

Phe Leu Leu Lys Lys Asn Asn Lys Tyr Tyr Thr Lys Tyr Ala Pro Glu
85 90 95

Glu Ile Ser Leu Phe Thr Glu Lys Gly Leu Ser Tyr Phe Asp Gly Trp
100 105 110

Trp Gln Tyr Pro Ala Leu Leu Arg Asp Thr Ile Asn Asn Met Glu Asp
115 120 125

37847-517 Sequence Listing_ST25 (003)

Phe Phe Ile Pro Lys Gln Pro Ile Pro Val Gln Ile Gln Lys Tyr Tyr
 130 135 140

Asn Glu Ile Leu Leu Asn Asn Phe Ala Val Ala Leu His Val Arg Arg
 145 150 155 160

Gly Asp Tyr Phe Thr Ser Lys Tyr Ala Lys Thr Tyr Ala Val Cys Asn
 165 170 175

Val Glu Tyr Tyr Thr Ser Ala Val Asn Leu Met Cys Glu Lys Leu Arg
 180 185 190

Ser Cys Lys Phe Tyr Val Phe Ser Asp Asp Leu Asp Trp Val Lys Ser
 195 200 205

Asn Leu Ile Leu Pro Ser Asn Thr Val Tyr Val Lys Asn Tyr Asp Ile
 210 215 220

Asn Ser Tyr Trp Tyr Ile Tyr Leu Met Ser Leu Cys Arg His Ile Ile
 225 230 235 240

Ile Ser Asn Ser Ser Phe Ser Trp Trp Gly Ala Thr Leu Asn Arg Asn
 245 250 255

Phe His Lys Ile Val Ile Ala Pro Lys Tyr Trp Ser Thr Lys Lys Asn
 260 265 270

Asn Thr Leu Cys Asp Asn Ser Trp Ile Lys Ile
 275 280

<210> 86

<211> 291

<212> PRT

<213> Bacteroides thetaiotaomicron

<400> 86

Met Lys Ile Ile Asn Ile Leu Gly Gly Leu Gly Asn Gln Met Phe Glu
 1 5 10 15

Tyr Ala Met Tyr Leu Ala Leu Lys Asn Ala His Ser Glu Glu Glu Ile

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

20 25 30

Leu Cys Ser Thr Arg Ser Phe Cys Gly Tyr Gly Leu His Asn Gly Tyr
35 40 45

Glu Leu Gly Arg Ile Phe Gly Ile Gln Val Lys Glu Ala Ser Leu Leu
50 55 60

Gln Leu Thr Lys Leu Ala Tyr Pro Phe Phe Asn Tyr Lys Ser Trp Gln
65 70 75 80

Val Met Arg His Trp Leu Pro Val Arg Lys Thr Met Thr Arg Gly Ala
85 90 95

Ile Asn Ile Pro Phe Asp Tyr Ser Gln Val Met Arg Glu Asp Ser Val
100 105 110

Tyr Tyr Asp Gly Tyr Trp Gln Asn Glu Lys Asn Phe Leu His Ile Arg
115 120 125

Glu Glu Ile Leu Thr Ala Tyr Thr Phe Pro Lys Phe Asp Asp Glu Lys
130 135 140

Asn Gln Glu Leu Ala Asp Ile Ile Val Lys Ser Asn Ala Val Ser Cys
145 150 155 160

His Ile Arg Arg Gly Asp Tyr Leu Lys Glu Ile Asn Met Cys Val Cys
165 170 175

Thr Ser Ser Tyr Tyr Ala His Ala Ile Ser Tyr Met Asn Glu Glu Ile
180 185 190

Asn Pro Asn Leu Tyr Cys Val Phe Ser Asp Asp Ile Glu Trp Cys Arg
195 200 205

Asn Asn Ile Cys Glu Leu Met Gly Glu Asp Lys Lys Ile Ile Phe Ile
210 215 220

Asp Trp Asn Lys Gly Glu Lys Ser Phe Arg Asp Met Gln Leu Met Ser

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

225 230 235 240

Leu Cys Lys His Asn Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Gly
245 250 255

Ala Trp Leu Asn Arg Asn Asp Lys Lys Ile Val Val Ala Pro Thr Arg
260 265 270

Trp Ile Ala Ser Glu Val Lys Asn Asp Pro Leu Cys Asp Ser Trp Lys
275 280 285

Arg Ile Glu
290

<210> 87
<211> 292
<212> PRT
<213> Desulfovibrio alaskensis

<400> 87

Met Lys Phe Val Gly Val Trp Ile Leu Gly Gly Leu Gly Asn Gln Met
1 5 10 15

Phe Gln Phe Ala Ala Ala Tyr Ala Leu Ala Lys Arg Met Gly Gly Glu
20 25 30

Leu Arg Leu Asp Leu Ser Gly Phe Lys Lys Tyr Pro Leu Arg Ser Tyr
35 40 45

Ser Leu Asp Leu Phe Thr Val Asp Thr Pro Leu Trp His Gly Leu Pro
50 55 60

Met Ser Gln Arg Arg Phe Arg Ile Pro Met Asp Ala Trp Thr Arg Gly
65 70 75 80

Ser Arg Leu Pro Leu Val Pro Ser Pro Pro Phe Val Met Ala Lys Glu
85 90 95

Lys Asn Phe Ala Phe Ser Pro Ile Val Tyr Glu Leu Gln Gln Ser Cys
100 105 110

37847-517 Sequence Listing_ST25 (003)

Tyr Leu Tyr Gly Tyr Trp Gln Ser Tyr Arg Tyr Phe Gln Asp Val Glu
115 120 125

Asp Asp Ile Arg Thr Leu Phe Ser Leu Ser Arg Phe Ala Thr Leu Glu
130 135 140

Leu Ala Pro Val Val Ala Gln Leu Asn Glu Val Glu Ser Val Ala Val
145 150 155 160

His Leu Arg Arg Gly Asp Tyr Ile Thr Asp Ala Ala Ser Asn Ala Val
165 170 175

His Gly Val Cys Gly Ile Asp Tyr Tyr Gln Arg Ser Met Ser Leu Val
180 185 190

Arg Arg Ser Thr Thr Lys Pro Ile Phe Tyr Ile Phe Ser Asp Glu Pro
195 200 205

Glu Val Ala Lys Lys Leu Phe Ala Thr Glu Asp Asp Val Val Val Met
210 215 220

Pro Ser Arg Arg Gln Glu Glu Asp Leu Leu Leu Met Ser Arg Cys Lys
225 230 235 240

His His Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Ala Ala Trp Leu
245 250 255

Gly Lys Arg Ala Ser Gly Leu Cys Ile Ala Pro Arg Tyr Trp Phe Ala
260 265 270

Arg Pro Lys Leu Glu Ser Thr Tyr Leu Phe Asp Leu Ile Pro Asp Glu
275 280 285

Trp Leu Leu Leu
290

<210> 88

<211> 296

37847-517 Sequence Listing_ST25 (003)

<212> PRT

<213> Prevotella oralis CC98A

<400> 88

Met Asp Ile Val Val Ile Phe Asn Gly Leu Gly Asn Gln Met Ser Gln
1 5 10 15

Tyr Ala Phe Tyr Leu Ala Lys Arg Lys Ser Gly Ser Arg Cys His Cys
20 25 30

Ile Phe His Asn Val Ser Thr Gly Phe His Asn Gly Ser Glu Leu Asp
35 40 45

Lys Val Phe Gly Ile Lys Tyr Glu Lys Gly Ile Phe Ser Lys Leu Leu
50 55 60

Ser Lys Ile Tyr Asp Ile Phe Asp Gly Ile Pro Lys Leu Arg Lys Lys
65 70 75 80

Leu Asn Ser Leu Gly Ile His Ile Ile Arg Glu Pro Arg Asn Tyr Asp
85 90 95

Tyr Thr Ala Ser Leu Leu Pro Arg Val Ser Arg Trp Gly Leu Asn Tyr
100 105 110

Phe Val Gly Gly Trp His Ser Glu Lys Tyr Tyr Thr Glu Ile Leu Gln
115 120 125

Glu Ile Lys Asn Thr Phe Ser Phe Lys Ile Asp Asp Glu Ile Lys Asp
130 135 140

Ile Asp Phe Tyr Glu Phe Tyr Ser Leu Ile His Asn Asp Ile Asn Ser
145 150 155 160

Val Ser Leu His Ile Arg Arg Gly Asp Tyr Val Gly Ala Asn Glu Tyr
165 170 175

Ser Tyr Phe Gln Phe Gly Gly Val Ala Thr Leu Glu Tyr Tyr His Lys
180 185 190

37847-517 Sequence Listing_ST25 (003)

Ala Ile Asp Glu Ile Tyr Gln Arg Ile Glu Asn Pro Thr Phe Tyr Val
195 200 205

Phe Ser Asp Asp Ile Gly Trp Cys Lys Thr Thr Phe Leu Lys Asn Asn
210 215 220

Phe Ile Phe Val Asp Cys Asn Cys Gly Glu Lys Ser Trp Arg Asp Met
225 230 235 240

Phe Leu Ile Ser Gln Cys Lys His His Ile Ile Ala Asn Ser Thr Phe
245 250 255

Ser Trp Trp Gly Ala Trp Leu Ser Ile Phe His Asn Ser Ile Thr Ile
260 265 270

Cys Pro Lys Glu Phe Ile Lys Gly Val Val Ile Arg Asp Val Tyr Pro
275 280 285

Asp Thr Trp Ile Lys Leu Ser Ser
290 295

<210> 89
<211> 303
<212> PRT
<213> Comamonadaceae bacterium CR

<400> 89

Met Ala Ser Lys Ile Ser Lys Ile Ile Pro Arg Ile Phe Gly Gly Leu
1 5 10 15

Gly Asn Gln Leu Phe Ile Tyr Ala Ala Ala Arg Arg Leu Ala Leu Val
20 25 30

Asn Gly Ala Glu Leu Ala Leu Asp Asp Val Ser Gly Phe Val Arg Asp
35 40 45

His Glu Tyr Asn Arg His Tyr Gln Leu Asp His Phe Asn Ile Pro Cys
50 55 60

37847-517 Sequence Listing_ST25 (003)

Arg Lys Ala Thr Ala Ala Glu Arg Leu Glu Pro Phe Ala Arg Val Arg
65 70 75 80

Arg Tyr Leu Lys Arg Lys Trp Asn Gln Arg Leu Pro Phe Glu Gln Arg
85 90 95

Lys Tyr Leu Val Gln Glu Ser Val Asp Phe Asp Glu Arg Leu Leu Thr
100 105 110

Phe Lys Pro Arg Gly Thr Val Tyr Leu Glu Gly Tyr Trp Gln Ser Glu
115 120 125

Asp Tyr Phe Lys Asp Ile Glu Pro Gln Ile Arg Ala Asp Leu Arg Ile
130 135 140

His Pro Pro Thr Asp Thr Val Asn Gln Gln Met Ala Glu Arg Ile Arg
145 150 155 160

Ala Thr Asn Ala Val Ala Val His Val Arg Phe Phe Asp Ala Pro Ala
165 170 175

Gln Ser Ala Leu Gly Val Gly Gly Asn Asn Ala Pro Gly Asp Tyr Tyr
180 185 190

Gln Arg Ala Ile Lys Val Met Gln Glu Gln Ala Pro Asp Ala Gln Tyr
195 200 205

Tyr Ile Phe Ser Asp Gln Pro Gln Ala Ala Arg Ala Arg Ile Pro Leu
210 215 220

Arg Asp Asp His Val Thr Leu Val Asn His Asn Gln Cys Asp Ala Val
225 230 235 240

Ala Tyr Ala Asp Leu Trp Leu Ile Ser Gln Cys Gln His Phe Ile Ile
245 250 255

Ala Asn Ser Thr Phe Ser Trp Trp Gly Ala Trp Leu Gly Lys Thr Pro
260 265 270

37847-517 Sequence Listing_ST25 (003)

Glu Ser Ile Val Ile Ala Pro Gly Phe Glu Lys Arg Glu Gly Ala Met
275 280 285

Phe Trp Gly Phe Arg Gly Leu Leu Pro Asp Arg Trp Val Lys Leu
290 295 300

<210> 90

<211> 292

<212> PRT

<213> *Vibrio nigripulchritudo*

<400> 90

Met Lys Asp Ser Arg Ile Val Lys Leu Asn Gly Gly Leu Gly Asn Gln
1 5 10 15

Met Phe Gln Phe Ala Leu Ala Phe Ala Leu Lys Lys Lys Leu Asn Val
20 25 30

Ala Val Lys Phe Asp Thr Glu Leu Leu Asp Thr Asn Arg Thr Glu Phe
35 40 45

Lys Leu Ser Leu Glu Arg Phe Gly Leu Ile Val Asp Lys Leu Thr Ile
50 55 60

Thr Glu Lys Phe Lys Tyr Lys Gly Leu Glu Ser Cys Lys Tyr Arg Lys
65 70 75 80

Ile Cys Asn Trp Ile Ser Asn Phe Thr Thr Ile Asn Ile His Lys Gly
85 90 95

Tyr Tyr Lys Glu Lys Glu Arg Gly Val Tyr Asp Arg Gly Ile Phe Asp
100 105 110

Ser Asn Val Lys Tyr Ile Asp Gly Tyr Trp Gln Asn Gln Glu Tyr Phe
115 120 125

Asn Asp Phe Arg Ser Glu Leu Leu Asn Lys Phe Asn Leu Asn Gly Lys
130 135 140

Val Ser Asn His Ala Ile Gln Tyr Leu Lys Glu Ile Thr Ser Val Gln

28 May 2018
2015259088

37847-517 Sequence Listing_ST25 (003)

145 150 155 160

Asn Ser Val Ser Ile His Val Arg Arg Gly Asp Tyr Leu Leu Leu Asp
165 170 175

Val Tyr Arg Asn Leu Thr Leu Asp Tyr Tyr Ser Glu Ala Ile Lys Leu
180 185 190

Val Arg Ile Thr Asn Pro Asp Ser Lys Phe Phe Ile Phe Ser Asn Asp
195 200 205

Ile Asn Trp Cys Lys Ser Asn Phe Lys Ser Val Asp Asn Ala Ile Phe
210 215 220

Val Asp Ser Thr Val Asp Glu Phe Asp Asp Met Phe Leu Met Ser Lys
225 230 235 240

Cys Lys Thr Asn Ile Ile Ala Asn Ser Thr Phe Ser Trp Trp Ala Ala
245 250 255

Trp Leu Asn Asn Asn Ser Gly Lys Ile Val Tyr Cys Pro Lys Lys Trp
260 265 270

Arg Asn Asp Thr Thr Glu Val His Lys Gly Leu Pro Glu Gly Trp Asn
275 280 285

Ile Ile Asp Lys
290

<210> 91
<211> 295
<212> PRT
<213> Sulfurospirillum deleyianum

<400> 91

Met Ile Ile Ile Lys Ile Met Gly Gly Leu Thr Ser Gln Met His Lys
1 5 10 15

Tyr Ala Leu Gly Arg Val Leu Ser Leu Lys Tyr Asn Val Pro Leu Lys
20 25 30

37847-517 Sequence Listing_ST25 (003)

Leu Asp Leu Thr Trp Phe Asp Asn Pro Lys Ser Asp Thr Pro Trp Glu
 35 40 45
 Tyr Gln Leu Asp Tyr Phe Asn Ile Asn Ala Thr Ile Ala Thr Val Ser
 50 55 60
 Glu Ile Lys Lys Leu Lys Gly Asn Asn Leu Phe Asn Arg Ile Ala Arg
 65 70 75 80
 Lys Ile Glu Lys Phe Phe Ser Ile Arg Ile Tyr Lys Lys Ser Tyr Ile
 85 90 95
 Asn Lys Ser Phe Ile Ser Ile Ser Asp Phe His Lys Leu Lys Ser Asp
 100 105 110
 Ile Tyr Leu Asp Gly Glu Trp Asn Gly Phe Lys Tyr Phe Glu Asp Tyr
 115 120 125
 Gln Asp Thr Ile Lys Asn Glu Leu Thr Leu Lys Arg Gly Ser Ser Ile
 130 135 140
 Asn Ile Gln Asn Thr Ile Lys Glu Leu Lys Ser Ser Asp Asn Ser Val
 145 150 155 160
 Phe Leu His Ile Arg Arg Gly Asp Tyr Leu Ser Asn Lys Asn Ala Ala
 165 170 175
 Ala Phe His Ala Lys Cys Ser Leu Asp Tyr Tyr Tyr Lys Ala Ile Gln
 180 185 190
 Ile Val Lys Glu Lys Ile Asp Asn Pro Ile Phe Tyr Ile Phe Ser Asp
 195 200 205
 Asp Ile Leu Trp Val Lys Lys Asn Phe Val Ile Asn Glu Ser Cys Arg
 210 215 220
 Phe Met Glu Lys Asn Gln Asn Phe Glu Asp Leu Leu Leu Met Ser Tyr
 225 230 235 240

37847-517 Sequence Listing_ST25 (003)

Cys Lys His Gly Ile Thr Ala Asn Ser Gly Phe Ser Leu Met Ala Gly
245 250 255

Trp Leu Asn Gln Asn Lys Asp Lys Met Ile Ile Val Pro Gln Thr Trp
260 265 270

Val Asn Asp Asp Arg Ile Asn Ile Asn Ile Leu Asn Ser Leu Glu Gln
275 280 285

Asp Asn Phe Thr Ile Ile Arg
290 295

<210> 92

<211> 298

<212> PRT

<213> Escherichia coli

<400> 92

Met Thr Phe Ile Val Arg Leu Thr Gly Gly Leu Gly Asn Gln Met Phe
1 5 10 15

Gln Tyr Ala Leu Ala Arg Ser Leu Ala Lys Lys Tyr Asn Ala Arg Leu
20 25 30

Lys Leu Asp Ile Ser Tyr Tyr His Asn Gln Pro His Lys Asp Thr Pro
35 40 45

Arg Thr Phe Glu Leu Asn Gln Leu Cys Ile Val Asp Asn Ile Leu Asn
50 55 60

Ser Ser Ser Phe Ser Glu Lys Phe Leu Tyr Ile Tyr Asp Lys Leu Arg
65 70 75 80

Val Lys Leu Ser Lys Lys Ile Ser Leu Pro Tyr Phe Arg Asn Ile Val
85 90 95

Thr Pro Val Asn Phe Asn Cys Ile Asp Phe Ala Glu Asp Lys Asp Tyr
100 105 110

37847-517 Sequence Listing_ST25 (003)

Tyr Phe Leu Gly His Phe Gln Glu Leu Ser Asn Ile Tyr Ser Ile Asp
 115 120 125

Glu Ser Leu Arg Ser Glu Phe Lys Pro Asn Gln Glu Ile Met Asn Leu
 130 135 140

Ala His Gln Ser Lys Ile Tyr Glu Leu Ile Lys Gln Ser Arg Gly Ser
 145 150 155 160

Val Ala Leu His Ile Arg Arg Gly Asp Tyr Val Thr Asn Lys Asn Ala
 165 170 175

Ala Glu His His Gly Val Ile Gly Leu Ser Tyr Tyr Val Asn Ala Leu
 180 185 190

Ser Tyr Leu Glu Asn Val Ser Glu Phe Phe Asp Val Phe Val Phe Ser
 195 200 205

Asp Asp Pro Glu Trp Ala Arg Lys Asn Ile Lys Asn Ser Arg Asn Leu
 210 215 220

Phe Phe Cys Asp Glu Gly Asn Cys Arg Tyr Ser Lys Lys Tyr Ser Thr
 225 230 235 240

Ile Asp Met Tyr Leu Met Ser Gln Cys Asp His Phe Ile Ile Ala Asn
 245 250 255

Ser Thr Tyr Ser Trp Trp Ala Ala Trp Leu Gly Asn Tyr Pro Ser Lys
 260 265 270

His Val Val Ala Pro Ala Arg Trp Asn Ala Asn Asn Ser Pro Tyr Pro
 275 280 285

Ile Leu Gln Asn Trp Lys Ala Ile His Glu
 290 295

<210> 93
 <211> 285
 <212> PRT

37847-517 Sequence Listing_ST25 (003)

<213> Firmicutes bacterium CAG:24

<400> 93

Met Val Gly Val Gln Leu Ser Gly Gly Leu Gly Asn Gln Met Phe Glu
1 5 10 15

Tyr Ala Leu Tyr Leu Lys Leu Lys Ser Met Gly Lys Asp Val Arg Ile
20 25 30

Asp Asp Val Thr Cys Tyr Gly Ala Gln Glu Lys Gln Arg Val Asn Gln
35 40 45

Leu Ser Val Phe Gly Val Ser Tyr Glu His Met Thr Lys Gln Glu Tyr
50 55 60

Glu Gln Ile Thr Asp Ser Ser Met Ser Pro Leu His Arg Ala Arg Arg
65 70 75 80

Leu Leu Cys Gly Arg Lys Asp Leu Ser Tyr Arg Glu Ala Ser Cys Asn
85 90 95

Tyr Asp Pro Glu Ile Leu Arg Arg Glu Pro Ala Leu Leu Leu Gly Tyr
100 105 110

Phe Gln Thr Glu Arg Tyr Phe Ala Asp Ile Lys Asp Gln Val Arg Glu
115 120 125

Ala Phe Thr Phe Arg Asn Leu Thr Leu Thr Lys Glu Ser Ala Ala Met
130 135 140

Glu Gln Gln Met Lys Glu Cys Glu Ser Val Ser Val His Ile Arg Arg
145 150 155 160

Gly Asp Tyr Leu Thr Pro Ala Asn Gln Ala Leu Phe Gly Gly Ile Cys
165 170 175

Asp Leu Asp Tyr Tyr His Arg Ala Val Ala Glu Ile Arg Lys Arg Lys
180 185 190

37847-517 Sequence Listing_ST25 (003)

Pro Asp Val Lys Phe Phe Leu Phe Ser Asn Asp Met Glu Trp Thr Lys
195 200 205

Glu His Phe Cys Gly Ser Glu Phe Val Pro Val Glu Gly Asn Ser Glu
210 215 220

Gln Ala Gly Glu Gln Asp Leu Tyr Leu Met Ser Cys Cys Lys Asn His
225 230 235 240

Ile Leu Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala Trp Leu Asp Asn
245 250 255

Gly Lys Asp Lys Leu Val Ile Ala Pro Glu Lys Trp Met Asn Gly Arg
260 265 270

Gly Cys Cys Asp Ile Tyr Thr Asp Glu Met Ile Arg Val
275 280 285

<210> 94
<211> 287
<212> PRT
<213> Amphritea japonica

<400> 94

Met Val Lys Ile Lys Ile Ile Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Ala Ala Lys Ser Leu Ala Val Leu Asn Asn Thr Arg Val Ser
20 25 30

Ala Asn Val Ser Val Phe Ser Asn Tyr Lys Thr His Pro Leu Arg Leu
35 40 45

Asn Lys Leu Asn Cys Asp Cys Glu Phe Asp Phe Thr Arg Asp Phe Arg
50 55 60

Leu Val Leu Ser Gly Phe Pro Leu Leu Gly Ser Ala Phe Ser Lys Lys
65 70 75 80

Ser Met Leu Leu Asn His Tyr Val Glu Lys Asp Leu Leu Phe Asp Ser

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

85 90 95

Ser Phe Phe Asp Leu Asp Asp Asn Val Leu Leu Ser Gly Tyr Phe Gln
100 105 110

Ser Glu Lys Tyr Phe Ser Asn Ile Arg Glu Leu Leu Ile Gln Glu Phe
115 120 125

Ser Leu Asp Asp Arg Leu Thr Glu Ala Glu Leu Ala Ile Asn Asn Lys
130 135 140

Ile Glu Ser Cys Asn Ser Ile Ala Ile His Ile Arg Arg Gly Asp Tyr
145 150 155 160

Ile Thr Asp Leu Ser Ala Asn Asn Ile His Gly Ile Cys Ser Glu Glu
165 170 175

Tyr Phe Glu Lys Ala Leu Asn Tyr Leu Asp Ser Ile Asn Val Leu Ser
180 185 190

Asp Pro Thr Thr Thr Leu Phe Ile Phe Ser Asp Asp Ile Leu Trp Cys
195 200 205

Lys Asp Asn Leu Ala Phe Lys Tyr Arg Thr Val Phe Val Glu Gly Ser
210 215 220

Val Asp Arg Pro Glu Val Asp Ile His Leu Met Ser Lys Cys Lys His
225 230 235 240

Gln Val Ile Ser Asn Ser Thr Phe Ser Trp Trp Gly Ala Trp Leu Asn
245 250 255

Thr Asn Leu Asp Lys Cys Val Ile Ala Pro Leu Lys Trp Phe Asn Ser
260 265 270

Leu His Asp Ser Thr Asp Ile Val Pro Lys Gln Trp Met Arg Leu
275 280 285

<210> 95

37847-517 Sequence Listing_ST25 (003)

<211> 273
 <212> PRT
 <213> Bacteroides salyersiae

<400> 95

Met Lys Gln Thr Ile Ile Met Ser Gly Gly Leu Gly Asn Gln Met Phe
 1 5 10 15

Gln Tyr Ala Leu Tyr Cys Ser Met Arg Glu Lys Gly Ile Arg Val Lys
 20 25 30

Ile Asp Ile Ser Leu Tyr Glu Phe Asn Arg Met His Asn Gly Tyr Met
 35 40 45

Leu Asp Tyr Ala Phe Gly Leu Asn Ile Ser His Asn Lys Ile Asn Lys
 50 55 60

Tyr Ser Val Leu Trp Thr Arg Leu Ile Arg Ser Asn Arg Ala Pro Phe
 65 70 75 80

Leu Leu Phe Arg Glu Asp Glu Ser Arg Phe Cys Asp Asp Val Phe Thr
 85 90 95

Thr Tyr Lys Pro Tyr Ile Asp Gly Cys Trp Ile Asp Glu Arg Tyr Phe
 100 105 110

Phe Asn Ile Lys Lys Lys Ile Ile Ser Gln Phe Ser Phe His Asn Ile
 115 120 125

Asp Gln Lys Asn Leu Met Val Ala Asn Met Met Lys Val Cys Asn Ser
 130 135 140

Val Ser Leu His Ile Arg Arg Gly Asp Tyr Leu Ser Gln Ser Met Tyr
 145 150 155 160

Asn Ile Cys Asn Glu Ser Tyr Tyr Lys Ser Ala Ile Glu Tyr Ile Ile
 165 170 175

Ser Arg Val Glu Asp Ser Lys Phe Phe Ile Phe Ser Asp Asp Pro Glu
 180 185 190

37847-517 Sequence Listing_ST25 (003)

Trp Cys Lys Tyr Phe Met Glu Lys Phe Asn Val Asp Tyr Glu Ile Ile
195 200 205

Gln His Asn Phe Gly Lys Asp Ser Tyr Lys Asp Met Tyr Leu Met Thr
210 215 220

Gln Cys Lys His Asn Ile Ile Ala Asn Ser Thr Phe Ser Trp Trp Gly
225 230 235 240

Ala Trp Leu Asn Asn Asn Ala Gly Lys Asn Val Val Cys Pro Ser Val
245 250 255

Trp Ile Asn Gly Arg Asp Phe Asn Pro Cys Leu Glu Glu Trp Tyr His
260 265 270

Ile

<210> 96

<211> 291

<212> PRT

<213> Bacteroides fragilis

<400> 96

Met Asp Ile Ile Leu Leu His Asn Gly Leu Gly Asn Gln Met Ser Gln
1 5 10 15

Tyr Ala Phe Tyr Leu Ser Lys Lys Lys Asn Gly Ile His Thr Ser Tyr
20 25 30

Ile Cys Leu Ser Asn Asp His Asn Gly Ile Glu Leu Asp Lys Val Phe
35 40 45

Gly Val Glu Cys Gln Met Gly Cys Lys Lys Ile Phe Leu Leu Phe Ile
50 55 60

Leu Arg Leu Leu Met Ser Asn Arg Thr Gly Phe Leu Ile Arg Lys Val
65 70 75 80

37847-517 Sequence Listing_ST25 (003)

Asn Leu Leu Phe Ser Lys Ile Lys Ile Lys Leu Ile Thr Glu Asn Leu
85 90 95

Asp Tyr Ser Phe His Pro Ser Phe Leu Ser Ala Ser Pro Tyr Cys Leu
100 105 110

Ala Phe Trp Val Gly Gly Trp His His Pro Gln Tyr Tyr Ser Glu Ile
115 120 125

Ser Ser Gln Ile Lys Glu Ala Phe Thr Phe Lys Arg Ser Leu Leu Asp
130 135 140

Glu Arg Asn Ile Cys Ile Glu Lys Arg Met Arg Glu Pro Asn Ser Val
145 150 155 160

Cys Leu His Ile Arg Arg Gly Asp Tyr Leu Thr Gly Ile Asn Tyr Glu
165 170 175

Leu Phe Gly Lys Val Cys Asn Glu Gln Tyr Tyr Gln Lys Ala Ile Asp
180 185 190

Tyr Ile Glu Gly Lys Leu Ser Asp Ile Cys Tyr Tyr Val Phe Ser Asn
195 200 205

Asp Met Glu Trp Ala Lys Lys Ile Leu Leu Gly Lys Asn Ala Val Phe
210 215 220

Val Asp Trp Asn Arg Gly Glu Glu Ser Trp Lys Asp Met Tyr Leu Met
225 230 235 240

Ser Lys Cys Ser Asn Leu Ile Ile Pro Asn Ser Thr Phe Ser Trp Trp
245 250 255

Ala Ala Trp Leu Cys Glu His Pro Val Asn Ile Val Cys Pro Lys Leu
260 265 270

Phe Val Tyr Gly Asp Glu Gln Ser Asp Ile Tyr Leu Asp Asn Trp His
275 280 285

37847-517 Sequence Listing_ST25 (003)

Lys Ile Glu
290

<210> 97
<211> 288
<212> PRT
<213> Bacteroides nordii

<400> 97

Met Met Gly Ile Glu Lys Thr Asn Met Val Ile Val Arg Leu Trp Gly
1 5 10 15

Gly Ile Gly Asn Gln Leu Phe Gln Tyr Ser Phe Gly Glu Phe Leu Arg
20 25 30

Glu Lys Tyr Gln Val Asp Val Ile Tyr Asp Ile Ala Ser Phe Gly Lys
35 40 45

Ser Asp Lys Leu Arg Lys Leu Glu Leu Ser Val Val Val Pro Gly Ile
50 55 60

Pro Val Thr Thr Asp Ile Ser Phe Ser Lys Tyr Val Gly Thr Lys Asn
65 70 75 80

Arg Leu Leu Arg Phe Ile Tyr Gly Leu Lys Asn Ser Phe Ile Glu Glu
85 90 95

Lys Tyr Phe Ser Asp Glu Gln Leu Phe Lys Tyr Leu Ser Lys Arg Gly
100 105 110

Asp Val Tyr Leu Gln Gly Tyr Trp Gln Lys Thr Ile Tyr Ala Glu Thr
115 120 125

Leu Arg Arg Lys Gly Ser Phe Phe Leu Ser Gln Glu Glu Pro Ile Val
130 135 140

Leu His Thr Ile Lys Ala Lys Ile Gln Glu Ala Glu Gly Ala Ile Ala
145 150 155 160

37847-517 Sequence Listing_ST25 (003)

Leu His Val Arg Arg Gly Asp Tyr Phe Ser Ser Lys His Ile Asn Thr
 165 170 175

Phe Gly Val Cys Asp Ala His Tyr Tyr Glu Lys Ala Val Asp Ile Met
 180 185 190

Arg Gly Arg Val Ser Asn Ala Met Ile Phe Val Phe Ser Asp Asp Leu
 195 200 205

Asp Trp Val Arg Arg Tyr Val Asn Leu Pro Thr Asn Val Ile Tyr Val
 210 215 220

Pro Asn Tyr Asp Ile Pro Gln Tyr Trp Tyr Ile Tyr Leu Met Ser Leu
 225 230 235 240

Cys Arg His Asn Ile Ile Ser Asn Ser Ser Phe Ser Trp Trp Gly Ala
 245 250 255

Phe Leu Asn Met Asn Thr Asn Lys Ile Val Val Ser Pro Ser Lys Trp
 260 265 270

Thr Leu Asn Ser Asp Lys Thr Ile Ala Leu Asp Glu Trp Phe Lys Ile
 275 280 285

<210> 98

<211> 281

<212> PRT

<213> Butyrivibrio proteoclasticus

<400> 98

Met Glu Cys Ser Met Ile Ile Ile Lys Phe Cys Gly Ala Leu Gly Asn
 1 5 10 15

Gln Leu Phe Gln Tyr Ala Leu Tyr Glu Lys Met Arg Ile Leu Gly Lys
 20 25 30

Asp Val Lys Ala Asp Ile Ser Ala Phe Gly Asp Gly Asn Glu Lys Arg
 35 40 45

Phe Phe Tyr Leu Asp Glu Leu Gly Ile Glu Phe Asn Ile Ala Ser Ala

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

50

55

60

Asp Glu Ile Ala Glu Tyr Leu Asn Arg Lys Thr Ile Arg Phe Val Pro
65 70 75 80

Gly Phe Leu Gln His Arg His Tyr Tyr Phe Glu Lys Lys Pro Tyr Val
85 90 95

Tyr Asn Lys Lys Ile Leu Ser Tyr Asp Asp Cys Tyr Leu Glu Gly Tyr
100 105 110

Trp Gln Asn Tyr Arg Tyr Phe Asp Asp Ile Lys Asp Glu Leu Leu Lys
115 120 125

His Met Lys Phe Pro Cys Leu Pro Leu Glu Gln Lys Lys Leu Ala Glu
130 135 140

Lys Met Glu Asn Glu Asn Ser Val Ala Val His Val Arg Met Gly Asp
145 150 155 160

Tyr Leu Asn Leu Gln Asp Leu Tyr Gly Gly Ile Cys Asp Ala Asp Tyr
165 170 175

Tyr Asp Arg Ala Phe Ser Tyr Ile Glu Gly Asn Ile Ser Asn Pro Val
180 185 190

Tyr Tyr Gly Phe Ser Asp Asp Val Asp Lys Ala Ser Ala Leu Leu Ala
195 200 205

Lys His Lys Ile Asn Trp Ile Asp Tyr Asn Ser Glu Lys Gly Ala Ile
210 215 220

Tyr Asp Leu Ile Leu Met Ser Lys Cys Lys Asn Asn Ile Ile Ala Asn
225 230 235 240

Ser Ser Phe Ser Trp Trp Gly Ala Tyr Leu Glu Tyr Asn Asn Gly Lys
245 250 255

Val Val Val Ser Pro Asn Arg Trp Met Asn Cys Phe Glu Asn Ser Asn

37847-517 Sequence Listing_ST25 (003)

260

265

270

Ile Ala Tyr Trp Gly Trp Ile Ser Leu
 275 280

<210> 99

<211> 289

<212> PRT

<213> Prevotella ruminicola

<400> 99

Met Arg Ile Val Lys Val Leu Gly Gly Leu Gly Asn Gln Met Phe Gln
 1 5 10 15

Phe Ala Leu Tyr Lys Ala Leu Gln Lys Gln Tyr Pro Glu Glu Arg Val
 20 25 30

Leu Leu Asp Leu His Cys Phe Asn Gly Tyr His Lys His Arg Gly Phe
 35 40 45

Glu Ile Asp Ser Val Phe Gly Val Thr Tyr Glu Lys Ala Thr Leu Lys
 50 55 60

Glu Val Ala Ser Leu Ala Tyr Pro Tyr Pro Asn Tyr Gln Cys Trp Arg
 65 70 75 80

Ile Gly Ser Arg Ile Leu Pro Val Arg Lys Thr Met Leu Lys Glu Glu
 85 90 95

Pro Asn Phe Thr Leu Glu Pro Ser Ala Leu Ser Leu Pro Asp Ser Thr
 100 105 110

Tyr Tyr Asp Gly Tyr Trp Gln His Glu Glu Tyr Phe Met His Ile Arg
 115 120 125

Glu Glu Ile Leu Ser Thr Tyr Ala Phe Pro Ala Phe Asp Asp Glu Arg
 130 135 140

Asn Lys Thr Thr Ala Gln Leu Ala Ala Ser Thr Asn Ser Cys Ser Ile
 145 150 155 160

37847-517 Sequence Listing_ST25 (003)

His Ile Arg Arg Gly Asp Tyr Leu Thr Asp Pro Leu Arg Lys Gly Thr
165 170 175

Thr Asn Gly Asn Tyr Val Ile Ala Ala Ile Lys Glu Met Gln Gln Glu
180 185 190

Val Lys Pro Glu Lys Trp Leu Val Phe Ser Asp Asp Ile Ala Trp Cys
195 200 205

Gln Gln His Leu Ala Ser Thr Leu Asp Ala Thr Asn Thr Ile Tyr Ile
210 215 220

Asp Trp Asn Thr Gly Ala Asn Ser Ile His Asp Met His Leu Met Ala
225 230 235 240

Leu Cys Arg His His Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Gly
245 250 255

Ala Trp Leu Ser Gln Gln Asp Gly Ile Thr Ile Ala Pro Ser Asn Trp
260 265 270

Met Asn Leu Lys Asp Val Cys Ser Pro Val Pro Asp Asn Trp Ile Lys
275 280 285

Ile

<210> 100

<211> 290

<212> PRT

<213> Prevotella salivae

<400> 100

Met Lys Ile Ile Lys Ile Ile Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Leu Ala Ile Ala Leu Gln Gln Gln Tyr Lys Asp Glu Glu Ile
20 25 30

37847-517 Sequence Listing_ST25 (003)

Arg Leu Asp Leu Asn Cys Phe Arg Gly Tyr Asn Lys His Gln Gly Tyr
 35 40 45
 Leu Leu Asp Glu Ile Phe Gly Arg Arg Phe Arg Ala Ala Ser Leu Gln
 50 55 60
 Glu Val Ala Arg Leu Ala Trp Pro Tyr Pro His Tyr Gln Leu Trp Arg
 65 70 75 80
 Val Gly Ser Arg Val Leu Pro Arg Arg Gln Thr Met Val Cys Glu Pro
 85 90 95
 Ala Asp Gly Ser Phe Ser Pro Asp Val Leu Thr Leu Glu Gly Asn Arg
 100 105 110
 Tyr Tyr Asp Gly Tyr Trp Gln Asp Glu Arg Tyr Phe Lys Ala Tyr Arg
 115 120 125
 Lys Glu Ile Ile Glu Ala Phe Lys Phe Ser Pro Phe Val Gly Asp Gly
 130 135 140
 Asn Arg His Val Glu Asn Met Leu Arg Asn Glu Arg Phe Ala Ser Leu
 145 150 155 160
 His Val Arg Arg Gly Asp Tyr Leu Asn Asp Ala Leu Tyr Gln Asn Thr
 165 170 175
 Cys Gly Ile Asp Tyr Tyr Gln Arg Ala Ile Ser Gln Met Asn Ala Met
 180 185 190
 Ala Asn Pro Ser Cys Tyr Phe Ile Phe Ser Asp Asp Ile Ala Trp Cys
 195 200 205
 Lys Thr His Ile Glu Pro Leu Cys Glu Gly His Arg Pro Tyr Tyr Ile
 210 215 220
 Asp Trp Asn Lys Gly Lys Glu Ala Tyr Arg Asp Met Gln Leu Met Ala
 225 230 235 240

37847-517 Sequence Listing_ST25 (003)

Leu Cys Lys Tyr His Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Gly
245 250 255

Ala Trp Leu Asn Asp Ala Glu Asp Gly Ile Thr Ile Ala Pro Gln Gln
260 265 270

Trp Tyr Ser His Gly Asn Lys Pro Ser Pro Ala Ser Glu Ser Trp Ile
275 280 285

Lys Val
290

<210> 101

<211> 293

<212> PRT

<213> Lachnospiraceae bacterium COE1

<400> 101

Met Asn Ile Val Arg Ile Ser Asp Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Tyr Ala Arg Lys Ile Ser Ile Leu Ser Arg Gln Arg Thr Tyr
20 25 30

Leu Asp Ile Arg Phe Ile Asn Asn Glu Asp Leu Val Lys Lys Gly Asn
35 40 45

His Val Gln Phe Arg Lys Lys Leu Gly His Arg Lys Tyr Gly Leu Ser
50 55 60

His Phe Asn Val Ser Leu Gln Ile Ala Asp Leu Lys Met Leu Ser His
65 70 75 80

Trp Glu Tyr Leu Ile Gln Ser Asn Cys Met Gln Gln Leu Ile Tyr Ser
85 90 95

Leu Ser Met Gln Asp Lys Trp Ile Trp Arg Tyr Arg His Glu Glu Val
100 105 110

37847-517 Sequence Listing_ST25 (003)

Asn Tyr Asp Gly Met Leu Ser Lys Val Glu Leu Leu Phe Pro Thr Tyr
 115 120 125

Tyr Gln Gly Tyr Phe Phe Ala Leu Lys Tyr Tyr Asp Asp Ile Lys His
 130 135 140

Ile Leu Gln His Asp Phe Ser Leu Lys Asp Lys Met Lys Leu Leu Pro
 145 150 155 160

Glu Leu Arg Asp Ala Leu Tyr Asn Arg Asn Thr Ile Ser Leu His Val
 165 170 175

Arg Arg Gly Asp Phe Leu Glu Ile Asn Arg Asp Ile Ser Gly Ser Glu
 180 185 190

Tyr Tyr Glu Lys Ala Val Gln Met Ile Gly Ser Lys Val Glu Ser Pro
 195 200 205

Ile Phe Leu Ile Phe Ser Asp Asp Ile Glu Trp Val Lys Glu His Ile
 210 215 220

Arg Ile Pro Asn Asp Lys Ile Tyr Val Ser Gly Ile Gly Tyr Glu Asp
 225 230 235 240

Tyr Glu Glu Leu Thr Ile Met Lys His Cys Lys His Asn Ile Ile Ala
 245 250 255

Asn Ser Thr Phe Ser Tyr Trp Ala Ala Tyr Leu Asn Ser Asn Lys Asp
 260 265 270

Lys Ile Val Ile Cys Pro Lys His Trp Arg Glu Arg Ile Ile Pro Lys
 275 280 285

Asp Trp Ile Cys Ile
 290

<210> 102
 <211> 300
 <212> PRT
 <213> Bacteroides dorei

37847-517 Sequence Listing_ST25 (003)

<400> 102

Met Ile Val Val Asn Val Asn Ala Gly Leu Ala Asn Gln Met Phe His
 1 5 10 15

Tyr Ala Phe Gly Arg Gly Leu Glu Ala Lys Gly Trp Asn Ile Tyr Phe
 20 25 30

Asp Gln Thr Asn Phe Lys Pro Arg Lys Glu Trp Ser Phe Glu Asn Val
 35 40 45

Gln Leu Gln Asp Ala Phe Pro Asn Leu Gly Leu Lys Met Met Pro Glu
 50 55 60

Gly Lys Phe Lys Trp Ile Cys Val Asn Asn Thr Asn Lys Leu Ser Lys
 65 70 75 80

Gly Leu His Leu Ala Met Ile Asn Leu His Asn Leu Ile Gly Asp Glu
 85 90 95

Lys Tyr Ile Phe Glu Thr Thr Tyr Gly Tyr Asp Pro Asp Ile Glu Lys
 100 105 110

Glu Ile Thr Lys Asn Cys Ile Leu Lys Gly Phe Trp Gln Ser Glu Lys
 115 120 125

Tyr Phe Ala His Cys Lys Asp Asp Ile Arg Lys Gln Phe Ser Phe Leu
 130 135 140

Pro Phe Asp Glu Glu Lys Asn Ile Val Ile Met Asn Lys Met Val Lys
 145 150 155 160

Glu Asn Ser Val Ala Ile His Leu Arg Lys Gly Ala Asp Tyr Leu Lys
 165 170 175

Ser Glu Leu Met Gly Lys Gly Leu Cys Gly Val Glu Tyr Tyr Ile Lys
 180 185 190

Ala Ile Glu Tyr Ile Lys Lys Asn Ile Asp Asn Pro Val Phe Tyr Val

37847-517 Sequence Listing_ST25 (003)

195 200 205

Phe Thr Asp Asn Pro Val Trp Val Lys Asn Asn Leu Pro Lys Phe Asp
210 215 220

Tyr Ile Leu Val Asp Trp Asn Glu Val Ala Gly Lys Lys Asn Phe Arg
225 230 235 240

Asp Met Gln Leu Met Ser Cys Ala Lys His Asn Ile Ile Ala Asn Ser
245 250 255

Thr Tyr Ser Trp Trp Gly Ala Trp Leu Asn Pro Asn Pro Asn Lys Ile
260 265 270

Val Ile Gly Pro Ala Lys Phe Phe Asn Pro Ile Asn Asn Phe Phe Ser
275 280 285

Ser Ser Asp Ile Met Cys Glu Asp Trp Val Lys Ile
290 295 300

<210> 103
<211> 288
<212> PRT
<213> Roseobacter sp. SK209-2-6

<400> 103

Met Leu Ser Lys Asp Pro Gly Met Ile Thr Thr Arg Leu His Gly Arg
1 5 10 15

Leu Gly Asn Gln Met Phe Gln Tyr Ala Ala Gly Arg Ala Leu Ala Ala
20 25 30

Arg Leu Gly Val Pro Leu Ala Leu Asp Ser Arg Gly Ala Lys Leu Arg
35 40 45

Gly Glu Gly Val Leu Thr Arg Val Phe Asp Leu Pro Leu Ala Gln Pro
50 55 60

Leu Ser Leu Pro Pro Leu Lys Gln Asp Ala Pro Leu Arg Tyr Ala Ala
65 70 75 80

37847-517 Sequence Listing_ST25 (003)

Trp Arg Leu Thr Gly Arg Thr Pro Arg Phe Arg Arg Glu Gln Gly Leu
 85 90 95

Gly Tyr Asn Pro Ala Phe Glu Thr Trp Gly Asp Asp Ser Tyr Leu His
 100 105 110

Gly Tyr Trp Gln Ser Glu Ala Tyr Phe Asp Ser Ile Ala Asp Gln Ile
 115 120 125

Arg Gln Asp Phe Thr Phe Pro Glu Phe Ser Asn Ser Gln Asn Arg Glu
 130 135 140

Met Ala Gln Arg Ile Ala Gly Ser Thr Ala Ile Ser Leu His Val Arg
 145 150 155 160

Arg Gly Asp Tyr Val Ala Leu Ala Ala His Val Leu Cys Asp Gln Ala
 165 170 175

Tyr Tyr Glu Ala Ala Leu Thr Arg Ile Leu Glu Gly Val Glu Gly Ser
 180 185 190

Pro Thr Val Tyr Val Phe Ser Asp Asp Pro Asn Trp Ala Lys Glu Asn
 195 200 205

Leu Pro Leu Pro Cys Glu Lys Val Val Val Asp Phe Asn Gly Pro Asp
 210 215 220

Thr Asp Phe Glu Asp Met Arg Leu Met Ser Leu Cys Gln His Asn Ile
 225 230 235 240

Ile Gly Asn Ser Ser Phe Ser Trp Trp Ala Ala Trp Leu Asn Thr His
 245 250 255

Asn Glu Lys Arg Val Ala Gly Pro Ala His Trp Phe Gly Asn Pro Lys
 260 265 270

Leu Gln Asn Pro Asp Ile Leu Pro Glu Ser Trp Leu Lys Ile Ser Val
 275 280 285

37847-517 Sequence Listing_ST25 (003)

28 May 2018
2015259088

<210> 104
 <211> 285
 <212> PRT
 <213> alpha proteobacterium SCGC AAA076-C03
 <400> 104

Met Ile Tyr Ser Arg Ile Arg Gly Gly Leu Gly Asn Gln Leu Phe Gln
 1 5 10 15

Tyr Cys Val Ala Arg Ser Leu Ala Asp Asn Leu Gly Thr Ser Leu Gly
 20 25 30

Leu Asp Val Arg Asp Phe Asn Glu Asn Ser Pro Tyr Leu Met Gly Leu
 35 40 45

Lys His Phe Asn Ile Arg Ala Asp Phe Asn Pro Pro Gly Met Ile Glu
 50 55 60

His Lys Lys Asn Gly Tyr Phe Arg Tyr Leu Ile Asp Val Val Asn Gly
 65 70 75 80

Lys Gln Lys Phe Val Tyr Lys Glu Pro His Leu Asn Phe Asp Lys Asn
 85 90 95

Ile Phe Ser Leu Pro Asn Ser Ser Tyr Leu Lys Gly Tyr Trp Gln Thr
 100 105 110

Glu Lys Tyr Phe Ile Lys Asn Lys Val Asn Ile Leu Asn Asp Leu Lys
 115 120 125

Ile Ile Ser His Gln Ser Asp Lys Asn Lys Thr Ile Ser Ser Lys Ile
 130 135 140

Ala Asn Asn Thr Ser Val Ser Leu His Ile Arg Arg Gly Asp Tyr Ile
 145 150 155 160

Ser Asn Ser Ala Tyr Asn Ser Thr His Gly Thr Cys Ser Leu Ala Tyr
 165 170 175

37847-517 Sequence Listing_ST25 (003)

Tyr Thr Asn Ala Val Asn Phe Leu Val Asn Lys Ile Gly Gly Asn Phe
180 185 190

Lys Val Phe Ala Phe Ser Asp Asp Pro Glu Trp Val Ser Ser Asn Leu
195 200 205

Lys Leu Pro Val Asp Ile Cys Phe Val Lys Asn Asn Ser Ser Glu Tyr
210 215 220

Asn Tyr Glu Asp Leu Arg Leu Met Ser Glu Cys Asn His Asn Ile Ile
225 230 235 240

Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala Trp Leu Asn Thr Asn His
245 250 255

Asn Lys Thr Val Ile Thr Pro Cys Lys Trp Tyr Ala Asp Asn Ser Thr
260 265 270

Lys Asn Ala Asp Ile Thr Pro Ser Asn Trp Ile Lys Ile
275 280 285

<210> 105
<211> 292
<212> PRT
<213> Helicobacter bilis

<400> 105

Met Gly Gly Gly Gly Gln Asp Leu Arg Leu Phe Glu Leu Met Leu Tyr
1 5 10 15

Asn Ile Ser Leu Pro Leu Cys Phe Asp Tyr Lys Thr Leu Val Lys Tyr
20 25 30

Phe Tyr Ser Asn Asp Lys Ser Leu Lys Tyr Asn Phe Pro Leu Gln Tyr
35 40 45

Ile Arg Tyr Ala Thr Arg Ser Lys Tyr His Lys Leu Tyr Trp Leu Ala
50 55 60

37847-517 Sequence Listing_ST25 (003)

Leu	Lys	His	Tyr	Lys	Tyr	Phe	Tyr	Asp	Glu	Asp	Pro	Gln	Gly	Asp	Asn	
65					70					75					80	
Ile	Val	Lys	Met	Tyr	Leu	Asn	Asn	Ser	Leu	Glu	Lys	His	Ala	Tyr	Pro	
				85					90					95		
Phe	Gly	Tyr	Phe	Gln	Asn	Leu	Ile	Tyr	Phe	Asp	Glu	Ile	Asp	Ser	Ile	
			100					105					110			
Ile	Arg	Glu	Glu	Phe	Cys	Leu	Lys	Ile	Pro	Leu	Lys	Pro	His	Asn	Gln	
		115					120					125				
Ala	Leu	Lys	Glu	Lys	Ile	Glu	Lys	Thr	Glu	Asn	Ser	Val	Phe	Leu	His	
		130				135					140					
Val	Arg	Leu	Gly	Asp	Tyr	Leu	Lys	Met	Glu	Ala	Thr	Asp	Gly	Gly	Tyr	
145					150					155					160	
Val	Arg	Leu	Gly	Lys	Thr	Tyr	Tyr	Gln	Ser	Ala	Leu	Glu	Ile	Leu	Lys	
				165				170						175		
Thr	Arg	Leu	Gly	Gln	Pro	His	Ile	Phe	Ile	Phe	Ser	Asn	Asp	Ile	Glu	
			180					185					190			
Trp	Cys	Glu	Lys	Asn	Leu	Cys	Asn	Leu	Leu	Asp	Phe	Thr	Gly	Cys	His	
		195					200					205				
Ile	Glu	Phe	Val	Lys	Ala	Asn	Gly	Glu	Gly	Asn	Ala	Ala	Glu	Glu	Met	
	210					215					220					
Glu	Leu	Met	Arg	Ala	Cys	Lys	His	Ala	Val	Ile	Ala	Asn	Ser	Thr	Phe	
225					230					235					240	
Ser	Trp	Trp	Ala	Ser	Tyr	Leu	Ile	Asp	Asn	Pro	Asp	Lys	Gln	Ile	Ile	
				245					250					255		
Met	Pro	Thr	Gln	Val	Phe	Asn	Asp	Thr	Arg	Arg	Ile	Pro	Lys	Ser	Asn	
			260					265					270			

37847-517 Sequence Listing_ST25 (003)

Met Leu Ala Lys Lys Gly Tyr Ile Leu Ile Asp Pro Phe Trp Gly Met
275 280 285

His Ser Ile Val
290

<210> 106
<211> 288
<212> PRT
<213> Ralstonia sp. GA3-3

<400> 106

Met Ile Val Thr Arg Val Ile Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Ala Gly Arg Ala Leu Ala Arg Arg Leu Gly Val Pro Leu Lys
20 25 30

Ile Asp Ser Ser Gly Phe Ala Asp Tyr Pro Leu His Asn Tyr Gly Leu
35 40 45

His His Phe Ala Leu Lys Ala Val Gln Ala Gly Asp Arg Glu Ile Pro
50 55 60

Ser Gly Arg Ala Glu Asn Arg Trp Ala Lys Ala Leu Arg Arg Phe Gly
65 70 75 80

Leu Gly Thr Glu Leu Arg Val Phe Arg Glu Arg Gly Phe Ala Val Asp
85 90 95

Pro Glu Val Met Lys Leu Pro Asp Gly Thr Tyr Leu Asp Gly Tyr Trp
100 105 110

Gln Ser Glu Ser Tyr Phe Ala Glu Met Thr Gln Glu Leu Arg Arg Asp
115 120 125

Phe Gln Ile Ala Thr Pro Pro Thr Ser Glu Asn Ala Glu Trp Leu Ala
130 135 140

Arg Ile Gly Gly Asp Glu Gly Ala Val Ser Ile His Val Arg Arg Gly

37847-517 Sequence Listing_ST25 (003)

145 150 155 160

Asp Tyr Val Thr Asn Ala Ser Ala Asn Ala Val His Gly Ile Cys Ser
165 170 175

Leu Asp Tyr Tyr Met Arg Ala Ala Arg Tyr Val Ala Glu Asn Ile Gly
180 185 190

Val Lys Pro Thr Phe Tyr Val Phe Ser Asp Asp Pro Asp Trp Val Ala
195 200 205

Gly Asn Leu His Leu Gly His Glu Thr Arg Tyr Val Arg His Asn Asp
210 215 220

Ser Ala Arg Asn Tyr Glu Asp Leu Arg Leu Met Ser Ala Cys Arg His
225 230 235 240

His Ile Ile Ala Asn Ser Thr Phe Ser Trp Trp Gly Ala Trp Leu Asn
245 250 255

Ala Ser Glu Lys Lys Val Val Ile Ala Pro Ala Gln Trp Phe Arg Asp
260 265 270

Glu Lys Tyr Asp Thr Arg Asp Leu Leu Pro Pro Thr Trp Thr Lys Leu
275 280 285

<210> 107
<211> 289
<212> PRT
<213> Bacteroides ovatus

<400> 107

Met Val Val Val Tyr Ile Ala Ala Gly Leu Ala Asn Lys Met Phe Gln
1 5 10 15

Tyr Ala Phe Ser Arg Gly Leu Met Ser His Gly Leu Asp Val Phe Leu
20 25 30

Asp Gln Thr Ser Phe Gln Pro Glu Trp Ser Phe Glu Asp Ile Ala Leu
35 40 45

37847-517 Sequence Listing_ST25 (003)

Glu Glu Val Phe Pro Asn Ile Glu Ile Lys Asn Ala Pro Asn Asn Met
50 55 60

Phe Ser Leu Ala Tyr Lys Lys Asp Leu Leu Ser Arg Ile Tyr Arg Arg
65 70 75 80

Met Ser Ala Phe Phe Pro Asn Asn Arg Tyr Leu Met Glu Arg Pro Phe
85 90 95

Ile Tyr Asp Glu Leu Ile Tyr Lys Lys Ala Thr Asn Asn Cys Ile Phe
100 105 110

Cys Gly Leu Trp Gln Thr Glu Leu Tyr Phe Asn Phe Cys Glu Arg Asp
115 120 125

Val Arg Arg Asn Phe Val Phe Thr Pro Phe Gln Asp Asp Gln Asn Ile
130 135 140

Lys Leu Ala Glu Lys Met Lys Asn Glu Asn Ser Val Ala Ile His Ile
145 150 155 160

Arg Lys Gly Ala Asp Tyr Leu Lys Arg Asn Ile Trp Asp Gly Thr Cys
165 170 175

Ser Val Glu Tyr Tyr Asn Gln Ala Ile Asn Tyr Leu Lys Glu His Val
180 185 190

Ser Asn Pro Val Phe Tyr Leu Phe Thr Asp Asn Pro Glu Trp Val Glu
195 200 205

Glu Asn Leu Lys Asn Ile Asp Tyr Lys Leu Val Asp Trp Asn Pro Val
210 215 220

Ser Gly Lys Gln Ser Tyr Arg Asp Met Gln Leu Met Ser Cys Ala Lys
225 230 235 240

His Asn Ile Ile Ala Asn Ser Thr Tyr Ser Trp Trp Gly Ala Trp Leu
245 250 255

37847-517 Sequence Listing_ST25 (003)

Asn Asn Asn Pro Gln Lys Ile Val Val Ala Pro Lys Ile Trp Phe Asn
260 265 270

Pro Lys Ile Glu Lys Ala Pro Tyr Ile Ile Pro Asp Arg Trp Ile Arg
275 280 285

Leu

<210> 108

<211> 294

<212> PRT

<213> Loktanella vestfoldensis

<400> 108

Met Ile Ile Thr Lys Leu Ile Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Ala Gly Arg Ser Leu Ala Met Arg His Gly Val Pro Leu Leu
20 25 30

Leu Asp Ile Thr Glu Leu Arg Ser Tyr Pro Lys His Gln Gly Tyr Gln
35 40 45

Phe Glu Asp Val Phe Ala Gly Arg Phe Glu Ile Ala Gly Leu Ile Pro
50 55 60

Leu Ile Arg Val Leu Gly Arg Lys Ala Arg Lys Val Pro Lys Thr Val
65 70 75 80

Ala Val Val Ser Pro Lys Trp Pro Pro Met Gly Asp His Val Trp Val
85 90 95

Arg Gln Arg Thr His Asp Tyr Asp Ala Ala Phe Glu Ser Ile Gly Ala
100 105 110

Asp Cys Tyr Leu Ser Gly Phe Trp Gln Ser Glu Lys Tyr Phe Ala Thr
115 120 125

37847-517 Sequence Listing_ST25 (003)

Ile Ala Pro Gln Ile Arg Glu Ser Phe Arg Phe Lys Glu Ala Leu Thr
 130 135 140

Gly Ala Asn Ala Ala Ile Ala Ser Arg Met Lys Glu Ala Pro Ser Ala
 145 150 155 160

Ala Ile His Ile Arg Arg Gly Asp Tyr Val Thr Asp Lys Gly Ala His
 165 170 175

Ala Phe His Gly Leu Cys Ala Trp Asp Tyr Tyr Asp Ala Ala Ile Asp
 180 185 190

His Ile Ser Arg His Glu Pro Asp Ala Arg Phe Phe Val Phe Ser Asp
 195 200 205

Asp Val Val Ala Ala Gln Glu Arg Phe Ala Asn Arg Gln Arg Ala Glu
 210 215 220

Val Val Ala Val Asn Ser Gly Arg His Ser Tyr Arg Asp Met Met Leu
 225 230 235 240

Met Ala Gln Cys Lys His Gln Ile Ile Ala Asn Ser Thr Phe Ser Trp
 245 250 255

Trp Ala Ala Trp Leu Asn Gln Asn Pro Asp Lys Ile Val Val Ala Pro
 260 265 270

Gly Thr Trp Phe Ser Gly Asn Asp Gly Gln Ile Lys Asp Ile Tyr Cys
 275 280 285

Lys Asp Trp Ile Val Ile
 290

<210> 109

<211> 292

<212> PRT

<213> Flavobacterium sp. ACAM 123

<400> 109

37847-517 Sequence Listing_ST25 (003)

Met	Asp	Val	Val	Ile	Ile	Phe	Asn	Gly	Leu	Gly	Asn	Gln	Met	Ser	Gln	
1				5				10					15			
Tyr	Ala	Phe	Tyr	Ser	Gln	Lys	Lys	Lys	Ile	Asn	Asn	Ser	Thr	Tyr	Phe	
			20					25					30			
Val	Pro	Phe	Cys	Lys	Asp	His	Asn	Gly	Leu	Glu	Leu	Glu	Thr	Val	Phe	
		35					40					45				
Ser	Leu	Asn	Thr	Lys	Glu	Thr	Leu	Ile	Gln	Lys	Ser	Leu	Tyr	Ile	Leu	
	50					55					60					
Phe	Arg	Ile	Leu	Leu	Thr	Asp	Arg	Leu	Lys	Ile	Val	Ser	Asp	Pro	Leu	
65					70					75					80	
Lys	Trp	Ile	Leu	Asn	Leu	Phe	Lys	Cys	Lys	Ile	Val	Lys	Glu	Ser	Phe	
			85						90					95		
Asn	Tyr	Asn	Tyr	Asn	Pro	Glu	Tyr	Leu	Lys	Pro	Ser	Lys	Gly	Ile	Thr	
			100					105					110			
Phe	Tyr	Tyr	Gly	Gly	Trp	His	Ala	Glu	Lys	Tyr	Phe	Ala	Lys	Glu	Asn	
		115					120					125				
Gln	Gln	Ile	Lys	Ser	Val	Phe	Glu	Phe	Thr	Gly	Asp	Leu	Gly	Lys	Ile	
	130					135					140					
Asn	Lys	Glu	His	Val	Lys	Asp	Ile	Ala	Ser	Thr	Asn	Ala	Val	Ser	Leu	
145					150					155					160	
His	Val	Arg	Arg	Gly	Asp	Phe	Met	Asn	Glu	Ala	Asn	Ile	Gly	Leu	Phe	
				165					170					175		
Gly	Gly	Val	Ser	Thr	Lys	Ala	Tyr	Phe	Glu	Gly	Ala	Ile	Lys	Leu	Ile	
			180					185					190			
Ala	Thr	Lys	Val	Asp	His	Pro	His	Phe	Phe	Val	Phe	Ser	Asn	Asp	Met	
		195					200					205				

37847-517 Sequence Listing_ST25 (003)

Asp Trp Val Lys Glu Asn Leu Ser Met Asp Thr Val Thr Tyr Val Thr
210 215 220

Cys Asn Ser Gly Lys Asp Ser Trp Lys Asp Met Cys Leu Met Ser Leu
225 230 235 240

Cys Gln His Asn Ile Ile Pro Asn Ser Thr Phe Ser Trp Trp Gly Ala
245 250 255

Trp Leu Asn Lys Asn Pro His Lys Ile Val Val Cys Pro Ser Arg Phe
260 265 270

Leu Asn Asn Asp Thr Tyr Thr Asp Ile Tyr Pro Asp Ser Trp Val Lys
275 280 285

Ile Ser Asp Tyr
290

<210> 110
<211> 283
<212> PRT
<213> Bacteroides fragilis

<400> 110

Met Met Lys Leu Val Arg Met Thr Gly Gly Leu Gly Asn Gln Met Phe
1 5 10 15

Ile Tyr Ala Phe Tyr Ile Gln Met Lys Thr Ile Phe Pro Glu Leu Arg
20 25 30

Ile Asp Met Ser Glu Met Lys Lys Tyr Lys Leu His Asn Gly Tyr Glu
35 40 45

Leu Glu Asp Val Phe Ser Ile Arg Pro Gln Thr Ile Ser Ala His Lys
50 55 60

Trp Leu Lys Arg Val Ile Val Tyr Ala Phe Phe Ser Ile Ile Arg Glu
65 70 75 80

Lys Ser Glu Glu Glu Leu Ser Ile His Lys Tyr Thr Gln His Lys Arg

37847-517 Sequence Listing_ST25 (003)

85 90 95

Trp Pro Leu Val Tyr Tyr Lys Gly Phe Phe Gln Ser Glu Leu Phe Phe
100 105 110

Lys Glu Ser Ser Asp Thr Ile Arg Asp Ile Phe Ser Phe Asn Thr Glu
115 120 125

Asn Ala Asn Phe Arg Thr Lys Glu Trp Ala Lys Ile Ile Lys Glu Gln
130 135 140

Arg Ser Ser Val Ser Ile His Ile Arg Arg Gly Asp Tyr Thr Ser Ala
145 150 155 160

Lys Asn Lys Ile Lys Tyr Gly Asn Ile Cys Thr Glu Glu Tyr Tyr Gln
165 170 175

Lys Ala Ile Ser Ile Ile Leu Lys Lys Glu Pro Lys Ala Phe Phe His
180 185 190

Ile Phe Ser Asp Asp Val Glu Trp Thr Lys Ala His Leu Lys Ile His
195 200 205

His Leu Pro His Gln Tyr Ile Ser Trp Asn Lys Gly Pro Asp Ser Trp
210 215 220

Gln Asp Met Met Leu Met Ser Leu Cys Arg His Asn Ile Ile Ala Asn
225 230 235 240

Ser Ser Phe Ser Trp Trp Gly Ala Trp Leu Asn Ala Tyr Lys Asp Lys
245 250 255

Thr Val Ile Ala Pro Ser Arg Trp Ser Asn Val Lys Lys Thr Pro His
260 265 270

Ile Leu Pro Glu Ser Trp Ile Ser Ile Asp Ile
275 280

<210> 111

37847-517 Sequence Listing_ST25 (003)

<211> 290
 <212> PRT
 <213> Spirosoma panaciterrae

<400> 111

Met Ile Ile Ser Arg Val Thr Ser Gly Leu Gly Asn Gln Leu Phe Gln
 1 5 10 15

Tyr Ala Ala Ala Arg Ser Leu Ser Leu Arg Asn Lys Thr Ala Phe Tyr
 20 25 30

Val Asp Leu Ser Tyr Tyr Leu Tyr Glu Tyr Pro Asp Asp Thr Ser Arg
 35 40 45

Ser Phe Lys Leu Gly Phe Phe Ser Val Pro Tyr Arg Ile Leu Gln Glu
 50 55 60

Ser Pro Val Glu Tyr Leu Ser Lys Ser Thr Lys Leu Phe Pro Asn Arg
 65 70 75 80

Ser Leu Arg Pro Phe Phe Leu Phe Leu Lys Glu Lys Gln Phe His Phe
 85 90 95

Asp Pro Thr Ile Leu Gln Ala His Ala Gly Cys Val Ile Met Glu Gly
 100 105 110

Phe Trp Gln Ser Glu Cys Tyr Phe Arg Asp His Ala Glu Ile Ile Arg
 115 120 125

Arg Glu Leu Gln Leu Ser Lys Ser Pro Ser Ser Glu Phe Glu Gly Tyr
 130 135 140

His Gln Gln Ile Gln Ala Thr Pro Val Pro Val Ser Val His Val Arg
 145 150 155 160

Arg Gly Asp Tyr Val Asn His Pro Glu Phe Ser Lys Thr Phe Gly Phe
 165 170 175

Ile Gly Leu Asp Tyr Tyr Lys Thr Ala Ile Arg His Leu Thr Lys Thr
 180 185 190

37847-517 Sequence Listing_ST25 (003)

Ile Lys Asn Pro His Phe Tyr Val Phe Ser Asp Asp Lys Glu Trp Ala
195 200 205

Arg Ala Asn Leu Pro Leu Pro Thr Asp Ser Val Phe Val Thr Asn Thr
210 215 220

Gly Pro Ser Gly Asp Val Ala Asp Leu Val Leu Met Ser Thr Cys His
225 230 235 240

His His Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala Trp Leu
245 250 255

Asn Pro Asn Pro Asp Lys Leu Val Ile Thr Pro Lys Leu Trp Tyr Lys
260 265 270

Asn Gln Pro Thr Trp Asn Thr Lys Asp Leu Leu Pro Pro Thr Trp Val
275 280 285

Ser Leu
290

<210> 112
<211> 288
<212> PRT
<213> Unknown

<220>
<223> uncultured bacterium

<400> 112

Met Ile Ile Thr Lys Leu Thr Gly Gly Leu Gly Asn Gln Leu Phe Gln
1 5 10 15

Tyr Ala Ile Gly Arg Asn Leu Ile Tyr Ile Asn Gly Ser Asp Leu Lys
20 25 30

Leu Asp Val Ser Glu Tyr Asp Val Ser Asn Lys Gly Asn Phe Arg His
35 40 45

37847-517 Sequence Listing_ST25 (003)

Tyr	Ala	Leu	Asp	Lys	Phe	Asn	Thr	Ile	Gln	Asn	Phe	Ala	Ser	Lys	Lys		
50						55					60						
Glu	Thr	Asn	Asn	Phe	Lys	Phe	Gly	Val	Phe	Lys	Lys	Trp	Leu	Tyr	Lys		
65					70					75					80		
Ser	Gly	Ile	Val	Lys	Asn	Lys	Asn	Tyr	Phe	Leu	Glu	Lys	Lys	Phe	Asn		
				85					90					95			
Phe	Asp	Lys	Glu	Ile	Leu	Lys	Ile	Lys	Asp	Asn	Ala	Phe	Leu	Gln	Gly		
			100					105						110			
Tyr	Trp	Gln	Ser	Glu	Lys	Tyr	Phe	Ile	Gly	Ile	Arg	Asp	Ile	Leu	Leu		
		115					120					125					
Gln	Glu	Phe	Ser	Leu	Lys	Glu	Asn	Ile	Glu	Leu	Lys	Phe	Gly	Glu	Ile		
	130					135					140						
Leu	Lys	Glu	Ile	Asn	Glu	Ser	Asn	Ser	Val	Ser	Ile	His	Val	Arg	Arg		
145					150					155					160		
Gly	Asp	Tyr	Val	Lys	Asn	Pro	Lys	Asn	Leu	Ser	Phe	His	Gly	Val	Cys		
				165					170					175			
Ser	Pro	Lys	Tyr	Tyr	Ser	Glu	Ser	Thr	Ser	Lys	Ile	Ala	Ser	Leu	Ile		
			180					185					190				
Glu	Lys	Pro	Val	Phe	Phe	Val	Phe	Ser	Asp	Asp	Ile	Glu	Trp	Val	Lys		
		195					200					205					
Glu	Asn	Leu	Asn	Ile	Thr	Phe	Pro	Val	Val	Tyr	Leu	Ser	Gly	Ile	Lys		
	210					215					220						
Asn	Ile	Lys	Ser	Tyr	Glu	Glu	Leu	Val	Leu	Met	Ser	Lys	Cys	Lys	His		
225					230					235					240		
Asn	Ile	Ile	Ala	Asn	Ser	Ser	Phe	Ser	Trp	Trp	Gly	Ala	Trp	Leu	Asn		
				245					250					255			

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

Thr Asn Gln Lys Lys Ile Val Ile Ala Pro Lys Arg Trp Phe Asn Asp
260 265 270

Val Lys Leu Asp Thr Thr Asp Leu Ile Pro Glu Asn Trp Ile Arg Ile
275 280 285

<210> 113

<211> 293

<212> PRT

<213> Thermosynechococcus elongatus

<400> 113

Met Ile Ile Val His Leu Cys Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Ala Gly Leu Ala Ala Ala His Arg Ile Gly Ser Glu Val Lys
20 25 30

Phe Asp Thr His Trp Phe Asp Ala Thr Cys Leu His Gln Gly Leu Glu
35 40 45

Leu Arg Arg Val Phe Gly Leu Glu Leu Pro Glu Pro Ser Ser Lys Asp
50 55 60

Leu Arg Lys Val Leu Gly Ala Cys Val His Pro Ala Val Arg Arg Leu
65 70 75 80

Leu Ala Gly His Phe Leu His Gly Leu Arg Pro Lys Ser Leu Val Ile
85 90 95

Gln Pro His Phe His Tyr Trp Thr Gly Phe Glu His Leu Pro Asp Asn
100 105 110

Val Tyr Leu Glu Gly Tyr Trp Gln Ser Glu Arg Tyr Phe Ser Asn Ile
115 120 125

Ala Asp Ile Ile Arg Gln Gln Phe Arg Phe Val Glu Pro Leu Asp Pro
130 135 140

His Asn Ala Ala Leu Met Asp Glu Met Gln Ser Gly Val Ser Val Ser

28 May 2018
2015259088

37847-517 Sequence Listing_ST25 (003)

145 150 155 160

Leu His Ile Arg Arg Gly Asp Tyr Phe Asn Asn Pro Gln Met Arg Arg
165 170 175

Val His Gly Val Asp Leu Ser Glu Tyr Tyr Pro Ala Ala Val Ala Thr
180 185 190

Met Ile Glu Lys Thr Asn Ala Glu Arg Phe Tyr Val Phe Ser Asp Asp
195 200 205

Pro Gln Trp Val Leu Glu His Leu Lys Leu Pro Val Ser Tyr Thr Val
210 215 220

Val Asp His Asn Arg Gly Ala Ala Ser Tyr Arg Asp Met Gln Leu Met
225 230 235 240

Ser Ala Cys Arg His His Ile Ile Ala Asn Ser Thr Phe Ser Trp Trp
245 250 255

Gly Ala Trp Leu Asn Pro Arg Pro Asp Lys Val Val Ile Ala Pro Arg
260 265 270

His Trp Phe Asn Val Asp Val Phe Asp Thr Arg Asp Leu Tyr Cys Pro
275 280 285

Gly Trp Ile Val Leu
290

<210> 114
<211> 293
<212> PRT
<213> Colwellia piezophila

<400> 114

Met Lys Ile Val Lys Ile Ala Gly Gly Phe Gly Asn Gln Leu Phe Gln
1 5 10 15

Tyr Ala Phe Tyr Leu Ala Leu Asp Lys Lys Tyr Ala Glu Gln Val Cys
20 25 30

37847-517 Sequence Listing_ST25 (003)

Leu Asp Ser Leu Asp Met Ala Lys Tyr Arg Leu His Asn Gly Tyr Glu
 35 40 45
 Leu Glu Gly Ile Phe Lys Leu Asp Ala Arg Tyr Cys Thr Glu Glu Gln
 50 55 60
 Arg Ile Ile Val Arg Lys Asp Asn Asn Ile Phe Thr Lys Leu Leu Ser
 65 70 75 80
 Ser Leu Lys Lys Lys Leu Gly Asn Asn Lys Asn Tyr Ile Leu Glu Pro
 85 90 95
 Lys Gln Glu His Phe Thr Phe His Glu Lys Ser Phe Gly Gln Ala Asn
 100 105 110
 Thr Pro Thr Tyr Tyr Lys Gly Tyr Trp Gln Asp Val Lys Tyr Leu Glu
 115 120 125
 Asn Ile Glu Glu Glu Leu Lys Ser Ser Leu Val Phe Pro Glu Phe Glu
 130 135 140
 Leu Gly Lys Asn Ile Glu Leu Ala Asn Phe Ile Ser Ser Asn Ser Ser
 145 150 155 160
 Val Ser Leu His Val Arg Arg Gly Asp Tyr Val Gln His Lys Ala Phe
 165 170 175
 Gly Gly Ile Cys Asp Leu Ser Tyr Tyr Gln Arg Ala Val Glu Gln Ile
 180 185 190
 Asn Thr Leu Val Lys Asp Pro Ile Phe Ile Val Phe Ser Asp Asp Ile
 195 200 205
 Gln Trp Cys Lys Asp Asn Leu Asn Leu Glu Lys Ala Lys Phe Val Asp
 210 215 220
 Trp Asn Ile Gly Glu Asn Ser Phe Arg Asp Met Gln Leu Met Thr Leu
 225 230 235 240

37847-517 Sequence Listing_ST25 (003)

Cys Lys His Asn Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala
245 250 255

Trp Leu Asn Ala Asn Asp Asp Lys Asn Val Ile Cys Pro Asp Lys Trp
260 265 270

Val His Tyr Thr Ser Ala Thr Gly Val Leu Pro Ser Glu Trp Ile Lys
275 280 285

Ile Lys Ala Ser Val
290

<210> 115

<211> 289

<212> PRT

<213> Prevotella maculosa

<400> 115

Met Lys Ile Val Lys Ile Ile Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Leu Ala Ile Ala Leu Gln Glu Arg Trp Lys Asp Glu Glu Ile
20 25 30

Lys Leu Asp Leu His Gly Phe Asn Gly Tyr His Lys His Gln Gly Tyr
35 40 45

Gln Leu Asp Met Leu Phe Gly His Arg Phe Glu Ala Ala Thr Leu Thr
50 55 60

Asp Val Ala Gln Leu Ala Trp Pro Tyr Pro His Tyr Gln Leu Trp Arg
65 70 75 80

Val Gly Ser Arg Leu Leu Pro Lys Arg Arg Ser Met Leu Cys Glu Pro
85 90 95

Ser Lys Gly Leu Leu Pro Ser Asp Val Leu Lys Gln Lys Gly Ser Leu
100 105 110

37847-517 Sequence Listing_ST25 (003)

Tyr Tyr Asp Gly Tyr Trp Gln Asp Glu Arg Tyr Phe Arg Ala Ile Arg
115 120 125

Pro Gln Ile Met Ala Ala Phe Lys Phe Pro Asp Phe Thr Asp Arg Arg
130 135 140

Asn Leu Glu Thr Glu Lys Arg Leu Lys Ala Ser Glu Ala Val Ser Ile
145 150 155 160

His Val Arg Arg Gly Asp Tyr Leu Asp Asp Val Leu Phe Gln Gly Thr
165 170 175

Cys Asn Ile Ala Tyr Tyr Gln Arg Ala Ile Ala Arg Leu Cys Gln Leu
180 185 190

Lys Thr Pro Val Phe Cys Ile Phe Ser Asn Asp Met Ala Trp Cys Lys
195 200 205

Val His Ile Glu Pro Leu Leu His Gly Lys Glu Ile Leu Tyr Val Asp
210 215 220

Trp Asn Arg Gly Lys Glu Ser Tyr Arg Asp Leu Gln Leu Met Thr Leu
225 230 235 240

Cys Arg His His Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala
245 250 255

Trp Leu Ser Lys Ala Glu Asp Gly Ile Thr Ile Ala Pro Arg His Trp
260 265 270

Tyr Ala His Asp Ala Lys Pro Ser Pro Ala Ala Glu Arg Trp Ile Lys
275 280 285

Val

<210> 116

<211> 298

<212> PRT

37847-517 Sequence Listing_ST25 (003)

<213> Salmonella enterica

<400> 116

Met Tyr Ser Cys Leu Ser Gly Gly Leu Gly Asn Gln Met Phe Gln Tyr
1 5 10 15

Ala Ala Ala Tyr Ile Leu Lys Gln Tyr Phe Gln Ser Thr Thr Leu Val
20 25 30

Leu Asp Asp Ser Tyr Tyr Tyr Ser Gln Pro Lys Arg Asp Thr Val Arg
35 40 45

Ser Leu Glu Leu Asn Gln Phe Asn Ile Ser Tyr Asp Arg Phe Ser Phe
50 55 60

Ala Asp Glu Lys Glu Lys Ile Lys Leu Leu Arg Lys Phe Lys Arg Asn
65 70 75 80

Pro Phe Pro Lys Gln Ile Ser Glu Ile Leu Ser Ile Ala Leu Phe Gly
85 90 95

Lys Tyr Ala Leu Ser Asp Arg Ala Phe Tyr Thr Phe Glu Thr Ile Lys
100 105 110

Asn Ile Asp Lys Ala Cys Leu Phe Ser Phe Tyr Gln Asp Ala Asp Leu
115 120 125

Leu Asn Lys His Lys Gln Leu Ile Leu Pro Leu Phe Glu Leu Arg Asp
130 135 140

Asp Leu Leu Asp Ile Cys Lys Asn Leu Glu Leu Tyr Ser Leu Ile Gln
145 150 155 160

Arg Ser Asn Asn Thr Thr Ala Leu His Ile Arg Arg Gly Asp Tyr Val
165 170 175

Thr Asn Gln His Ala Ala Lys Tyr His Gly Val Leu Asp Ile Ser Tyr
180 185 190

37847-517 Sequence Listing_ST25 (003)

Tyr Asn His Ala Met Glu Tyr Val Glu Arg Glu Arg Gly Lys Gln Asn
195 200 205

Phe Ile Ile Phe Ser Asp Asp Val Arg Trp Ala Gln Lys Ala Phe Leu
210 215 220

Glu Asn Asp Asn Cys Tyr Val Ile Asn Asn Ser Asp Tyr Asp Phe Ser
225 230 235 240

Ala Ile Asp Met Tyr Leu Met Ser Leu Cys Lys Asn Asn Ile Ile Ala
245 250 255

Asn Ser Thr Tyr Ser Trp Trp Gly Ala Trp Leu Asn Lys Tyr Glu Asp
260 265 270

Lys Leu Val Ile Ser Pro Lys Gln Trp Phe Leu Gly Asn Asn Glu Thr
275 280 285

Ser Leu Arg Asn Ala Ser Trp Ile Thr Leu
290 295

<210> 117

<211> 279

<212> PRT

<213> Bacteroides sp. 3_2_5

<400> 117

Met Lys Lys Val Ile Phe Ser Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Phe Tyr Leu Phe Leu Lys Lys Lys Gly Ile Lys Ala Val Ile
20 25 30

Asp Asn Ser Leu Tyr Ser Glu Phe Lys Met His Asn Gly Phe Glu Leu
35 40 45

Ile Lys Val Phe Asp Ile Lys Glu Ser Ile Tyr Arg Thr Tyr Phe Leu
50 55 60

Lys Val His Leu Ile Phe Ile Lys Leu Leu Met Lys Ile Pro Pro Val

37847-517 Sequence Listing_ST25 (003)

65		70		75		80
Arg	Lys	Leu	Ser	Cys	Lys	Asp
			85			90
						95
Phe	Asp	Pro	Pro	Tyr	Ala	Arg
			100			105
						110
Lys	Ile	Val	Asn	Tyr	Val	Ile
		115				120
						125
Arg	Asn	Ile	Pro	Gln	Met	Thr
						135
						140
Ile	Asn	Ser	Val	Ser	Ile	His
						150
						155
						160
Pro	Ala	Tyr	Gln	Gly	Ile	Cys
						165
						170
						175
Ser	Phe	Met	Lys	Glu	His	Phe
						180
						185
						190
Asn	Asp	Ser	Ile	Trp	Ala	Lys
						195
						200
						205
Met	Glu	Ile	Ile	Val	Thr	Pro
						210
						215
						220
Leu	Met	Ser	Arg	Cys	Arg	Asn
						225
						230
						235
						240
Trp	Trp	Ala	Ala	Val	Leu	Asn
						245
						250
						255
Pro	Thr	Ile	Phe	Lys	Lys	Asp
						260
						265
						270
Trp	Val	Lys	Ile	Ser	Asn	Ile

37847-517 Sequence Listing_ST25 (003)

275

<210> 118
 <211> 293
 <212> PRT
 <213> Clostridium sp. CAG:510

<400> 118

Met Ile Met Leu Gln Met Thr Gly Gly Met Gly Asn Gln Met Phe Thr
 1 5 10 15

Tyr Ala Leu Tyr Arg Ser Leu Arg Gln Lys Gly Lys Glu Val Cys Ile
 20 25 30

Glu Asp Phe Thr His Tyr Asp Thr Pro Glu Lys Asn Cys Leu Gln Thr
 35 40 45

Val Phe His Leu Asp Tyr Arg Lys Ala Asp Arg Glu Val Tyr Gln Arg
 50 55 60

Leu Thr Asp Ser Glu Pro Asp Phe Leu His Lys Val Lys Arg Lys Leu
 65 70 75 80

Thr Gly Arg Lys Glu Lys Ile Tyr Gln Glu Lys Asp Ala Ile Ile Phe
 85 90 95

Glu Pro Glu Val Phe Gln Thr Asp Asp Val Tyr Met Ile Gly Tyr Phe
 100 105 110

Gln Ser Gly Arg Tyr Phe Glu Lys Ala Val Phe Asp Leu Arg Lys Asp
 115 120 125

Phe Thr Phe Ala Trp Asn Thr Phe Pro Glu Lys Ala Lys Lys Leu Arg
 130 135 140

Glu Gln Met Gln Ala Glu Ser Ser Val Ser Leu His Ile Arg Arg Gly
 145 150 155 160

Asp Tyr Met Asn Gly Lys Phe Ala Ser Ile Tyr Gly Asn Ile Cys Thr
 165 170 175

37847-517 Sequence Listing_ST25 (003)

Asp Ala Tyr Tyr Glu Ala Ala Arg Arg Tyr Met Lys Glu His Phe Gly
180 185 190

Asp Cys Arg Phe Tyr Leu Phe Thr Asp Asp Ala Glu Trp Gly Arg Gln
195 200 205

Gln Glu Ser Glu Asp Thr Val Tyr Val Asp Ala Ser Glu Gly Ala Gly
210 215 220

Ala Tyr Val Asp Met Ala Leu Met Ser Cys Cys Arg His His Ile Ile
225 230 235 240

Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala Trp Leu Asp Glu Asn Pro
245 250 255

Asp Lys Thr Val Ile Ala Pro Ala Lys Trp Leu Asn Ile Ser Glu Gly
260 265 270

Lys Asp Ile Tyr Ala Gly Leu Cys Asn Cys Leu Ile Asp Ala Asn Gly
275 280 285

Ser Val Gln Gly Glu
290

<210> 119

<211> 311

<212> PRT

<213> Rhodopirellula europaea

<400> 119

Met Ile Val Thr Arg Leu Ile Gly Gly Leu Gly Asn Gln Leu Phe Gln
1 5 10 15

Tyr Ala Phe Gly His Ser Leu Ala Arg Ser Thr Tyr Gln Thr Leu Leu
20 25 30

Ile Asp Asp Ser Ala Phe Ile Asp Tyr Arg Leu His Pro Leu Ala Ile
35 40 45

37847-517 Sequence Listing_ST25 (003)

Asp His Phe Thr Ile Ser Ala Ser Arg Leu Ser Asp Ala Asp Arg Ser
50 55 60

Arg Val Pro Gly Lys Phe Leu Arg Thr Pro Val Gly Arg Ala Leu Asp
65 70 75 80

Lys Val Ser Arg Phe Val Pro Gly Tyr Gln Gly Val Leu Pro Val Arg
85 90 95

Arg Glu Lys Pro Phe Gly Phe Arg Glu Ser Leu Leu Ala Arg Glu Ser
100 105 110

Asp Leu Tyr Leu Asp Gly Tyr Trp Gln Ser Glu Lys Phe Phe Pro Gly
115 120 125

Leu Arg Gly Ser Leu Arg Glu Glu Phe Gln Leu Arg Glu Gln Pro Ser
130 135 140

Glu Thr Thr Arg Arg Leu Ser Ala Gln Met Lys Ser Glu Asn Ser Val
145 150 155 160

Ala Ile His Val Arg Arg Gly Asp Tyr Val Thr Ser Ala Lys Ala Lys
165 170 175

Gln Ile Tyr Arg Thr Leu Asp Ala Asp Tyr Tyr Arg Arg Cys Leu Leu
180 185 190

Asp Leu Ala Ala His Glu Thr Asp Leu Lys Leu Tyr Leu Phe Ser Asn
195 200 205

Asp Val Pro Trp Cys Glu Ser Asn Leu Asp Val Gly Ile Pro Phe Thr
210 215 220

Pro Val Gln His Thr Asp Gly Ala Thr Ala His Glu Asp Leu His Leu
225 230 235 240

Ile Ala Gln Cys Arg His Val Val Ile Ala Asn Ser Thr Phe Ser Trp
245 250 255

37847-517 Sequence Listing_ST25 (003)

Trp Gly Ala Tyr Leu Gly Gln Leu His Pro Thr Arg Arg Val Tyr Tyr
260 265 270

Pro Glu Pro Trp Phe His Pro Gly Thr Leu Asp Gly Ser Ala Met Gly
275 280 285

Cys Asp Asp Trp Ile Ser Glu Ala Ser Leu Glu Glu Gln Ser Ser Leu
290 295 300

Lys Ser Ser Arg Arg Ala Ala
305 310

<210> 120

<211> 297

<212> PRT

<213> Unknown

<220>

<223> uncultured bacterium

<400> 120

Met Ile Ile Val Lys Leu Lys Gly Gly Met Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Ile Gly Arg Asn Leu Ala Thr Lys Leu Gly Thr Gln Leu Arg
20 25 30

Leu Asp Leu Thr Phe Leu Leu Asp Arg Ser Pro Arg Lys Asp Phe Val
35 40 45

Phe Arg Asp Tyr Asp Leu Asp Ile Phe Ala Leu Asp Val Ala Phe Ala
50 55 60

Gly Pro Thr Asp Leu Lys Pro Phe Thr Gln Phe Arg Ile Ser His Leu
65 70 75 80

Thr Lys Ile Tyr Asn Ile Phe Pro Arg Leu Leu Gly Arg Pro Tyr Val
85 90 95

Ile Ser Glu Pro His Phe His Phe Ser Glu Ala Ile Leu Lys Ser Ser

37847-517 Sequence Listing_ST25 (003)

100

105

110

Asp Asn Val Tyr Leu Asp Gly Tyr Trp Gln Ser Glu Lys Tyr Phe Lys
115 120 125

Glu Ile Glu Asn Ser Ile Arg Asp Asp Phe Lys Phe Arg Gln Pro Leu
130 135 140

Glu Gly Arg Ala Ala Glu Met Ala Ala Gln Ile Lys Asn Glu Asp Arg
145 150 155 160

Ala Val Cys Leu Asn Val Arg Arg Ala Asp Phe Val Thr Ser Lys Lys
165 170 175

Ala Gln Glu Phe His Gly Phe Ile Gly Leu Asp Tyr Tyr Gln Lys Ala
180 185 190

Val Asp Leu Leu Val Ser Lys Val Gly Pro Leu His Leu Phe Ile Phe
195 200 205

Ser Asp Asp Val Asp Trp Cys Ala Ala Asn Leu Lys Phe Asn Tyr Pro
210 215 220

Thr Thr Phe Val Thr Lys Asp Tyr Ser Gly Lys Lys Tyr Glu Ala Tyr
225 230 235 240

Leu Gln Leu Met Thr Leu Cys Arg His Tyr Ile Ile Pro Asn Ser Thr
245 250 255

Phe Ala Trp Trp Gly Ala Trp Leu Asn Ser Asp Pro Asn Lys Ile Val
260 265 270

Ile Ala Pro Lys Gln Trp Phe Lys Glu Ala Ser Ile Asp Thr Thr Asp
275 280 285

Ile Ile Pro Ser Thr Trp Ile Arg Leu
290 295

<210> 121

37847-517 Sequence Listing_ST25 (003)

<211> 292
 <212> PRT
 <213> Bacillus cereus

<400> 121

Met Ile Ile Val Lys Leu Lys Gly Gly Leu Gly Asn Gln Met Phe Gln
 1 5 10 15

Tyr Ala Leu Gly Lys Ser Leu Ala Leu Tyr Tyr Asp Lys Pro Leu Lys
 20 25 30

Ile Asp Ala Asp Tyr Ile Lys Asn Asn Glu Gly Tyr Val Pro Arg Asp
 35 40 45

Phe Ser Leu Ser Lys Phe Asn Ile Glu Leu Asp Leu Tyr Gln Glu Ala
 50 55 60

Asp Lys Glu Arg Val Gly Phe Ile Leu Lys Asn Asn Phe Leu Ala Lys
 65 70 75 80

Lys Leu Arg Asn Tyr Phe Leu Lys Lys Gly Lys Tyr Lys Gly Lys Tyr
 85 90 95

Ile Ile Glu Asn Pro Asp Asn Leu Gly Leu Phe Lys Lys Glu Leu Phe
 100 105 110

Glu Asn His Asn Glu Ser Met Tyr Ile Asp Gly Tyr Trp Gln Ser Tyr
 115 120 125

Leu Tyr Phe Asn Asn Ile Arg Glu Cys Leu Ile Lys Glu Phe Asn Leu
 130 135 140

Lys Pro Glu Tyr Thr Lys Glu Met Thr Glu Ile Met Gln Arg Ile Asn
 145 150 155 160

Glu Thr Asn Ser Val Ala Val His Ile Arg Arg Gly Asp Tyr Val Lys
 165 170 175

Leu Gly Trp Thr Leu Asp Thr Thr Tyr Tyr Lys Lys Ala Ile Ala Glu
 180 185 190

37847-517 Sequence Listing_ST25 (003)

Ile Val Lys Asn Val Asp Asn Pro Lys Phe Tyr Val Phe Ser Asp Asp
195 200 205

Thr Asp Trp Val Arg Ser Asn Leu Gln Glu Leu Asp Asn Ala Val Phe
210 215 220

Ile Gly Glu Cys Asn Leu Phe Asp Tyr Gln Glu Leu Trp Leu Met Ser
225 230 235 240

Thr Cys Lys His Asn Ile Ile Ser Asn Ser Thr Phe Ser Trp Trp Gly
245 250 255

Ala Trp Leu Asn Gln Asn Asp His Gln Val Val Val Ser Pro Ser Ala
260 265 270

Trp Ile Asn Gly Met Ser Val Glu Thr Thr Ser Leu Ile Pro Asp Ser
275 280 285

Trp Lys Arg Val
290

<210> 122
<211> 302
<212> PRT
<213> Firmicutes bacterium CAG:95

<400> 122

Met Asp Ile Ile Arg Met Glu Gly Gly Leu Gly Asn Gln Leu Phe Gln
1 5 10 15

Tyr Ala Leu Tyr Arg Gln Leu Gln Phe Met Gly Arg Thr Val Lys Met
20 25 30

Asp Val Thr Thr Glu Tyr Gly Arg Glu His Asp Arg Gln Gln Met Leu
35 40 45

Trp Ala Phe Asp Val His Tyr Glu Glu Ala Thr Gln Glu Glu Ile Asn
50 55 60

37847-517 Sequence Listing_ST25 (003)

Arg Leu Thr Asp Gly Phe Met Asp Leu Pro Ser Arg Ile Arg Arg Lys
65 70 75 80

Leu Thr Gly Arg Arg Thr Lys Lys Tyr Ala Glu Ala Asp Ser Asn Phe
85 90 95

Asp Pro Gln Val Leu Leu Lys Thr Pro Val Tyr Leu Thr Gly Tyr Phe
100 105 110

Gln Ser Glu Lys Tyr Phe Lys Asp Val Glu Gly Ile Leu His Thr Glu
115 120 125

Leu Gly Phe Ser Asp Arg Ile Tyr Asp Gly Ile Ser Glu Val Phe Ala
130 135 140

Asp Gln Ile Arg Asn Tyr Gln Lys Gln Ile Arg Glu Thr Glu Ser Val
145 150 155 160

Ser Leu His Val Arg Arg Gly Asp Tyr Leu Glu His Pro Glu Ile Tyr
165 170 175

Gly Met Ser Cys Thr Met Glu Tyr Tyr Gln Ala Gly Val Arg Tyr Ile
180 185 190

Arg Glu Arg His Pro Asp Ala Glu Ile Phe Val Phe Thr Asn Asp Pro
195 200 205

Val Phe Thr Glu Lys Trp Leu Gln Glu Asn Phe Leu Gly Asp Phe Thr
210 215 220

Leu Ile Gln Gly Thr Ser Glu Glu Thr Gly Tyr Leu Asp Leu Met Leu
225 230 235 240

Met Ser Gln Cys Lys His Gln Ile Met Ala Asn Ser Ser Phe Ser Trp
245 250 255

Trp Gly Ala Trp Leu Asn Pro Asn Lys Asp Lys Ile Val Val Ala Pro
260 265 270

37847-517 Sequence Listing_ST25 (003)

Glu Pro Trp Phe Gly Asp Arg Asn Phe His Asp Ile Tyr Thr Glu Glu
275 280 285

Met Ile Arg Ile Ser Pro Arg Gly Glu Val Lys Lys His Gly
290 295 300

<210> 123
<211> 296
<212> PRT
<213> Prevotella oris

<400> 123

Met Ile Ala Ala Thr Leu Phe Gly Gly Leu Gly Asn Gln Met Phe Ile
1 5 10 15

Tyr Ala Thr Val Lys Ala Leu Ser Leu His Tyr Gln Val Pro Met Ala
20 25 30

Phe Asn Leu Asn His Gly Phe Ala Asn Asp Tyr Lys Tyr His Arg Lys
35 40 45

Leu Glu Leu Cys Lys Phe Asn Cys Gln Leu Pro Thr Ala Lys Trp Ile
50 55 60

Thr Phe Asp Tyr Arg Gly Glu Leu Asn Ile Lys Arg Ile Ser Arg Arg
65 70 75 80

Ile Gly Arg Asn Leu Leu Cys Pro Asn Tyr Gln Phe Val Ile Glu Glu
85 90 95

Glu Pro Phe His Tyr Glu Lys Arg Leu Phe Glu Phe Thr Asn Lys Asn
100 105 110

Ile Phe Leu Glu Gly Tyr Trp Gln Ser Pro Cys Tyr Phe Glu Asn Tyr
115 120 125

Ser Lys Glu Ile Arg Ala Asp Phe Gln Leu Lys Val Pro Leu Ser Lys
130 135 140

37847-517 Sequence Listing_ST25 (003)

Glu Met Leu Glu Glu Ile Tyr Ala Leu Lys Ala Thr Gly Lys Thr Leu
145 150 155 160

Val Met Leu Gly Ile Arg Arg Tyr Gln Glu Val Glu Gly Arg Asp Ile
165 170 175

Cys Thr Tyr Lys Leu Cys Asp Lys Glu Tyr Tyr Ile Lys Ala Ile Thr
180 185 190

Tyr Ile Gln Glu Arg Ile Pro Asn Ala Leu Phe Val Val Phe Thr Gln
195 200 205

Asp Lys Glu Trp Ala Thr Thr His Leu Pro Lys Gly Ala Glu Phe Tyr
210 215 220

Phe Val Lys Asp Lys Gln Asp Glu Tyr Ala Thr Val Ala Asp Met Phe
225 230 235 240

Leu Met Thr Gln Cys Thr His Ala Ile Ile Ser Asn Ser Thr Phe Tyr
245 250 255

Trp Trp Gly Ala Trp Leu Gln Cys Thr Thr Lys Asn His Ile Val Ile
260 265 270

Ala Pro Asp Ser Phe Ile Asn Ser Asp Cys Val Cys Lys Glu Trp Ile
275 280 285

Ile Leu Lys Arg Asn Ser Leu Cys
290 295

<210> 124
<211> 302
<212> PRT
<213> Escherichia coli

<400> 124

Met Tyr Ser Cys Leu Ser Gly Gly Leu Gly Asn Gln Met Phe Gln Tyr
1 5 10 15

Ala Ala Ala Tyr Ile Leu Gln Arg Lys Leu Lys Gln Arg Ser Leu Val

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

20

25

30

Leu Asp Asp Ser Tyr Phe Leu Asp Cys Ser Asn Arg Asp Thr Arg Arg
35 40 45

Arg Phe Glu Leu Asn Gln Phe Asn Ile Cys Tyr Asp Arg Leu Thr Thr
50 55 60

Ser Lys Glu Lys Lys Glu Ile Ser Ile Ile Arg His Val Asn Arg Tyr
65 70 75 80

Arg Leu Pro Leu Phe Val Thr Asn Ser Ile Phe Gly Val Leu Leu Lys
85 90 95

Lys Asn Tyr Leu Pro Glu Ala Lys Phe Tyr Glu Phe Leu Asn Asn Cys
100 105 110

Lys Leu Gln Val Lys Asn Gly Tyr Cys Leu Phe Ser Tyr Phe Gln Asp
115 120 125

Ala Thr Leu Ile Asp Ser His Arg Asp Met Ile Leu Pro Leu Phe Gln
130 135 140

Ile Asn Glu Asp Leu Leu His Leu Cys Asn Asp Leu His Ile Tyr Lys
145 150 155 160

Lys Val Ile Cys Glu Asn Ala Asn Thr Thr Ser Leu His Ile Arg Arg
165 170 175

Gly Asp Tyr Ile Thr Asn Pro His Ala Ser Lys Phe His Gly Val Leu
180 185 190

Pro Met Asp Tyr Tyr Glu Lys Ala Ile Arg Tyr Ile Glu Asp Val Gln
195 200 205

Gly Glu Gln Val Ile Ile Val Phe Ser Asp Asp Val Lys Trp Ala Glu
210 215 220

Asn Thr Phe Ala Asn Gln Pro Asn Tyr Tyr Val Val Asn Asn Ser Glu

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

225 230 235 240

Cys Glu Tyr Ser Ala Ile Asp Met Phe Leu Met Ser Lys Cys Lys Asn
245 250 255

Asn Ile Ile Ala Asn Ser Thr Tyr Ser Trp Trp Gly Ala Trp Leu Asn
260 265 270

Thr Phe Glu Asp Lys Ile Val Val Ser Pro Arg Lys Trp Phe Ala Gly
275 280 285

Asn Asn Lys Ser Lys Leu Thr Met Asp Ser Trp Ile Asn Leu
290 295 300

<210> 125
<211> 2102
<212> PRT
<213> Leeia oryzae

<400> 125

Met Ile Ile Val Lys Ile Ile Gly Gly Leu Gly Asn Gln Met Met Gln
1 5 10 15

Tyr Ala Phe Ala His Ala Cys Ala Lys Arg Leu Gly Val Pro Phe Lys
20 25 30

Leu Asp Ile Thr Ala Phe Glu Ser Tyr Lys Leu Trp Pro Tyr Gly Leu
35 40 45

His Asn Phe Glu Ile Thr Ala Pro Ile Ala Ser Leu Glu Glu Ile Glu
50 55 60

His Ala Lys Ser Met Gly Val Ile Thr Glu Thr Ser Phe Arg Phe Asp
65 70 75 80

Asp Ser Leu Val Ser Ala Val Lys Asp Gly Met Tyr Leu Asp Gly Tyr
85 90 95

Trp Ala Asp Tyr Arg Tyr Ser Glu Ser Val Trp Gly Glu Leu Lys Pro
100 105 110

37847-517 Sequence Listing_ST25 (003)

Val Phe Thr Leu Met Asp Pro Leu Thr Pro Glu Gln Gln Ala Leu Ala
115 120 125

Met Asn Leu Ser Ala Pro Asn Ala Val Ala Leu His Val Arg Arg Gly
130 135 140

Asp Tyr Val Thr Asn Pro Asn Cys Phe Leu Leu Pro Gln Gln Tyr Tyr
145 150 155 160

Arg Asp Ala Ile Lys Leu Val Leu Asp Gln Gln Pro Asp Ala Val Phe
165 170 175

Tyr Cys Phe Ser Asp Asp Pro Asp Trp Val Glu Ala His Leu Asp Ile
180 185 190

Pro Ala Pro Lys Val Val Val Arg Gly Gln Gly Ile Asp Asn Gly Phe
195 200 205

Val Asp Met Ile Leu Met Ser Lys Ala Arg His Arg Ile Val Ala Asn
210 215 220

Ser Thr Phe Ser Ile Trp Ala Ser Arg Leu Ala Asp Gln Asp Gly Leu
225 230 235 240

Thr Ile Val Pro Ser Gln Phe Phe Arg Lys Asp Asp Pro Trp Leu Leu
245 250 255

Gln Val Tyr Gly Glu Val Leu Gln Pro Cys Tyr Pro Pro Gln Trp Arg
260 265 270

Val Val Asp Val Thr Gly Asp Gly Lys Lys Glu Ala Glu Asn Thr Ser
275 280 285

Thr Ala Leu Leu Gln Ile Ala Gly Gly Asp Val Arg Gly Arg Lys Leu
290 295 300

Arg Ile Gly Val Trp Gly Phe Tyr Glu Glu Phe Tyr Gln Asn Asn Tyr
305 310 315 320

37847-517 Sequence Listing_ST25 (003)

Ile Phe Leu Asn Lys Asn Ala Pro Ile Gly His Glu Leu Leu Lys Pro
 325 330 335

Phe Asn Gln Leu Tyr Gln Tyr Gly Gln Ala His Asn Leu Glu Phe Val
 340 345 350

Thr Leu Asp Leu Val Ala Asp Leu Ser Thr Leu Asp Ala Val Leu Phe
 355 360 365

Phe Asp Ala Pro Asn Met Arg Ser Pro Leu Val Ser Ser Val Met Gln
 370 375 380

Leu Asp Ile Lys Lys Tyr Leu Cys Leu Leu Glu Cys Glu Leu Ile Lys
 385 390 395 400

Pro Asp Asn Trp Gln Gln Ser Leu His Glu Leu Phe Thr Arg Ile Phe
 405 410 415

Thr Trp His Asp Gly Leu Val Asp Asn His Arg Tyr Ile Lys Val Asn
 420 425 430

Tyr Val Thr Asp Leu Met Pro Trp Ile Glu Ser Ala Gln Ser Leu Thr
 435 440 445

Ala Pro Phe Glu Glu Thr Ala Arg Lys Gly Tyr Leu Gln Lys Lys Leu
 450 455 460

Ile Cys Asn Ile Ser Gly Asn Lys Leu Val Ser His Pro Phe Glu Leu
 465 470 475 480

Tyr Ser Lys Arg Ile Glu Val Ile Arg Trp Phe Glu Ser His His Pro
 485 490 495

Glu His Phe Asp Leu Tyr Gly Met Gly Trp Ser Ala Ser Asp Tyr Pro
 500 505 510

Ser Tyr Lys Gly Lys Ile Asp Asp Lys Leu Glu Val Leu Lys Gly Tyr
 515 520 525

37847-517 Sequence Listing_ST25 (003)

Arg Phe Ser Leu Cys Tyr Glu Asn Ala Lys Glu Leu Pro Gly Tyr Ile
530 535 540

Thr Glu Lys Ile Ile Asp Cys Phe Lys Ala Gly Val Val Pro Val Tyr
545 550 555 560

Ser Gly Ala Pro Asn Ile Ala Asp Trp Ile Pro Asp Asn Cys Phe Ile
565 570 575

Asp Ser Gly Lys Phe Pro Asp Thr Asp Ala Leu Tyr Thr Tyr Leu Ile
580 585 590

Ser Met Thr Glu Glu Val His Ala Asp Tyr Leu Glu Asn Ile Arg Gln
595 600 605

Phe Phe Leu Gly Gly Lys Ala Tyr Pro Phe Ser Ala Asp Ala Phe Ile
610 615 620

Asn Thr Ile Thr Arg Thr Ile Val Gln Asp Cys Leu Phe Pro His Glu
625 630 635 640

Arg Thr Asp Val Ser Val Val Val Pro Asn Tyr Asn His Gly Asn Phe
645 650 655

Val Val Ser Ala Ile Thr Ser Ala Leu Asn Gln Asn Val Ser Val Glu
660 665 670

Leu Leu Val Leu Asp Asn Ala Ser Thr Asp Asp Ser Trp Ser Gln Leu
675 680 685

Gln Phe Phe Ala Asp Tyr Pro Gln Val Arg Leu Ile Arg Asn Arg Trp
690 695 700

Asn Ile Gly Val Gln His Asn Trp Asn His Ala Thr Trp Leu Ala Thr
705 710 715 720

Gly Arg Tyr Val Val Met Leu Ser Ala Asp Asp Leu Leu Leu Pro Gly
725 730 735

37847-517 Sequence Listing_ST25 (003)

His Leu Glu Gln Ala Val Lys Arg Leu Asp Glu Asn Pro Ala Ser Ser
740 745 750

Leu Tyr Tyr Thr Pro Cys Leu Trp Ile Asn Glu His Asp Gln Pro Leu
755 760 765

Gly Thr Leu Asn His Pro Gly His Leu Glu Ser Asp Tyr Val Gly Gly
770 775 780

Arg Asp Glu Ile Ser Asp Leu Leu Lys Phe Asp Ser Tyr Ile Thr Pro
785 790 795 800

Ser Ala Ala Val Ile Arg Arg Glu Thr Leu Asn Arg Ile Gly Ser Met
805 810 815

Asn Leu His Leu Lys Gly Ala Ile Asp Trp Asp Leu Trp Ile Arg Ile
820 825 830

Ala Glu Ile Ser Pro Ala Phe Ile Phe Arg Lys Gln Pro Gly Val Cys
835 840 845

Tyr Arg Gln His Ser Gly Asn Asn Ser Val Asp Phe Tyr Ala Ser Thr
850 855 860

Ala Pro Leu Glu Asp His Ile Arg Ile Val Glu Ser Ile Ile Asp Arg
865 870 875 880

Lys Val Ala Val Lys Tyr Leu Leu Lys Ala Lys Glu Glu Ile Ile Ala
885 890 895

His Leu Asp Asn Arg Ala Ser Ser Tyr Pro Glu Asn Gln Ile Gln His
900 905 910

Leu Leu Ser Arg Ile Asn Asn Ile Lys Asp Tyr Leu Arg Lys Gly Ala
915 920 925

Gly Pro Val Ile Ser Val Ile Ile Pro Thr Lys Asn Arg Pro Gly Leu
930 935 940

37847-517 Sequence Listing_ST25 (003)

Leu Ala Asn Ala Leu Glu Ser Leu Thr Tyr Gln Thr Phe Lys Asp Phe
 945 950 955 960
 Glu Val Val Ile His Asn Asp Gly Gly Cys Asp Ile Gly Gly Ile Val
 965 970 975
 Asp Phe Phe Ser Asp Gln Leu Gln Ile Ser Tyr Val Arg Ser Ser Gln
 980 985 990
 Ser Gly Gly Ala Ala Ala Ser Arg Asn Arg Ala Leu Lys Leu Ala Lys
 995 1000 1005
 Gly Arg Ile Ile Ala Tyr Leu Asp Asp Asp Asp Val Tyr Leu Asp
 1010 1015 1020
 Ser His Leu Glu Lys Leu Val Asp Ala Tyr Lys Gly Arg Ser Glu
 1025 1030 1035
 Lys Phe Ile Tyr Thr Asn Cys Glu Tyr Leu Ile Gln Glu Arg Lys
 1040 1045 1050
 Glu Gly Arg Leu Ile Glu Leu Gly Arg Glu Arg Arg Tyr Ala Gly
 1055 1060 1065
 Ile Ser Tyr Ser Arg Ala Gln Leu Leu Val Ser Asn Phe Ile Pro
 1070 1075 1080
 Thr Pro Thr Trp Ser His Thr Lys Glu Leu Ile Asp Thr Ile Gly
 1085 1090 1095
 Asp Phe Asp Glu Ser Leu Glu Ile Leu Glu Asp Trp Asp Phe Leu
 1100 1105 1110
 Leu Arg Ala Ser Lys Val Thr Glu Phe Tyr Gln Val Asn Ala Thr
 1115 1120 1125
 Thr Val Glu Val Arg Ser Asp Arg Ser Arg Asp Asp His Thr Leu
 1130 1135 1140

37847-517 Sequence Listing_ST25 (003)

Arg	Ala	Asn	Ala	Asp	Lys	Leu	Leu	Ala	Tyr	His	Gln	Lys	Ile	Tyr
	1145					1150					1155			
Ala	Lys	His	Pro	Val	Glu	Asn	Glu	Ser	Ile	Leu	Ala	Asn	Arg	Gln
	1160					1165					1170			
Ser	Leu	Ile	Asn	Ser	Leu	Ser	Asn	Arg	Gln	Asp	Val	Thr	Pro	Lys
	1175					1180					1185			
Asn	Glu	Asn	Ser	Tyr	Gln	Gly	Trp	Val	Asn	Ala	Arg	Gln	Pro	Asn
	1190					1195					1200			
Glu	Leu	Ala	Val	Gln	Ile	Leu	Ala	Glu	Arg	Met	Met	Leu	Gln	Trp
	1205					1210					1215			
Ser	Lys	Gln	Tyr	Gln	Phe	Met	Ile	Val	Met	Val	Val	Lys	Gln	Ser
	1220					1225					1230			
Gln	Gln	Asn	Leu	Leu	Ala	Asn	Thr	Ile	Asp	Ser	Phe	Cys	Gln	Gln
	1235					1240					1245			
Leu	Tyr	Ser	Gly	Trp	Lys	Leu	Ile	Val	Ile	Ser	Asp	Phe	Glu	Ala
	1250					1255					1260			
Pro	Asp	Glu	Ser	Phe	Ile	Asn	Asn	Glu	Val	Leu	Gly	Trp	Leu	Thr
	1265					1270					1275			
Leu	Glu	Thr	Val	Glu	Asp	Glu	Asn	Leu	Leu	Thr	Gln	Ala	Phe	Asn
	1280					1285					1290			
Gly	Val	Leu	Ala	Glu	Val	Pro	Ser	Asp	Trp	Val	Thr	Ile	Leu	Pro
	1295					1300					1305			
Val	Gly	Thr	Arg	Leu	Thr	Ser	Thr	Ala	Leu	Leu	Lys	Val	Gly	Asp
	1310					1315					1320			
Arg	Leu	Leu	Leu	Asn	Gly	Gly	Ala	Cys	Val	Ile	Tyr	Thr	Asp	His
	1325					1330					1335			

37847-517 Sequence Listing_ST25 (003)

Asp	Tyr	Val	Ser	Asp	Asp	Gly	Met	Ile	Lys	Asp	Pro	Val	Leu	Lys
1340						1345					1350			
Pro	Ala	Phe	Asn	Leu	Asp	Met	Leu	Arg	Ser	Gln	Asp	Tyr	Ile	Gly
1355						1360					1365			
Ser	Ser	Ile	Phe	Phe	Arg	Thr	Asp	Ser	Leu	Ala	Ala	Val	Gly	Gly
1370						1375					1380			
Phe	Ala	Ser	Phe	Pro	Gly	Ala	Arg	Thr	Tyr	Glu	Ala	Cys	Phe	Arg
1385						1390					1395			
Met	Leu	Asp	Asn	Tyr	Gly	Pro	Gln	Thr	Ile	Glu	His	Leu	Pro	Glu
1400						1405					1410			
Pro	Val	Met	Thr	Phe	Pro	Glu	Asn	Gln	Pro	Glu	Asn	Ser	Leu	Arg
1415						1420					1425			
Val	Ala	Ala	Met	Gln	Leu	Ala	Leu	Glu	Glu	His	Leu	His	Arg	Asn
1430						1435					1440			
Asn	Ile	Ser	Ala	Ser	Ile	Glu	Glu	Gly	Tyr	Val	Thr	Gly	Thr	Phe
1445						1450					1455			
Leu	Val	Gln	Tyr	His	His	Ser	Glu	Gln	Pro	Phe	Val	Ser	Ile	Ile
1460						1465					1470			
Ile	Pro	Asn	Lys	Asp	Lys	His	Glu	Phe	Leu	Ala	Pro	Cys	Ile	Glu
1475						1480					1485			
Thr	Leu	Met	Lys	Val	Thr	Gln	Tyr	Pro	Ala	Phe	Glu	Val	Ile	Ile
1490						1495					1500			
Val	Asp	Asn	Gln	Ser	Thr	Asp	Pro	Asp	Thr	Leu	Ser	Tyr	Tyr	Glu
1505						1510					1515			
Glu	Ile	Glu	Ser	Arg	Phe	Ala	Asn	Asn	Val	Lys	Val	Ile	Gln	Tyr
1520						1525					1530			

37847-517 Sequence Listing_ST25 (003)

Asp	Asn	Pro	Phe	Asn	Phe	Ser	Ala	Gln	Cys	Asn	Leu	Gly	Ala	Glu
1535						1540					1545			
Ser	Ala	Arg	Gly	Asp	Phe	Ile	Leu	Phe	Leu	Asn	Asn	Asp	Thr	Glu
1550						1555					1560			
Ile	Val	Gln	Ala	Asn	Trp	Leu	Glu	Arg	Met	Met	Gln	His	Ala	Gln
1565						1570					1575			
Arg	Asn	Asp	Val	Gly	Val	Val	Gly	Ala	Arg	Leu	Val	Phe	Pro	Glu
1580						1585					1590			
Thr	Val	Thr	Ile	Gln	His	Ala	Gly	Ile	Val	Leu	Gly	Gly	Lys	Tyr
1595						1600					1605			
Pro	Asp	Glu	Val	Phe	Gln	Phe	Pro	Tyr	Met	Asn	Phe	Pro	Val	Asp
1610						1615					1620			
Lys	Asp	Val	Ser	Leu	Asn	Arg	Thr	Lys	Val	Val	Gln	Asn	Tyr	Ser
1625						1630					1635			
Ala	Val	Thr	Gly	Ala	Cys	Leu	Leu	Val	Arg	Lys	Ser	Leu	Tyr	Gln
1640						1645					1650			
Gln	Val	Gly	Gly	Met	Asn	Glu	Gln	Asn	Leu	Ala	Val	Leu	Tyr	Gly
1655						1660					1665			
Asp	Val	Asp	Leu	Cys	Leu	Arg	Ile	Arg	Gln	Leu	His	Lys	Ser	Val
1670						1675					1680			
Val	Trp	Thr	Pro	Phe	Ser	Thr	Leu	Val	His	His	Thr	Gly	Lys	Thr
1685						1690					1695			
Leu	Asn	Ser	Asn	Ser	Asp	His	Glu	Lys	His	Leu	Met	Met	Val	Ile
1700						1705					1710			
Gln	Thr	Arg	Gln	Glu	Arg	Glu	Tyr	Met	Leu	Ser	His	Trp	Leu	Asp
1715						1720					1725			

37847-517 Sequence Listing_ST25 (003)

Ile	Ile	Ala	Asn	Asp	Pro	Tyr	Tyr	His	Arg	Leu	Leu	Asp	Lys	Ser
1730						1735					1740			
Glu	Cys	Asn	Gly	Thr	Ile	Asp	Cys	Thr	His	Thr	Pro	Leu	Trp	Asp
1745						1750					1755			
Asp	Ile	Pro	Ser	Ala	Arg	Pro	Arg	Leu	Gln	Gly	Met	Ala	Leu	Val
1760						1765					1770			
Gly	Gly	Ser	Gly	Glu	Tyr	Arg	Val	Asn	Met	Pro	Phe	Arg	Thr	Leu
1775						1780					1785			
Glu	Arg	Ser	Ala	Leu	Ala	Glu	Ile	Val	Leu	Ser	Asn	Met	Thr	Ser
1790						1795					1800			
Lys	Ala	Arg	Leu	Pro	Ser	Ile	Thr	Glu	Leu	Ala	Arg	Asn	Ala	Pro
1805						1810					1815			
Asp	Val	Phe	Val	Val	Gln	Asn	Ala	Leu	Ala	Asp	Glu	Phe	Ile	Arg
1820						1825					1830			
Met	Leu	Glu	Met	Tyr	Lys	Lys	Tyr	Leu	Pro	Ser	Val	Phe	Arg	Ile
1835						1840					1845			
Gln	Met	Leu	Asp	Asp	Leu	Leu	Thr	Glu	Ile	Pro	Asp	Ala	Ser	Ser
1850						1855					1860			
Phe	Lys	Arg	His	Phe	Gln	Lys	Asn	Trp	Arg	Asp	Ala	Lys	Ala	Arg
1865						1870					1875			
Leu	Arg	Lys	Ser	Leu	Lys	Phe	Cys	Asp	Arg	Leu	Ile	Val	Ser	Thr
1880						1885					1890			
Glu	Pro	Leu	Arg	Thr	Phe	Ala	Glu	Asp	Met	Ile	Asp	Asp	Ile	Ile
1895						1900					1905			
Val	Val	Pro	Asn	Met	Leu	Glu	Arg	Ser	Val	Trp	Gly	Asp	Leu	Val
1910						1915					1920			

37847-517 Sequence Listing_ST25 (003)

Ser	Lys	Arg	Arg	Ala	Gly	Lys	Lys	Pro	Arg	Val	Gly	Trp	Val	Gly	1925	1930	1935
Ala	Gln	Gln	His	Ala	Gly	Asp	Leu	Ala	Leu	Met	Thr	Asp	Val	Val	1940	1945	1950
Lys	Ala	Thr	Gly	His	Glu	Val	Asp	Trp	Val	Phe	Gln	Gly	Met	Cys	1955	1960	1965
Pro	Asp	Asp	Ile	Arg	Pro	Tyr	Val	Ala	Glu	Val	Asn	Thr	Glu	Trp	1970	1975	1980
Leu	Thr	Tyr	Asp	Lys	Tyr	Pro	Gln	Gly	Ile	Ala	Ala	Leu	Asn	Leu	1985	1990	1995
Asp	Leu	Ala	Ile	Ala	Pro	Leu	Glu	Ile	Asn	Ala	Phe	Asn	Glu	Ala	2000	2005	2010
Lys	Ser	Asn	Leu	Arg	Leu	Leu	Glu	Tyr	Gly	Ala	Leu	Gly	Trp	Pro	2015	2020	2025
Val	Ile	Cys	Thr	Asp	Ile	Tyr	Pro	Tyr	Gln	Thr	Asn	Asn	Ala	Pro	2030	2035	2040
Val	Cys	Arg	Val	Pro	Asn	Asp	Ala	Ser	Ala	Trp	Ile	Glu	Ala	Ile	2045	2050	2055
Arg	Ser	His	Ile	Ala	Asp	Leu	Asp	Ala	Thr	Ala	Gln	Lys	Gly	Asp	2060	2065	2070
Gln	Leu	Arg	Gln	Trp	Val	His	Asp	His	Tyr	Met	Ile	Glu	Asp	His	2075	2080	2085
Ala	Gln	Glu	Trp	Leu	Ser	Ala	Leu	Thr	Arg	Pro	Ala	Gly	Lys		2090	2095	2100

<210> 126

<211> 284

37847-517 Sequence Listing_ST25 (003)

<212> PRT

<213> *Desulfovibrio africanus*

<400> 126

Met Phe Gln Tyr Ala Ala Ala Arg Ala Leu Ser Leu Arg His Ser Ala
 1 5 10 15

Ser Leu Ala Ala Asp Leu Thr Trp Phe Ser Gln Gln Phe Asp Val Gln
 20 25 30

Thr Thr Pro Arg Glu Tyr Ala Leu Pro Ala Phe Arg Leu Asn Leu Pro
 35 40 45

Glu Ala Asp Lys Arg Ile Val Ala Thr Phe Arg Leu Asn Pro Thr Glu
 50 55 60

Leu Arg Ile Val Ser Phe Leu Arg His Arg Ile Cys Phe Pro Ser Arg
 65 70 75 80

Phe Leu Pro Arg His Ile Thr Glu Leu Ser Phe Asp Tyr Trp Asp Gly
 85 90 95

Phe Arg Asp Ile Leu Pro Pro Ala Tyr Leu Asp Gly Tyr Trp Gln Ser
 100 105 110

Glu Arg Tyr Phe Ser Asp Tyr Pro Asp Ile Ile Arg Ala Asp Phe Ser
 115 120 125

Met Leu Ser Ile Ser Glu Gln Ala Ala Trp Met Ser Ala Lys Ile Ala
 130 135 140

Ser Val Gln Asp Ser Ile Ser Leu His Ile Arg Arg Gly Asp Tyr Val
 145 150 155 160

Asn Ser Leu Ala Thr Arg Lys Ala His Gly Ile Asp Thr Glu Arg Tyr
 165 170 175

Tyr Ala Lys Ala Leu Glu Trp Ile Ala Asp Arg Ile Gly Ala Ala Thr
 180 185 190

37847-517 Sequence Listing_ST25 (003)

Ile Phe Ala Phe Ser Asp Asp Pro Arg Trp Val Arg Ala Asn Phe Asp
 195 200 205

Phe Gly Lys His Lys Gly Ile Val Val Asp Gly Ser Trp Thr Ala His
 210 215 220

Glu Asp Met His Leu Met Ser Leu Cys Ser His His Ile Ile Ala Asn
 225 230 235 240

Ser Ser Phe Ser Trp Trp Gly Ala Trp Leu Ser Thr Ser Gln Gly Ile
 245 250 255

Thr Ile Ala Pro Lys Ser Trp Phe Ser Asn Pro His Ile Trp Thr Pro
 260 265 270

Asp Val Cys Pro Ala Thr Trp Glu Arg Ile Pro Cys
 275 280

<210> 127

<211> 298

<212> PRT

<213> Akkermansia muciniphila CAG:154

<400> 127

Met Ala Lys Gly Lys Ile Ile Val Met Arg Leu Phe Gly Gly Leu Gly
 1 5 10 15

Asn Gln Leu Phe Gln Tyr Ala Phe Leu Phe Ala Leu Ser Arg Gln Gly
 20 25 30

Gly Lys Ala Arg Leu Glu Thr Ser Ser Tyr Glu His Asp Asp Lys Arg
 35 40 45

Val Cys Glu Leu His His Phe Arg Val Ser Leu Pro Ile Glu Gly Gly
 50 55 60

Pro Pro Pro Trp Ala Phe Arg Lys Ser Arg Ile Pro Ala Cys Leu Arg
 65 70 75 80

37847-517 Sequence Listing_ST25 (003)

Ser	Leu	Phe	Ala	Ala	Pro	Lys	Tyr	Pro	His	Phe	Arg	Glu	Glu	Lys	Arg	
			85					90						95		
His	Gly	Phe	Asp	Pro	Gly	Leu	Ala	Ala	Pro	Pro	Arg	Arg	His	Thr	Tyr	
			100					105					110			
Phe	Lys	Gly	Tyr	Phe	Gln	Thr	Glu	Gln	Tyr	Phe	Leu	His	Cys	Arg	Glu	
		115					120					125				
Gln	Leu	Cys	Arg	Glu	Phe	Arg	Leu	Lys	Thr	Pro	Leu	Thr	Pro	Glu	Asn	
	130					135					140					
Ala	Arg	Ile	Leu	Glu	Asp	Ile	Arg	Ser	Cys	Cys	Ser	Ile	Ser	Leu	His	
145					150					155					160	
Ile	Arg	Arg	Thr	Asp	Tyr	Leu	Ser	Asn	Pro	Tyr	Leu	Ser	Pro	Pro	Pro	
			165					170						175		
Leu	Glu	Tyr	Tyr	Leu	Arg	Ser	Met	Ala	Glu	Met	Glu	Gly	Arg	Leu	Arg	
		180						185					190			
Ala	Ala	Asp	Ala	Pro	Gln	Glu	Ser	Leu	Arg	Tyr	Phe	Ile	Phe	Ser	Asp	
		195					200					205				
Asp	Ile	Glu	Trp	Ala	Arg	Gln	Asn	Leu	Arg	Pro	Ala	Leu	Pro	His	Val	
	210					215					220					
His	Val	Asp	Ile	Asn	Asp	Gly	Gly	Thr	Gly	Tyr	Phe	Asp	Leu	Glu	Leu	
225					230					235					240	
Met	Arg	Asn	Cys	Arg	His	His	Ile	Ile	Ala	Asn	Ser	Thr	Phe	Ser	Trp	
			245						250					255		
Trp	Ala	Ala	Trp	Leu	Asn	Glu	His	Ala	Glu	Lys	Ile	Val	Ile	Ala	Pro	
		260						265					270			
Arg	Ile	Trp	Phe	Asn	Arg	Glu	Glu	Gly	Asp	Arg	Tyr	His	Thr	Asp	Asp	
	275						280					285				

37847-517 Sequence Listing_ST25 (003)

Ala Leu Ile Pro Gly Ser Trp Leu Arg Ile
290 295

<210> 128

<211> 292

<212> PRT

<213> Dysgonomonas mossii

<400> 128

Met Lys Ile Val Lys Leu Gln Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Ile Ala Arg Thr Leu Glu Thr Asn Lys Lys Lys Asp Ile Phe
20 25 30

Leu Asp Leu Ser Phe Leu Arg Met Asn Asn Val Ser Thr Asp Cys Phe
35 40 45

Thr Ala Arg Asp Phe Glu Leu Ser Ile Phe Pro His Leu Arg Ala Lys
50 55 60

Lys Leu Asn Ser Leu Gln Glu Lys Phe Leu Leu Ser Asp Arg Val Arg
65 70 75 80

Tyr Lys Phe Ile Arg Lys Ile Ala Asn Ile Asn Phe His Lys Ile Asn
85 90 95

Gln Leu Glu Asn Glu Ile Val Gly Ile Pro Phe Gly Ile Lys Asn Val
100 105 110

Tyr Leu Asp Gly Phe Phe Gln Ser Glu Ser Tyr Phe Lys His Ile Arg
115 120 125

Phe Asp Leu Ile Lys Asp Phe Glu Phe Pro Glu Leu Asp Thr Arg Asn
130 135 140

Glu Ala Leu Lys Lys Thr Ile Val Asn Asn Asn Ser Val Ser Ile His
145 150 155 160

Ile Arg Arg Gly Asp Tyr Val His Leu Lys Asn Ala Asn Thr Tyr His

37847-517 Sequence Listing_ST25 (003)

165 170 175

Gly Val Leu Ser Leu Glu Tyr Tyr Leu Asn Cys Ile Lys Arg Ile Gly
180 185 190

Glu Glu Thr Lys Glu Gln Leu Ser Phe Phe Ile Phe Ser Asp Asp Pro
195 200 205

Glu Tyr Ala Ser Lys Ser Leu Ser Phe Leu Pro Asn Met Gln Ile Val
210 215 220

Asp Trp Asn Leu Gly Lys Asn Ser Trp Lys Asp Met Ala Leu Met Leu
225 230 235 240

Ala Cys Lys His His Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Gly
245 250 255

Ala Trp Leu Ser Glu Arg Asn Gly Ile Thr Tyr Ala Pro Val Lys Trp
260 265 270

Phe Asn Asn Glu Ser Gln Tyr Asn Ile Asn Asn Ile Ile Pro Ser Asp
275 280 285

Trp Val Ile Ile
290

<210> 129
<211> 291
<212> PRT
<213> Prevotella oris

<400> 129

Met Asp Ile Val Leu Ile Phe Asn Gly Leu Gly Asn Gln Met Ser Gln
1 5 10 15

Tyr Ala Phe Tyr Met Ser Lys Lys Lys Phe Val Pro Gln Ser Lys Cys
20 25 30

Met Tyr Tyr Lys Gly Ala Ser Asn Asn His Asn Gly Ser Glu Leu Asp
35 40 45

37847-517 Sequence Listing_ST25 (003)

Lys Leu Phe Asp Ile Lys Tyr Ser Glu Thr Phe Phe Cys Lys Leu Ile
50 55 60

Leu Leu Leu Phe Lys Leu Tyr Glu Asn Ile Pro Arg Leu Arg Lys Tyr
65 70 75 80

Phe His Ile Leu Gly Ile Asn Ile Val Ser Glu Pro Gln Asn Tyr Asp
85 90 95

Tyr Asn Glu Ser Ile Leu Lys Lys Lys Thr Arg Phe Gly Ile Thr Leu
100 105 110

Tyr Lys Gly Gly Trp His Ser Glu Lys Tyr Phe Leu Ala Asn Lys Gln
115 120 125

Asp Val Leu Asn Thr Phe Ser Phe Lys Ile Ala Lys Glu Asp Lys Asn
130 135 140

Phe Ile Asp Leu Ala Lys Ser Ile Glu Glu Asp Thr Asn Ser Val Ser
145 150 155 160

Leu His Val Arg Arg Gly Asp Tyr Leu Asn Ile Ser Pro Thr Asp His
165 170 175

Tyr Gln Phe Gly Gly Val Ala Thr Thr Asn Tyr Tyr Lys Asn Ala Val
180 185 190

Ser Tyr Met Leu Lys Arg Asn Lys Gln Ala His Phe Tyr Ile Phe Ser
195 200 205

Asp Asp Ile Thr Trp Cys Lys Ala Glu Tyr Lys Asp Leu Met Pro Thr
210 215 220

Phe Ile Glu Cys Asn Lys Lys Asn Lys Ser Trp Arg Asp Met Leu Leu
225 230 235 240

Met Ser Leu Cys Thr Asn His Ile Asn Ala Asn Ser Thr Phe Ser Trp
245 250 255

37847-517 Sequence Listing_ST25 (003)

Trp Gly Ala Trp Leu Ser Thr Lys Asn Gly Ile Thr Ile Cys Pro Thr
260 265 270

Glu Phe Ile His Asn Val Val Thr Arg Asp Ile Tyr Pro Glu Thr Trp
275 280 285

Val Gln Leu
290

<210> 130

<211> 294

<212> PRT

<213> Pseudogulbenkiania ferrooxidans

<400> 130

Met Ile Ile Val Arg Leu Met Gly Gly Met Gly Asn Gln Leu Phe Gln
1 5 10 15

Tyr Ala Thr Ala Phe Ala Leu Ser Lys Arg Lys Ser Glu Pro Leu Val
20 25 30

Leu Asp Thr Arg Phe Phe Asp His Tyr Thr Leu His Gly Gly Tyr Lys
35 40 45

Leu Asp His Phe Asn Ile Ser Ala Arg Ile Leu Ser Lys Glu Glu Glu
50 55 60

Ser Leu Tyr Pro Asn Trp Gln Ala Asn Leu Leu Leu Arg Tyr Pro Ile
65 70 75 80

Ile Asp Arg Ala Phe Lys Lys Trp His Val Glu Arg Gln Phe Thr Tyr
85 90 95

Gln Asp Arg Ile Tyr Arg Met Lys Arg Gly Gln Ala Leu Leu Gly Tyr
100 105 110

Trp Gln Ser Glu Leu Tyr Phe Gln Glu Tyr Arg Lys Glu Ile Ser Ala
115 120 125

37847-517 Sequence Listing_ST25 (003)

Glu Phe Thr Leu Lys Glu Gln Ser Ser Val Thr Ala Gln Gln Ile Ser
130 135 140

Val Ala Met Gln Gly Gly Asn Ser Val Ala Val His Ile Arg Arg Gly
145 150 155 160

Asp Tyr Leu Ser Asn Pro Ser Ala Leu Arg Thr His Gly Ile Cys Ser
165 170 175

Leu Gly Tyr Tyr Asn His Ala Met Ser Leu Leu Asn Glu Arg Ile Asn
180 185 190

Asp Ala Gln Phe Tyr Ile Phe Ser Asp Asp Ile Ala Trp Ala Lys Glu
195 200 205

Asn Ile Lys Ile Gly Lys Thr Ser Lys Asn Leu Ile Phe Ile Glu Gly
210 215 220

Glu Ser Val Glu Thr Asp Phe Trp Leu Met Thr Gln Ser Lys His His
225 230 235 240

Ile Ile Ala Asn Ser Thr Phe Ser Trp Trp Gly Ala Trp Leu Ala Asn
245 250 255

Asn Thr Asp Glu Gln Leu Val Ile Cys Pro Ser Pro Trp Phe Asp Asp
260 265 270

Lys Asn Leu Ser Glu Thr Asp Leu Ile Pro Lys Ser Trp Ile Arg Leu
275 280 285

Asn Lys Asp Leu Pro Val
290

<210> 131
<211> 298
<212> PRT
<213> Salmonella enterica

<400> 131

37847-517 Sequence Listing_ST25 (003)

Met	Tyr	Ser	Cys	Leu	Ser	Gly	Gly	Leu	Gly	Asn	Gln	Met	Phe	Gln	Tyr	1	5	10	15
Ala	Ala	Ala	Tyr	Ile	Leu	Lys	Gln	Tyr	Phe	Gln	Ser	Thr	Thr	Leu	Val	20	25	30	
Leu	Asp	Asp	Ser	Tyr	Tyr	Tyr	Ser	Gln	Pro	Lys	Arg	Asp	Thr	Val	Arg	35	40	45	
Ser	Leu	Glu	Leu	Asn	Gln	Phe	Asn	Ile	Ser	Tyr	Asp	Arg	Phe	Ser	Phe	50	55	60	
Thr	Asp	Glu	Lys	Glu	Lys	Ile	Lys	Leu	Leu	Arg	Lys	Phe	Lys	Arg	Asn	65	70	75	80
Pro	Phe	Pro	Lys	Lys	Ile	Ser	Glu	Ile	Leu	Ser	Ile	Ala	Leu	Phe	Gly	85	90	95	
Lys	Tyr	Ala	Leu	Ser	Asp	Ser	Ala	Phe	Tyr	Ala	Val	Glu	Thr	Ile	Lys	100	105	110	
Asn	Ile	Asp	Lys	Ala	Cys	Leu	Phe	Ser	Phe	Tyr	Gln	Asp	Ala	Asp	Leu	115	120	125	
Leu	Asn	Lys	His	Lys	Gln	Leu	Ile	Leu	Pro	Leu	Phe	Glu	Leu	Arg	Asp	130	135	140	
Asp	Leu	Leu	Asp	Ile	Cys	Lys	Asn	Leu	Asp	Val	Tyr	Pro	Leu	Ile	Leu	145	150	155	160
Arg	Asn	Asn	Asn	Thr	Thr	Ala	Leu	His	Ile	Arg	Arg	Gly	Asp	Tyr	Leu	165	170	175	
Thr	Asn	Gln	His	Ala	Ala	Lys	Tyr	His	Gly	Val	Leu	Asp	Thr	Ser	Tyr	180	185	190	
Tyr	Asn	Asn	Ala	Met	Glu	Tyr	Val	Glu	Arg	Glu	Arg	Gly	Lys	Gln	Asn	195	200	205	

37847-517 Sequence Listing_ST25 (003)

Phe Ile Ile Phe Ser Asp Asp Val Lys Trp Ala Gln Lys Ala Phe Leu
210 215 220

Gly Asn Glu Asn Cys Tyr Ile Val Asn Asn Gly Asp Tyr Asp Tyr Ser
225 230 235 240

Ala Ile Asp Met Tyr Leu Met Ser Leu Cys Lys Asn Asn Ile Ile Ala
245 250 255

Asn Ser Thr Tyr Ser Trp Trp Gly Ala Trp Leu Asn Lys Ser Glu Asp
260 265 270

Lys Leu Val Ile Ser Pro Lys Gln Trp Phe Leu Gly Asn Asn Glu Thr
275 280 285

Ser Leu Arg Asn Ala Ser Trp Ile Ile Leu
290 295

<210> 132
<211> 287
<212> PRT
<213> Carnobacterium sp. WN1359

<400> 132

Met Leu Ile Val Lys Val Tyr Gly Gly Ile Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ser Phe Tyr Lys Tyr Leu Gln Lys Asn Asn Asp Asp Val Phe Leu
20 25 30

Asp Ile Ser Asp Tyr Lys Val His Asn His His Asn Gly Phe Glu Leu
35 40 45

Ile Asp Val Phe Asn Ile Glu Val Lys Gln Ala Asp Met Ser Lys Phe
50 55 60

Lys Gly His Val Ser Ser Lys Asn Ser Ile Phe Tyr Arg Leu Thr Ser
65 70 75 80

Lys Leu Phe Lys Arg Asn Ile Leu Gly Tyr Ser Glu Phe Met Asp Ser

37847-517 Sequence Listing_ST25 (003)

85 90 95

Asn Gly Ile Ser Ile Val Arg Asn Glu Lys Ile Leu Thr Asp His Tyr
100 105 110

Phe Ile Gly Phe Trp Gln Asp Val Leu Tyr Leu Gln Ser Val Glu Glu
115 120 125

Glu Ile Lys Glu Ala Phe Asn Phe Lys Asn Val Ala Ile Gly Lys Gln
130 135 140

Asn Leu Glu Leu Ile Ser Leu Ser Glu Ser Val Glu Ser Val Ser Val
145 150 155 160

His Ile Arg Lys Gly Asp Tyr Ala Asn Asn Ser Asp Leu Ser Asp Ile
165 170 175

Cys Asp Leu Glu Tyr Tyr Glu Glu Ala Met Lys Ile Ile Asp Ser Lys
180 185 190

Val Ser Glu Pro Leu Tyr Phe Ile Phe Ser Asp Asp Ile Glu Trp Cys
195 200 205

Lys Gln Lys Phe Gly Lys Arg Asp Asn Leu Ile Tyr Val Asp Trp Asn
210 215 220

Ile Ala Lys Lys Ser Tyr Ile Asp Met Leu Leu Met Ser Lys Cys Lys
225 230 235 240

His Asn Ile Ile Ala Asn Ser Thr Phe Ser Trp Trp Gly Ala Trp Leu
245 250 255

Asn Asn Asn Ser Lys Lys Ile Val Ile Cys Pro Lys Thr Trp Asp Arg
260 265 270

Lys Lys Asn Glu Asn His Leu Leu Leu Asn Asp Trp Ile Ala Ile
275 280 285

<210> 133

37847-517 Sequence Listing_ST25 (003)

<211> 291
 <212> PRT
 <213> Prevotella sp. CAG:1185

<400> 133

Met Met Lys Ile Ile Val Asn Met Ala Cys Gly Leu Ala Asn Arg Met
 1 5 10 15

Phe Gln Tyr Ser Tyr Tyr Leu Phe Leu Met His Lys Gly Tyr Asn Val
 20 25 30

Lys Val Asp Phe Tyr Asn Ser Ala Lys Leu Ala His Glu Lys Val Ala
 35 40 45

Trp Asn Asp Ile Phe Pro Lys Ala Arg Ile Glu Gln Ala Ser Phe Ser
 50 55 60

Asp Ile Leu Lys Ser Gly Gly Gly Ser Asp Val Ile Ser Lys Ile Arg
 65 70 75 80

Arg Lys Tyr Leu Pro Phe Leu Ser Ser Val Val Asn Met Pro Thr Ala
 85 90 95

Phe Asp Ala Asn Leu Pro Val Glu Asn Lys Lys Leu Gln Tyr Ile Ile
 100 105 110

Gly Val Phe Gln Asn Ala Asn Met Val Glu Ala Val Glu Glu Asp Val
 115 120 125

Lys Arg Cys Phe Lys Phe Gln Pro Phe Thr Asp Glu Arg Asn Leu Lys
 130 135 140

Leu Gln Asn Glu Met Gln Ser Cys Glu Ser Val Ala Ile His Val Arg
 145 150 155 160

Lys Gly Lys Asp Tyr Ala Gln Arg Ile Trp Tyr Gln Asn Thr Cys Pro
 165 170 175

Ile Glu Tyr Tyr Gln Asn Ala Ile Arg Leu Ile Ser Glu Lys Val Asn
 180 185 190

37847-517 Sequence Listing_ST25 (003)

28 May 2018
2015259088

Asn Pro Lys Leu Tyr Val Phe Thr Asp Asn Pro Glu Trp Val Lys Glu
195 200 205

His Phe Lys Asp Phe Pro Tyr Thr Leu Val Glu Gly Asn Pro Ala Ser
210 215 220

Gly Trp Gly Ser His Phe Asp Met Gln Leu Met Ser Val Cys Lys His
225 230 235 240

Asn Ile Ile Ser Asn Ser Thr Tyr Ser Trp Trp Ser Ala Phe Leu Asn
245 250 255

Val His Asn Glu Lys Ile Val Ile Gly Pro Lys Val Trp Phe Asn Pro
260 265 270

Asp Ser Cys Ser Glu Phe Thr Ser Glu Arg Ile Leu Cys Lys Asp Trp
275 280 285

Ile Ala Val
290

<210> 134
<211> 294
<212> PRT
<213> Selenomonas sp. CM52

<400> 134

Met Phe Gln Tyr Ala Met Ala Ser Ser Val Ala Arg Arg Ala Gly Glu
1 5 10 15

Ile Leu Lys Leu Asp Leu Ser Trp Ile Arg Gln Met Glu Lys Lys Leu
20 25 30

Ser Ala Asp Asp Ile Tyr Gly Leu Gly Ile Phe Ser Phe Asp Glu Lys
35 40 45

Phe Ser Thr Ser Asn Glu Val Gln Lys Phe Leu Pro Ser Gly Lys Phe
50 55 60

37847-517 Sequence Listing_ST25 (003)

Ser Ala Lys Ile Tyr Arg Ala Val Asn Arg Arg Met Pro Phe Ser Trp
65 70 75 80

Arg Arg Val Leu Glu Glu Gly Gly Met Gly Trp His Pro Gln Ile Met
85 90 95

Glu Ile Arg Arg Ser Val Tyr Phe Tyr Met Gly Tyr Trp Gln Ser Glu
100 105 110

Lys Tyr Phe Ser Asp Phe Ile Gln Glu Ile Arg Lys Asp Phe Thr Phe
115 120 125

Arg Glu Glu Val Arg Gln Ser Ile Glu Glu Arg Arg Pro Ile Val Glu
130 135 140

Lys Ile Arg Lys Ser Asp Ala Val Ser Leu His Ile Arg Arg Gly Asp
145 150 155 160

Tyr Ala Gln Asn Pro Ala Leu Gly Glu Ile Phe Leu Ser Phe Thr Met
165 170 175

Gln Tyr Tyr Ile Asp Ala Ala Arg Tyr Ile Ser Glu Arg Val Lys Thr
180 185 190

Pro Val Phe Phe Ile Phe Ser Asp Asp Ile Pro Trp Ala Lys Glu Asn
195 200 205

Leu Pro Leu Pro Tyr Glu Val Cys Tyr Ile Asp Asp Asn Ile Gln Thr
210 215 220

Asn Glu Arg Glu Ile Gly His Lys Ser Lys Gly Tyr Glu Asp Met Tyr
225 230 235 240

Leu Met Thr Gln Cys Gln His Asn Ile Ile Ala Asn Ser Ser Phe Ser
245 250 255

Trp Trp Gly Ala Trp Leu Asn His Asn Pro Asn Lys Ile Val Val Ala
260 265 270

37847-517 Sequence Listing_ST25 (003)

Pro Lys Lys Trp Cys Asn Gly Ser Phe Asn Tyr Ala Asp Ile Val Pro
 275 280 285

Glu Gln Trp Val Lys Leu
 290

<210> 135
 <211> 293
 <212> PRT
 <213> Bacteroides nordii

<400> 135

Met Glu Ile Val Phe Ile Phe Asn Gly Leu Gly Asn Gln Met Ser Gln
 1 5 10 15

Tyr Ala Leu Tyr Leu Ser Lys Arg Asn Leu Gly Cys Lys Val Arg Tyr
 20 25 30

Ala Tyr Asn Ile Arg Ser Leu Ser Asp His Asn Gly Phe Glu Leu Asp
 35 40 45

Arg Val Phe Gly Ile Thr Tyr Pro Asn Asn Leu Phe Asn Lys Cys Ile
 50 55 60

Asn Ile Ile Tyr Arg Leu Leu Phe Ala Asn Lys Tyr Leu Phe Leu Val
 65 70 75 80

Gln Lys Met Ile Tyr Val Leu Arg Gln Met Asn Val Tyr Ser Ile Lys
 85 90 95

Glu Lys Asp Asn Tyr Asp Tyr Asp Tyr Lys Ile Leu Thr Arg His Lys
 100 105 110

Gly Ile Val Leu Tyr Tyr Gly Gly Trp His Ser Glu Lys Tyr Phe Leu
 115 120 125

Ser Asn Ala Asp Ile Ile Lys Asp Lys Phe Arg Phe Asn Ile Ser Lys
 130 135 140

37847-517 Sequence Listing_ST25 (003)

Leu Asn Ser Glu Ser Leu Val Leu Tyr His Arg Leu Ser Ser Leu Asn
145 150 155 160

Ala Val Ala Leu His Val Arg Arg Gly Asp Tyr Met Ala Pro Glu His
165 170 175

Tyr Asn Val Phe Gly Cys Val Cys Gly Ile Glu Tyr Tyr Lys Ala Ala
180 185 190

Ile Gln Tyr Ile Gln Ser Gln Ile Leu Asn Pro Val Phe Ile Val Phe
195 200 205

Ser Asn Asp Ile Glu Trp Val Lys Glu Asn Ile Thr Gly Ile Gln Met
210 215 220

Ile Phe Val Asp Phe Asn Lys Lys Glu Asn Ser Trp Met Asp Met Cys
225 230 235 240

Leu Met Ser Cys Cys Glu His Asn Ile Ile Ser Asn Ser Thr Phe Ser
245 250 255

Trp Trp Gly Ala Trp Leu Asn Asn Asn Lys Asn Lys Ile Val Val Cys
260 265 270

Pro Lys Tyr Phe Met Ser Asn Ile Asp Thr Lys Asp Ile Tyr Pro Glu
275 280 285

Ser Trp Ile Lys Ile
290

<210> 136
<211> 276
<212> PRT
<213> Parabacteroides merdae

<400> 136

Met Lys Lys Lys Asp Ile Ile Leu Arg Val Trp Gly Gly Val Gly Asn
1 5 10 15

Gln Leu Phe Ile Tyr Ala Phe Ala Lys Val Leu Ser Leu Ile Thr Asp

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

20

25

30

Cys Lys Val Thr Leu Asp Ile Arg Thr Gly Phe Ala Asn Asp Gly Tyr
35 40 45

Lys Arg Val Tyr Arg Leu Gly Asp Phe Ser Ile Ser Leu Leu Pro Ala
50 55 60

Leu Arg Phe Tyr Thr Leu Leu Ser Phe Ala Gln Arg Lys Met Pro Tyr
65 70 75 80

Ile Arg His Leu Leu Ala Tyr Lys Phe Asp Phe Phe Glu Glu Asp Gln
85 90 95

Lys Tyr Pro Leu Glu Thr Leu Asp Ser Phe Phe Lys Ile Tyr Ser Asp
100 105 110

Lys Asn Leu Tyr Leu Gln Gly Tyr Trp Gln Tyr Phe Asp Phe Ser Ser
115 120 125

Tyr Arg Asp Val Leu Leu Lys Asp Leu Arg Phe Glu Val Glu Ile Asn
130 135 140

Asn Thr Tyr Leu Tyr Tyr Ser Asp Leu Ile Glu Lys Ser Asn Ala Val
145 150 155 160

Ala Ile His Phe Arg Arg Ile Gln Tyr Glu Pro Val Ile Ser Ile Asp
165 170 175

Tyr Tyr Lys Lys Ala Ile Lys Tyr Ile Ser Glu Asn Val Glu Asn Pro
180 185 190

Thr Phe Phe Ile Phe Ser Asp Asp Ile Asn Trp Cys Arg Glu Asn Leu
195 200 205

Ser Ile Asn Gly Ile Cys Phe Phe Val Glu Asn Phe Lys Asp Glu Leu
210 215 220

Tyr Glu Leu Lys Leu Met Ser Gln Cys Asn His Phe Ile Ile Ala Asn

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

225 230 235 240

Ser Thr Phe Ser Trp Trp Gly Ala Trp Leu Ser Val Asn Ala Asp Lys
245 250 255

Lys Val Ile Met Pro Asp Gly Tyr Thr Asp Val Ser Met Asn Gly Ser
260 265 270

Ile Val His Ile
275

<210> 137

<211> 290

<212> PRT

<213> Butyrivibrio sp. NC2007

<400> 137

Met Ile Ile Ile Gln Leu Lys Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Leu Tyr Lys Glu Leu Lys His Arg Gly Arg Asp Val Lys Ile
20 25 30

Asp Asp Glu Ser Gly Phe Ile Gly Asp Lys Leu Arg Val Pro Val Leu
35 40 45

Asp Arg Phe Gly Val Glu Tyr Asp Arg Ala Thr Lys Asp Glu Val Ile
50 55 60

Ala Leu Thr Asp Ser Lys Met Asp Ile Phe Ser Arg Ile Arg Arg Lys
65 70 75 80

Leu Thr Gly Arg Lys Thr Phe Arg Ile Asp Glu Met Glu Gly Ile Phe
85 90 95

Asp Pro Lys Ile Leu Glu Thr Glu Asn Ala Tyr Leu Val Gly Tyr Trp
100 105 110

Gln Ser Glu Lys Tyr Phe Thr Ser Pro Glu Val Ile Glu Gln Ile Gln
115 120 125

37847-517 Sequence Listing_ST25 (003)

Glu Ala Phe Gly Lys Arg Pro Gln Glu Ile Met His Asp Ser Val Ser
130 135 140

Trp Ser Thr Leu Gln Gln Ile Glu Cys Cys Glu Ser Val Ser Ile His
145 150 155 160

Val Arg Arg Thr Asp Tyr Met Asp Ala Glu His Ile Lys Ile His Asn
165 170 175

Leu Cys Ser Glu Lys Tyr Tyr Lys Asn Ala Ile Ser Lys Ile Arg Glu
180 185 190

Glu His Pro Asn Ala Val Phe Phe Ile Phe Thr Asp Asp Lys Glu Trp
195 200 205

Cys Lys Glu His Phe Lys Gly Pro Lys Phe Ile Thr Val Glu Leu Gln
210 215 220

Glu Gly Glu Phe Thr Asp Val Ala Asp Met Leu Leu Met Ser Arg Cys
225 230 235 240

Lys His His Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Ser Ala Trp
245 250 255

Leu Asn Asp Ser Pro Glu Lys Ile Val Ile Ala Pro Ser Lys Trp Ile
260 265 270

Asn Asn Lys Lys Met Asp Asp Ile Tyr Thr Glu Arg Met Thr Lys Val
275 280 285

Ala Ile
290

<210> 138

<211> 286

<212> PRT

<213> Bacteroides ovatus

<400> 138

37847-517 Sequence Listing_ST25 (003)

Met Ile Val Val Tyr Ser Asn Ala Gly Leu Ala Asn Arg Met Phe His
1 5 10 15

Tyr Ala Leu Tyr Lys Ala Leu Glu Val Lys Gly Ile Asp Val Tyr Phe
20 25 30

Asp Glu Lys Ser Tyr Val Pro Glu Trp Ser Phe Glu Thr Thr Thr Leu
35 40 45

Met Asp Val Phe Pro Asn Ile Gln Tyr Arg Glu Ser Leu Gln Phe Lys
50 55 60

Arg Ala Ser Lys Lys Thr Phe Leu Asp Lys Ile Val Ile His Cys Ser
65 70 75 80

Asn Leu Phe Gly Gly Arg Tyr Tyr Val Asn Tyr Arg Phe Lys Tyr Asp
85 90 95

Asp Lys Leu Phe Thr Lys Leu Glu Thr Asn Gln Asp Leu Cys Leu Ile
100 105 110

Gly Leu Trp Gln Ser Glu Lys Tyr Phe Met Asp Val Arg Gln Glu Ile
115 120 125

Gln Lys Cys Phe Gln Tyr Arg Ser Phe Val Asp Asp Lys Asn Val Lys
130 135 140

Thr Ala Gln Gln Met Leu Ser Glu Asn Ser Val Ala Ile His Val Arg
145 150 155 160

Lys Gly Ala Asp Tyr Gln Gln Asn Arg Ile Trp Lys Asn Thr Cys Thr
165 170 175

Ile Asp Tyr Tyr Arg Leu Ala Ile Asp Tyr Ile Arg Met His Val Gln
180 185 190

Asn Pro Val Phe Tyr Val Phe Thr Asp Asn Lys Asp Trp Val Ile Glu
195 200 205

37847-517 Sequence Listing_ST25 (003)

Asn Phe Thr Asp Leu Asp Tyr Thr Leu Cys Asp Trp Asn Pro Thr Ser
210 215 220

Gly Lys Gln Asn Tyr Leu Asp Met Gln Leu Met Ser Cys Ala Lys His
225 230 235 240

Asn Val Ile Ala Asn Ser Thr Tyr Ser Trp Trp Gly Ala Trp Leu Asn
245 250 255

Glu Asn Ser Asp Lys Ile Val Ile Ala Pro Lys Arg Trp Phe Asn Lys
260 265 270

Ile Val Thr Pro Asp Ile Leu Pro Glu Gln Trp Ile Lys Ile
275 280 285

<210> 139

<211> 288

<212> PRT

<213> Mesotoga prima MesG1.Ag.4.2;Mesotoga prima

<400> 139

Met Arg Val Val Trp Phe Gly Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Gly Leu Tyr Cys Phe Leu Lys Lys Asn Asn Gln Glu Val Lys Ala
20 25 30

Asp Cys Thr Gln Tyr Ser Thr Thr Pro Met Asn Asn Gly Phe Glu Leu
35 40 45

Glu Arg Leu Phe Asn Leu Asp Ile Ala His Ala Asn Leu Asp Val Ile
50 55 60

Ser Lys Leu Thr Gly Gly Asn Arg Leu Ser Pro Arg Lys Val Ile Trp
65 70 75 80

Lys Leu Phe Arg Lys Pro Lys Val Tyr Phe Glu Glu Lys Ile Pro Phe
85 90 95

37847-517 Sequence Listing_ST25 (003)

Ser Phe Asp Pro Asp Val Leu Lys Gly Asn Asn Arg Tyr Leu Lys Gly
100 105 110

Tyr Trp Gln Asn Met Asn Tyr Leu Glu Pro Cys Ala Lys Glu Leu Arg
115 120 125

Asp Val Phe Thr Phe Pro Ala Phe Ser Ser Asp Asn Asn Lys Arg Leu
130 135 140

Ala Asp Glu Ile Ala Lys Val Glu Ala Val Gly Val His Phe Arg Arg
145 150 155 160

Gly Asp Phe Leu Lys Ser Ser Asn Leu Gly Leu Phe Gly Gly Ile Cys
165 170 175

Ser Asp Gln Tyr Tyr Leu Arg Ala Ile Gln Thr Met Glu Asn Thr Val
180 185 190

Val Glu Pro Val Phe Tyr Val Phe Cys Asp Asp Pro Gln Trp Ala Lys
195 200 205

Asn Ser Phe Ser Asp Ala Arg Phe Thr Val Ile Asp Trp Asn Ile Gly
210 215 220

Ser Asn Ser Tyr Arg Asp Met Gln Leu Met Ser Leu Cys Lys His Asn
225 230 235 240

Ile Ile Ala Asn Ser Thr Phe Ser Trp Trp Ala Ala Trp Leu Asn Arg
245 250 255

Asn Pro Asn Arg Thr Val Ile Ala Pro Glu Arg Met Val Asn Arg Asp
260 265 270

Leu Asp Phe Ser Gly Ile Phe Pro Asn Asp Trp Ile Arg Leu Gln Gly
275 280 285

<210> 140
<211> 296
<212> PRT
<213> Clostridium sp. KLE 1755

37847-517 Sequence Listing_ST25 (003)

<400> 140

Met Ile Val Leu Lys Leu Gln Gly Gly Leu Gly Asn Gln Met Phe Glu
 1 5 10 15

Tyr Ala Phe Ala Arg Thr Ile Gln Glu Gln Lys Lys Asp Lys Lys Leu
 20 25 30

Ile Leu Asp Thr Ser Asp Phe Gln Tyr Asp Lys Gln Arg Glu Tyr Ser
 35 40 45

Leu Gly His Phe Ile Leu Asn Glu Asn Ile Glu Ile Asp Ser Ser Gly
 50 55 60

Lys Phe Asn Leu Trp Tyr Asp Gln Arg Lys Asn Pro Leu Leu Lys Val
 65 70 75 80

Gly Phe Lys Phe Trp Pro Lys Phe Gln Phe Gln Thr Leu Lys Leu Phe
 85 90 95

Gly Ile Tyr Val Trp Asp Tyr Ala Lys Tyr Ile Pro Val Asp Val Ser
 100 105 110

Lys Lys His Lys Asn Ile Leu Leu His Gly Leu Trp Gln Ser Asp Lys
 115 120 125

Tyr Phe Ser Gln Ile Ser Glu Ile Ile Arg Lys Glu Phe Ala Val Lys
 130 135 140

Asp Glu Pro Ser Gln Gly Asn Lys Ala Trp Leu Glu Arg Ile Ser Ser
 145 150 155 160

Ala Asn Ala Val Cys Val His Ile Arg Arg Gly Asp Phe Leu Ala Lys
 165 170 175

Gly Ser Val Leu Leu Thr Cys Ser Asn Ser Tyr Tyr Leu Lys Ala Met
 180 185 190

Glu Ile Ile Ser Lys Lys Val Asn Glu Pro Glu Phe Phe Ile Phe Ser

37847-517 Sequence Listing_ST25 (003)

195

200

205

Asp Asp Ile Glu Asp Val Lys Lys Ile Phe Glu Phe Pro Gly Tyr Gln
 210 215 220

Ile Thr Leu Val Asn Gln Ser Asn Pro Asp Tyr Glu Glu Leu Arg Leu
 225 230 235 240

Met Ser Lys Cys Lys His Phe Ile Ile Ala Asn Ser Thr Phe Ser Trp
 245 250 255

Trp Ser Ser Leu Leu Ser Glu Asn Glu Asp Lys Val Ile Val Ala Pro
 260 265 270

Arg Leu Trp Tyr Ser Asp Gly Arg Asp Thr Ser Ala Leu Met Arg Asp
 275 280 285

Glu Trp Ile Ile Ile Asp Asn Glu
 290 295

<210> 141

<211> 289

<212> PRT

<213> Bacteroides plebeius CAG:211

<400> 141

Met Asp Leu Val Thr Leu Ser Gly Gly Leu Gly Asn Gln Met Phe Gln
 1 5 10 15

Phe Ala Phe Tyr Trp Ala Leu Lys Lys Arg Gly Lys Lys Val Phe Leu
 20 25 30

Tyr Lys Asn Lys Leu Ala Ala Lys Glu His Asn Gly Tyr Glu Leu Gln
 35 40 45

Thr Leu Phe Gly Val Glu Glu Lys Cys Val Asp Gly Leu Trp Met Thr
 50 55 60

Arg Leu Leu Gly Cys Pro Leu Leu Gly Lys Ile Leu Lys His Ile Leu
 65 70 75 80

37847-517 Sequence Listing_ST25 (003)

Phe Pro His Lys Ile Arg Glu Arg Val Leu Tyr Asn Tyr Ser Ile Tyr
85 90 95

Leu Pro Leu Phe Glu Arg Asn Gly Leu His Trp Val Gly Tyr Trp Gln
100 105 110

Ser Glu Lys Tyr Phe Gln Asp Val Ala Asp Asp Ile Arg Arg Ile Phe
115 120 125

Cys Phe Asp His Leu Ser Leu Asn Pro Ala Thr Ser Ala Ala Leu Lys
130 135 140

Cys Met Ser Glu Gln Val Ala Val Ser Val His Ile Arg Arg Gly Asp
145 150 155 160

Tyr Tyr Leu Pro Cys Asn Val Ala Thr Tyr Gly Gly Leu Cys Thr Val
165 170 175

Glu Tyr Tyr Glu Asn Ala Ile Arg Tyr Val Lys Glu Arg Tyr Pro Gln
180 185 190

Ala Val Phe Tyr Val Phe Ser Asp Asp Leu Asp Trp Val Arg Glu Asn
195 200 205

Ile Pro Ser Ala Gly Lys Met Val Phe Val Asp Trp Asn Arg Gly Lys
210 215 220

Asp Ser Trp Gln Asp Met Phe Leu Met Ser Lys Cys His His Asn Ile
225 230 235 240

Leu Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala Trp Leu Asn Thr His
245 250 255

Pro Glu Lys Leu Val Ile Ala Pro Glu Arg Trp Ala Asn Cys Pro Ala
260 265 270

Pro Asp Ala Leu Pro Asp Gly Trp Val Arg Ile Glu Gly Val Ser Arg
275 280 285

Arg

<210> 142
 <211> 308
 <212> PRT
 <213> Treponema lecithinolytikum

<400> 142

Met Ala Ile Lys Ile Val Lys Ile Ser Gly Gly Leu Gly Asn Gln Met
 1 5 10 15

Phe Cys Tyr Ala Phe Ala Cys Ala Leu Gln Lys Cys Gly His Lys Val
 20 25 30

Tyr Val Asp Thr Ser Leu Tyr Arg Lys Ala Thr Val His Ser Gly Ile
 35 40 45

Asp Phe Cys His Asn Gly Leu Glu Thr Glu Arg Leu Phe Gly Ile Lys
 50 55 60

Phe Asp Glu Ala Asp Thr Ala Asp Val Arg Arg Leu Ser Thr Ser Ala
 65 70 75 80

Glu Gly Leu Leu Asn Arg Ile Arg Arg Lys Tyr Phe Thr Lys Lys Thr
 85 90 95

His Tyr Ile Asp Thr Val Phe Lys Tyr Thr Pro Glu Leu Leu Ser Asp
 100 105 110

Lys Asn Asp Cys Tyr Leu Glu Gly Tyr Trp Gln Thr Glu Lys Tyr Phe
 115 120 125

Leu Pro Ile Glu Lys Asp Ile Arg Arg Leu Phe Thr Phe Arg Pro Thr
 130 135 140

Leu Ser Glu Lys Ser Ala Ala Val Gln Ser Ala Leu Gln Ala Gln Gln
 145 150 155 160

37847-517 Sequence Listing_ST25 (003)

Ala Ala Val Leu Ser Ala Ser Ile His Val Arg Arg Gly Asp Phe Leu
165 170 175

Asn Thr Lys Thr Leu Asn Val Cys Thr Glu Thr Tyr Tyr Asn Asn Ala
180 185 190

Ile Lys Tyr Ala Val Lys Lys His Ala Val Ser Arg Phe Tyr Ile Phe
195 200 205

Ser Asp Asp Ile Pro Trp Cys Arg Glu His Leu Cys Phe Cys Asn Ala
210 215 220

His Ala Val Phe Ile Asp Trp Asn Thr Gly Asn Asp Ser Trp Gln Asp
225 230 235 240

Met Ala Leu Met Ser Met Cys Arg Cys Asn Ile Ile Ala Asn Ser Ser
245 250 255

Phe Ser Trp Trp Ala Ala Trp Leu Asn Asn Ala Ser Asp Lys Thr Val
260 265 270

Leu Ala Pro Ala Ile Trp Asn Arg Arg Gln Leu Glu Tyr Val Asp Arg
275 280 285

Tyr Tyr Gly Tyr Asp Tyr Ser Asp Ile Val Pro Glu Ser Trp Ile Arg
290 295 300

Ile Pro Ile Asp
305

<210> 143
<211> 281
<212> PRT
<213> Bacteroides eggerthii

<400> 143

Met Arg Leu Ile Lys Met Thr Gly Gly Leu Gly Asn Gln Met Phe Ile
1 5 10 15

37847-517 Sequence Listing_ST25 (003)

Tyr	Ala	Phe	Tyr	Leu	Arg	Met	Lys	Lys	Arg	His	Thr	Asn	Thr	Arg	Ile
			20					25					30		
Asp	Leu	Ser	Asp	Met	Met	His	Tyr	Asn	Val	His	His	Gly	Tyr	Glu	Met
		35					40					45			
His	Arg	Val	Phe	Asn	Leu	Pro	Lys	Thr	Glu	Phe	Cys	Ile	Asn	Gln	Pro
	50					55					60				
Leu	Lys	Lys	Val	Ile	Glu	Phe	Leu	Phe	Phe	Lys	Lys	Ile	Tyr	Glu	Arg
65					70					75					80
Lys	Gln	Asp	Pro	Ser	Ser	Leu	Leu	Pro	Phe	Asp	Lys	Lys	Tyr	Leu	Trp
				85					90					95	
Pro	Leu	Leu	Tyr	Phe	Lys	Gly	Phe	Tyr	Gln	Ser	Glu	Arg	Phe	Phe	Ala
			100					105					110		
Asp	Met	Glu	Asn	Asp	Ile	Arg	Ile	Ala	Phe	Thr	Phe	Asn	Ser	Asp	Leu
		115					120					125			
Phe	Asn	Glu	Lys	Thr	Gln	Ala	Met	Leu	Thr	Gln	Ile	Lys	His	Asn	Glu
	130					135					140				
His	Ala	Val	Ser	Leu	His	Ile	Arg	Arg	Gly	Asp	Tyr	Leu	Glu	Pro	Lys
145					150					155					160
His	Trp	Lys	Thr	Thr	Gly	Ser	Val	Cys	Gln	Leu	Pro	Tyr	Tyr	Leu	Asn
				165					170					175	
Ala	Ile	Thr	Glu	Met	Asn	Lys	Arg	Ile	Glu	Gln	Pro	Ser	Tyr	Tyr	Val
			180					185					190		
Phe	Ser	Asp	Asp	Ile	Ala	Trp	Val	Lys	Glu	Asn	Leu	Pro	Leu	Pro	Gln
		195					200					205			
Ala	Val	Phe	Ile	Asp	Trp	Asn	Lys	Gly	Ala	Glu	Ser	Trp	Gln	Asp	Met
	210					215					220				

37847-517 Sequence Listing_ST25 (003)

Met Leu Met Ser His Cys Arg His His Ile Ile Cys Asn Ser Thr Phe
 225 230 235 240

Ser Trp Trp Gly Ala Trp Leu Asn Pro Arg Glu Asn Lys Thr Val Ile
 245 250 255

Met Pro Glu Arg Trp Phe Gln His Cys Asp Thr Pro Asn Ile Tyr Pro
 260 265 270

Asp Gly Trp Ile Lys Val Pro Val Asn
 275 280

<210> 144
 <211> 287
 <212> PRT
 <213> Bacteroides stercoris

<400> 144

Met Arg Phe Ile Lys Met Thr Gly Gly Leu Gly Asn Gln Met Phe Ile
 1 5 10 15

Tyr Ala Phe Tyr Met Arg Met Lys Lys His Tyr Ser Asn Thr Arg Ile
 20 25 30

Asp Leu Ser Asp Met Val His Tyr Lys Ala His Asn Gly Tyr Glu Met
 35 40 45

His Arg Val Phe Asn Leu Pro Pro Ile Glu Phe Arg Ile Asn Gln Pro
 50 55 60

Leu Lys Lys Val Ile Glu Phe Leu Phe Phe Lys Lys Ile Tyr Glu Arg
 65 70 75 80

Lys Gln Val Pro Ser Ser Leu Val Pro Tyr Asp Lys Lys Tyr Phe Trp
 85 90 95

Pro Leu Leu Tyr Phe Lys Gly Phe Tyr Gln Ser Glu Arg Phe Phe Ala
 100 105 110

Asp Met Ala Asp Asp Ile Arg Lys Ala Phe Thr Phe Asn Pro Arg Leu

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

115

120

125

Ser Asn Arg Lys Thr Lys Glu Met Ser Glu Gln Ile Asp His Asp Glu
130 135 140

Asn Ala Val Ser Ile His Val Arg Arg Gly Asp Tyr Leu Glu Pro Lys
145 150 155 160

Tyr Trp Lys Thr Thr Gly Cys Val Cys Gln Leu Pro Tyr Tyr Leu Asn
165 170 175

Ala Ile Ala Glu Met Asn Lys Arg Ile Ser Gln Pro Ser Tyr Tyr Val
180 185 190

Phe Ser Asp Asp Ile Ala Trp Val Lys Glu Asn Leu Pro Leu Pro Lys
195 200 205

Ala Phe Phe Ile Asp Trp Asn Lys Gly Ala Glu Ser Trp Gln Asp Met
210 215 220

Met Leu Met Ser Arg Cys Arg His His Ile Ile Cys Asn Ser Thr Phe
225 230 235 240

Ser Trp Trp Gly Ala Trp Leu Asn Pro Arg Glu Asn Lys Thr Val Ile
245 250 255

Met Pro Glu Arg Trp Phe Arg His Cys Glu Thr Pro Asp Ile Cys Pro
260 265 270

Asp Lys Trp Ile Lys Val Pro Ile Asn Gln Pro Asp Ser Ile Gln
275 280 285

<210> 145

<211> 290

<212> PRT

<213> Butyrivibrio proteoclasticus

<400> 145

Met Ile Ile Ile Gln Leu Lys Gly Gly Met Gly Asn Gln Met Phe Gln
1 5 10 15

37847-517 Sequence Listing_ST25 (003)

Tyr Ala Leu Tyr Arg Gln Leu Lys Lys Leu Gly Arg Glu Val Lys Ile
 20 25 30
 Asp Asp Glu Thr Gly Phe Val Asp Asp Glu Leu Arg Ile Pro Val Leu
 35 40 45
 Gln Arg Phe Gly Ile Ser Tyr Asp Lys Ala Thr Arg Glu Glu Ile Val
 50 55 60
 Lys Leu Thr Asp Ser Lys Met Asp Ile Phe Ser Arg Ile Arg Arg Lys
 65 70 75 80
 Leu Thr Gly Arg Lys Thr Phe Arg Ile Asp Glu Glu Ser Gly Ile Phe
 85 90 95
 Asp Pro Arg Ile Leu Glu Val Glu Asp Ala Tyr Leu Val Gly Tyr Trp
 100 105 110
 Gln Ser Asp Lys Tyr Phe Ala Asn Glu Glu Val Glu Lys Glu Ile Arg
 115 120 125
 Glu Ala Phe Glu Lys Arg Pro Gln Glu Val Met Gln Asp Ser Val Ser
 130 135 140
 Trp Thr Ile Leu Gln Gln Ile Glu Cys Cys Glu Ser Val Ser Leu His
 145 150 155 160
 Ile Arg Arg Thr Asp Tyr Ile Asp Glu Glu His Ile His Ile His Asn
 165 170 175
 Ile Cys Thr Glu Lys Tyr Tyr Lys Ser Ala Ile Asp Glu Val Arg Asn
 180 185 190
 Gln Tyr Pro Ser Ala Val Phe Phe Ile Phe Thr Asp Asp Lys Asp Trp
 195 200 205
 Cys Arg Gln His Phe Arg Gly Pro Asn Phe Phe Val Val Asp Leu Asp
 210 215 220

37847-517 Sequence Listing_ST25 (003)

Glu Asp Thr Asn Thr Asp Ile Ala Glu Met Thr Leu Met Ser Arg Cys
225 230 235 240

Lys His His Ile Leu Ala Asn Ser Ser Phe Ser Trp Trp Ala Ala Trp
245 250 255

Leu Asn Asp Asn Pro Gly Lys Ile Val Ile Ala Pro Ser Lys Trp Ile
260 265 270

Asn Asn Arg Lys Met Asp Asp Ile Tyr Thr Ala Arg Met Lys Lys Ile
275 280 285

Ala Ile
290

<210> 146

<211> 303

<212> PRT

<213> Prevotella oulora

<400> 146

Met Ser Pro Ile Val His Phe Pro Ser Asp Arg Leu Leu Arg Tyr Glu
1 5 10 15

His Leu Asn Ser Leu Trp Lys Thr Ala Met Ile Tyr Thr Arg Leu Leu
20 25 30

Ala Arg Leu Gly Asn Gln Met Phe Gln Tyr Ala Ala Gly Arg Gly Leu
35 40 45

Ala Ala Arg Leu Gly Val Asp Phe Thr Val Asp Ser Arg Arg Ala Val
50 55 60

His Lys Gly Asp Gly Val Leu Thr Arg Val Phe Asp Leu Asp Trp Ala
65 70 75 80

Ala Pro Glu Asn Met Pro Pro Ala Gln His Glu Arg Pro Leu Ala Tyr
85 90 95

37847-517 Sequence Listing_ST25 (003)

Tyr Ala Trp Arg Gly Leu Arg Arg Asp Pro Lys Ile Tyr Arg Glu Asn
 100 105 110

Gly Leu Gly Tyr Asn Ala Ala Phe Glu Thr Leu Pro Asp Asn Thr Tyr
 115 120 125

Leu His Gly Tyr Trp Gln Cys Glu Arg Tyr Phe Ala His Ile Ala Asp
 130 135 140

Asp Ile Arg Ala Ala Phe Val Pro Arg His Pro Met Ser Ala Gln Asn
 145 150 155 160

Ala Asp Met Ala Arg Arg Ile Ala Ser Gly Pro Ser Val Ser Leu His
 165 170 175

Val Arg Arg Gly Asp Tyr Leu Thr Val Gly Ala His Gly Ile Cys Asp
 180 185 190

Gln Thr Tyr Tyr Asp Ala Ala Leu Ala Ala Val Met Gln Gly Leu Pro
 195 200 205

Ser Pro Thr Val Tyr Val Phe Ser Asp Asp Pro Gln Trp Ala Lys Asp
 210 215 220

Asn Leu Pro Leu Thr Phe Glu Lys Val Val Val Asp Phe Asn Gly Pro
 225 230 235 240

Asp Ser Asp Tyr Glu Asp Met Arg Leu Met Ser Leu Cys Gln His Asn
 245 250 255

Val Ile Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala Trp Leu Asn Ala
 260 265 270

Asn Pro Gln Lys Arg Val Ala Gly Pro Ala Asn Trp Phe Ser Asn Pro
 275 280 285

Lys Leu Ser Asn Pro Asp Ile Leu Pro Ser Arg Trp Ile Arg Ile
 290 295 300

37847-517 Sequence Listing_ST25 (003)

<210> 147
 <211> 289
 <212> PRT
 <213> *Thalassobacter arenae*

<400> 147

Met Gly Gln Asp Met Ile Tyr Ser Arg Ile Phe Gly Gly Leu Gly Asn
 1 5 10 15

Gln Leu Phe Gln Tyr Ala Thr Ala Arg Ala Val Ser Leu Arg Gln Gly
 20 25 30

Val Glu Leu Val Leu Asp Thr Arg Leu Ala Pro Pro Gly Ser His Trp
 35 40 45

Ala Phe Gly Leu Asp His Phe Asn Ile Ser Ala Arg Ile Ala Glu Pro
 50 55 60

Ser Glu Leu Pro Pro Ser Lys Asp Asn Phe Phe Lys Tyr Val Met Trp
 65 70 75 80

Arg Ala Phe Gly His Asp Pro Ala Phe Met Arg Glu Arg Gly Leu Gly
 85 90 95

Tyr Gln Ser Arg Ile Ala Gln Ala Pro Asp Gly Thr Tyr Leu His Gly
 100 105 110

Tyr Phe Gln Ser Glu Arg Tyr Phe Ala Asp Val Leu Asp His Leu Glu
 115 120 125

Asn Glu Leu Arg Ile Val Thr Pro Pro Asp Thr Arg Asn Ala Glu Tyr
 130 135 140

Ala Asp Arg Ile Ala Ser Ala Gly His Thr Val Ser Leu His Val Arg
 145 150 155 160

Arg Gly Asp Tyr Val Glu Thr Ser Lys Ser Asn Ser Thr His Ala Thr
 165 170 175

37847-517 Sequence Listing_ST25 (003)

Cys Asp Glu Ala Tyr Tyr Leu Arg Ala Leu Ala Arg Leu Ser Glu Gly
180 185 190

Lys Ser Asp Leu Lys Val Phe Val Phe Ser Asp Asp Pro Glu Trp Val
195 200 205

Arg Asp Asn Leu Lys Leu Pro Tyr Asp Thr Thr Pro Val Gly His Asn
210 215 220

Gly Pro Asp Lys Pro His Glu Asp Leu Arg Leu Met Ser Cys Cys Ser
225 230 235 240

Asp His Val Ile Ala Asn Ser Thr Phe Ser Trp Trp Gly Gly Trp Leu
245 250 255

Asp Arg Arg Pro Glu Ala Arg Val Val Gly Pro Ala Lys Trp Phe Asn
260 265 270

Asn Pro Lys Leu Val Asn Pro Asp Ile Leu Pro Glu Arg Trp Ile Ala
275 280 285

Ile

<210> 148
<211> 290
<212> PRT
<213> Prevotella oris

<400> 148

Met Lys Ile Ile Lys Ile Ile Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Leu Ala Val Ala Leu Gln Lys Lys Trp Lys Asp Glu Glu Ile
20 25 30

Lys Leu Asp Leu His Gly Phe Asn Gly Tyr His Lys His Gln Gly Tyr
35 40 45

Gln Leu Asp Glu Ile Phe Gly His Arg Phe Lys Ala Ala Ser Leu Lys

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

50

55

60

Glu Val Ala Gln Leu Ala Trp Pro Tyr Pro His Tyr Gln Leu Trp Arg
65 70 75 80

Val Gly Ser Arg Leu Leu Pro Lys Arg Lys Thr Met Val Cys Glu Ser
85 90 95

Ala Asp Cys Arg Phe Gln Ser Asp Leu Leu Asn Leu Glu Gly Ser Leu
100 105 110

Tyr Tyr Asp Gly Tyr Trp Gln Asp Glu Arg Tyr Phe Lys Ala Phe Arg
115 120 125

Thr Glu Ile Ile Glu Ala Phe Lys Phe Thr Pro Leu Val Gly Asp Ser
130 135 140

Asn Arg Lys Val Glu Asn Met Leu Lys Glu Gly Arg Phe Ala Ser Leu
145 150 155 160

His Val Arg Arg Gly Asp Tyr Leu Lys Glu Pro Leu Phe Gln Ser Thr
165 170 175

Cys Asp Ile Ala Tyr Tyr Gln Arg Ala Ile Ser Arg Leu Asn Gln Met
180 185 190

Ala Asp Pro Tyr Cys Tyr Leu Ile Phe Ser Asn Asp Ile Ala Trp Cys
195 200 205

Lys Thr His Ile Glu Pro Leu Cys Asp Gly Arg Arg Thr His Tyr Val
210 215 220

Asp Trp Asn His Gly Lys Glu Ser Tyr Arg Asp Met Gln Leu Met Thr
225 230 235 240

Phe Cys Lys His His Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Gly
245 250 255

Ala Trp Leu Ser Thr Ala Asn Asp Gly Ile Thr Ile Ala Pro His Gln

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

260

265

270

Trp Tyr Ala Asn Asp Arg Lys Pro Ser Pro Ala Ala Glu Ala Trp Leu
275 280 285

Lys Leu
290

<210> 149
<211> 290
<212> PRT
<213> Prevotella oulora

<400> 149

Met Lys Ile Val Arg Ile Ile Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Leu Ala Leu Ala Leu Lys Gln Gln Gln Glu Asn Glu Glu Val
20 25 30

Lys Leu Asp Leu Ser Ala Phe Arg Gly Tyr Lys Lys His Gly Gly Phe
35 40 45

Gln Leu Val Gln Cys Phe Gly Thr Thr Leu Pro Ala Ala Thr Trp Gln
50 55 60

Glu Val Ala Gln Leu Ala Trp Tyr Tyr Pro His Tyr Gln Leu Trp Arg
65 70 75 80

Leu Gly His Arg Val Leu Pro Val Cys Lys Thr Met Leu Lys Glu Pro
85 90 95

Asp Asn Gly Ala Phe Leu Pro Glu Val Leu Gln Arg Lys Gly Asp Ala
100 105 110

Tyr Tyr Glu Gly Cys Trp Gln Asp Glu Arg Tyr Phe Ser His Tyr Arg
115 120 125

Pro Ala Ile Leu Gln Ala Phe Thr Phe Pro Thr Phe Thr Asn Pro Arg
130 135 140

37847-517 Sequence Listing_ST25 (003)

Asn Leu Ala Met Gln Gln Gln Ile Asn Thr Thr Glu Ser Val Ala Ile
145 150 155 160

His Val Arg Arg Gly Asp Tyr Leu His Asp Ala Leu Phe Arg Asn Thr
165 170 175

Cys Gly Leu Ala Tyr Phe Gln Arg Ala Ile Thr Cys Ile Leu Gln His
180 185 190

Val Ala His Pro Val Phe Tyr Val Phe Ser Asp Asp Met Ala Trp Cys
195 200 205

Arg Gln His Ile Gln Pro Leu Leu Gln Thr Asn Glu Ala Val Phe Val
210 215 220

Asp Trp Asn His Gly Lys Ala Ser Ile Cys Asp Leu His Leu Met Thr
225 230 235 240

Leu Cys Arg His His Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Gly
245 250 255

Ala Trp Leu Ser Pro His Gln Ala Gly Trp Ile Ile Ala Pro Lys Gln
260 265 270

Trp Tyr Ala His Glu Glu Lys Met Ser Pro Ala Ala Glu Arg Trp Leu
275 280 285

Lys Leu
290

<210> 150
<211> 298
<212> PRT
<213> Spirosoma panaciterrae

<400> 150

Met Asn Arg Arg Val Ala Val Gln Leu Lys Gly Gly Leu Gly Asn Gln
1 5 10 15

37847-517 Sequence Listing_ST25 (003)

Leu Phe Gln Tyr Ala Leu Gly Arg Arg Leu Ser Leu Gln Leu Glu Ala
 20 25 30
 Glu Leu Leu Phe Asp Cys Ser Val Leu Glu Asn Arg Ile Pro Val Thr
 35 40 45
 Asn Phe Thr Phe Arg Ser Phe Asp Leu Asp Met Phe Arg Ile Ala Gly
 50 55 60
 Arg Val Ala Thr Pro Ser Asp Leu Pro Leu Phe Pro Lys Ser Ala Ser
 65 70 75 80
 Ile Arg Ser Pro Trp Pro His Leu Val Gln Leu Ala Arg Leu Trp Lys
 85 90 95
 Gln Gly Tyr Ser Tyr Val Tyr Glu Arg Gly Phe Ala Tyr Asn Pro Lys
 100 105 110
 Met Leu Arg Gln Leu Ser Asp Arg Val Tyr Leu Asn Gly Tyr Trp Gln
 115 120 125
 Ser Tyr Arg Tyr Phe Glu Asp Ile Ala Ala Thr Leu Arg Ala Asp Cys
 130 135 140
 Ser Phe Pro Asp Pro Leu Pro Asp Ser Ala Val Gly Leu Ala Gly Gln
 145 150 155 160
 Ile Asn Ala Thr Asn Ser Ile Cys Leu His Ile Arg Arg Thr Asp Phe
 165 170 175
 Leu Gln Val Pro Leu His Gln Val Ser Asn Ala Asp Tyr Val Gly Arg
 180 185 190
 Ala Ile Ala Tyr Met Ala Glu Arg Val Asn Asp Pro His Phe Phe Val
 195 200 205
 Phe Ser Asp Asp Ile Ala Trp Cys Gln Thr Asn Leu Arg Leu Ser Tyr
 210 215 220

37847-517 Sequence Listing_ST25 (003)

Pro Val Val Phe Val Pro Asn Glu Leu Ala Gly Pro Lys Asn Ser Leu
225 230 235 240

His Phe Arg Leu Met Arg Tyr Cys Lys His Phe Ile Thr Ala Asn Ser
245 250 255

Thr Phe Ser Trp Trp Ala Ala Trp Leu Ser Glu Pro Ser Asp Gly Lys
260 265 270

Val Ile Val Thr Pro Gln Thr Trp Phe Ser Asp Ser Arg Ser Ile Asp
275 280 285

Asp Leu Ile Pro Ala Asn Trp Ile Arg Leu
290 295

<210> 151
<211> 267
<212> PRT
<213> Butyrivibrio proteoclasticus

<400> 151

Met Asn Tyr Val Glu Val Lys Gly Gly Leu Gly Asn Gln Leu Phe Gln
1 5 10 15

Tyr Thr Phe Tyr Lys Tyr Leu Glu Lys Lys Ser Gly His Lys Val Leu
20 25 30

Leu His Thr Asp Phe Phe Lys Asn Ile Asp Ser Phe Glu Glu Ala Thr
35 40 45

Lys Arg Lys Leu Gly Leu Asp Arg Phe Asp Cys Asp Phe Val Ala Val
50 55 60

Ser Gly Phe Ile Ser Cys Glu Lys Leu Val Lys Glu Ser Asp Tyr Lys
65 70 75 80

Asp Ser Met Leu Ser Gln Asp Glu Val Phe Tyr Ser Gly Tyr Trp Gln
85 90 95

37847-517 Sequence Listing_ST25 (003)

Asn Lys Arg Phe Phe Leu Glu Val Met Asp Asp Ile Arg Lys Asp Leu
100 105 110

Leu Leu Lys Asp Glu Asn Ile Gln Asp Glu Val Lys Glu Leu Ala Lys
115 120 125

Glu Leu Arg Ala Val Asp Ser Val Ala Ile His Phe Arg Arg Gly Asp
130 135 140

Tyr Leu Ser Glu Gln Asn Lys Lys Ile Phe Thr Ser Leu Ser Val Asp
145 150 155 160

Tyr Tyr Gln Lys Ala Ile Ala Gln Leu Ala Glu Arg Asn Gly Ala Asp
165 170 175

Leu Lys Gly Tyr Ile Phe Thr Asp Glu Pro Glu Tyr Val Ser Gly Ile
180 185 190

Ile Asp Gln Leu Gly Ser Ile Asp Ile Lys Leu Met Pro Val Arg Glu
195 200 205

Asp Tyr Glu Asp Leu Tyr Leu Met Ser Cys Ala Arg His His Ile Ile
210 215 220

Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala Ala Leu Gly Asp Thr Glu
225 230 235 240

Ser Gly Ile Thr Ile Ala Pro Ala Lys Trp Tyr Val Asp Gly Arg Thr
245 250 255

Pro Asp Leu Tyr Leu Arg Asn Trp Ile Ser Ile
260 265

<210> 152
<211> 290
<212> PRT
<213> Butyrivibrio sp. XPD2006

<400> 152

Met Ile Ile Ile Gln Leu Lys Gly Gly Leu Gly Asn Gln Met Phe Gln

28 May 2018
2015259088

37847-517 Sequence Listing_ST25 (003)

1				5						10					15
Tyr	Ala	Leu	Tyr	Lys	Glu	Leu	Lys	His	Arg	Gly	Arg	Glu	Val	Lys	Ile
			20					25					30		
Asp	Asp	Val	Ser	Gly	Phe	Val	Asn	Asp	Lys	Leu	Arg	Val	Pro	Val	Leu
		35					40					45			
Asp	Arg	Phe	Gly	Val	Glu	Tyr	Glu	Arg	Ala	Thr	Arg	Glu	Glu	Val	Val
	50					55					60				
Glu	Leu	Thr	Asp	Ser	Arg	Met	Asp	Ile	Phe	Ser	Arg	Ile	Arg	Arg	Lys
65					70					75					80
Leu	Thr	Gly	Arg	Lys	Thr	Tyr	Arg	Ile	Asp	Glu	Met	Glu	Gly	Ile	Phe
				85					90					95	
Asp	Pro	Ala	Ile	Leu	Glu	Thr	Glu	Asn	Ala	Tyr	Leu	Val	Gly	Tyr	Trp
			100					105					110		
Gln	Ser	Glu	Lys	Tyr	Phe	Thr	Ser	Pro	Glu	Val	Ile	Glu	Gln	Ile	Gln
		115					120					125			
Glu	Ala	Phe	Gly	Lys	Arg	Pro	Gln	Glu	Ile	Met	His	Asp	Ser	Val	Ser
	130					135					140				
Trp	Ser	Thr	Leu	Gln	Gln	Ile	Glu	Cys	Cys	Glu	Ser	Val	Ser	Ile	His
145				150						155					160
Val	Arg	Arg	Thr	Asp	Tyr	Val	Asp	Ala	Glu	His	Ile	Lys	Ile	His	Asn
				165					170					175	
Leu	Cys	Ser	Glu	Lys	Tyr	Tyr	Lys	Asn	Ala	Ile	Gly	Lys	Ile	Arg	Glu
			180					185					190		
Lys	His	Pro	Asn	Ala	Val	Phe	Phe	Ile	Phe	Thr	Asp	Asp	Lys	Glu	Trp
		195					200					205			
Cys	Lys	Asp	His	Phe	Lys	Gly	Pro	Asn	Phe	Ile	Thr	Val	Glu	Leu	Gln

37847-517 Sequence Listing_ST25 (003)

210

215

220

Glu Gly Glu Phe Thr Asp Val Ala Asp Met Leu Leu Met Ser Arg Cys
 225 230 235 240

Lys His His Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Ser Ala Trp
 245 250 255

Leu Asn Asp Ser Pro Glu Lys Met Val Ile Ala Pro Ser Lys Trp Ile
 260 265 270

Asn Asn Lys Lys Met Asp Asp Ile Tyr Thr Glu Arg Met Thr Arg Val
 275 280 285

Ala Ile
 290

<210> 153
 <211> 306
 <212> PRT
 <213> Bacteroides sp. 1_1_6

<400> 153

Met Lys Ile Val Asn Ile Thr Gly Gly Leu Gly Asn Gln Met Phe Gln
 1 5 10 15

Tyr Ala Phe Ala Met Ala Leu Lys Tyr Arg Asn Pro Gln Glu Glu Val
 20 25 30

Phe Val Asp Ile Gln His Tyr Asn Thr Ile Phe Phe Lys Lys Phe Lys
 35 40 45

Gly Ile Asn Leu His Asn Gly Tyr Glu Ile Asp Lys Val Phe Pro Lys
 50 55 60

Ala Lys Leu Pro Val Ala Gly Val Arg Gln Leu Met Lys Phe Ser Tyr
 65 70 75 80

Trp Ile Pro Asn Tyr Ile Leu Ser Arg Leu Gly Arg Lys Phe Leu Pro
 85 90 95

37847-517 Sequence Listing_ST25 (003)

Ile Arg Lys Lys Glu Tyr Ile Pro Pro Tyr Ser Met Asn Tyr Ser Tyr
100 105 110

Asp Glu Lys Ala Leu Asn Trp Lys Gly Asp Gly Tyr Phe Glu Gly Tyr
115 120 125

Trp Gln Ser Tyr Asn His Phe Gly Asp Ile Lys Glu Glu Leu Gln Lys
130 135 140

Val Tyr Ala His Pro Lys Pro Asn Gln Tyr Asn Ala Ala Leu Ile Ser
145 150 155 160

Asn Leu Glu Ser Cys Asn Ser Val Gly Ile His Val Arg Arg Gly Asp
165 170 175

Tyr Leu Ala Glu Pro Glu Phe Arg Gly Ile Cys Gly Leu Asp Tyr Tyr
180 185 190

Glu Lys Gly Ile Lys Glu Ile Leu Ser Asp Glu Lys Lys Tyr Val Phe
195 200 205

Phe Ile Phe Ser Asn Asp Met Gln Trp Cys Gln Glu Asn Ile Ala Pro
210 215 220

Leu Val Gly Asp Asn Arg Ile Val Phe Ile Ser Gly Asn Lys Gly Lys
225 230 235 240

Asp Ser Cys Trp Asp Met Phe Leu Met Thr His Cys Lys Asp Leu Ile
245 250 255

Ile Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala Phe Leu Asn Lys Lys
260 265 270

Val Asp Arg Val Ile Cys Pro Lys Pro Trp Leu Asn Arg Asp Cys Asn
275 280 285

Ile Asp Ile Tyr Asn Pro Ser Trp Ile Leu Val Pro Cys Tyr Ser Glu
290 295 300

37847-517 Sequence Listing_ST25 (003)

Asp Trp
305

<210> 154
<211> 271
<212> PRT
<213> Bacteroides fragilis

<400> 154

Met Lys Ile Val Thr Phe Gln Gly Gly Leu Gly Asn Gln Leu Phe Gln
1 5 10 15

Tyr Val Phe Tyr Leu Trp Leu Asp Met Arg Cys Asp Lys Asp Asn Ile
20 25 30

Tyr Gly Tyr Tyr Pro Lys Lys Gly Leu Arg Ala His Asn Gly Leu Glu
35 40 45

Ile Glu Lys Val Phe Glu Val Lys Leu Pro Asn Ser Ser Leu Ser Thr
50 55 60

Asp Leu Ile Val Lys Ser Ile Lys Leu Ile Asn Lys Ile Phe Lys Asn
65 70 75 80

Arg Gln Tyr Ile Ser Thr Asp Gly Arg Leu Asp Val Asn Gly Val Leu
85 90 95

Phe Glu Gly Phe Trp Gln Asp Lys Tyr Phe Trp Glu Asp Val Asp Ile
100 105 110

Val Leu Asn Phe Arg Trp Pro Leu Lys Leu Asp Val Thr Asn Ser Phe
115 120 125

Ile Met Thr Lys Ile Gln Ala Asn Asn Ser Ile Ser Ile His Ile Arg
130 135 140

Arg Gly Asp Tyr Leu Leu Pro Lys Tyr Arg Asn Ile Tyr Gly Asp Ile
145 150 155 160

37847-517 Sequence Listing_ST25 (003)

Cys Asn Glu Glu Tyr Tyr Gln Lys Ala Ile Glu Tyr Ile Leu Lys Cys
 165 170 175

Val Asp Asp Pro Phe Phe Phe Val Phe Ser Asp Asp Ile Asp Trp Ala
 180 185 190

Lys Ser Ile Ile Asn Val Ser Asn Val Thr Phe Val Asn Asn Asn Lys
 195 200 205

Gly Lys Asp Ser Tyr Ile Asp Met Phe Leu Met Ser Leu Cys His His
 210 215 220

Asn Ile Ile Ala Asn Ser Thr Phe Ser Trp Trp Ala Ala Gln Leu Asn
 225 230 235 240

Lys His Ser Asp Lys Ile Met Ile Ala Pro Ile Arg Trp Phe Lys Ser
 245 250 255

Leu Phe Lys Asp Pro Asn Ile Phe Thr Glu Ser Trp Ile Arg Ile
 260 265 270

<210> 155
 <211> 266
 <212> PRT
 <213> Bacteroides sp. 9_1_42FAA

<400> 155

Met Ile Lys Ile Val Ser Phe Ser Gly Gly Leu Gly Asn Gln Leu Phe
 1 5 10 15

Gln Tyr Leu Leu Val Val Tyr Leu Arg Glu Cys Gly His Gln Val Tyr
 20 25 30

Gly Tyr Tyr Asn Arg Lys Trp Leu Ile Gly His Asn Gly Leu Glu Val
 35 40 45

Asn Asn Val Phe Asp Ile Tyr Leu Pro Lys Thr Asn Phe Ile Val Asn
 50 55 60

37847-517 Sequence Listing_ST25 (003)

Ala Leu Val Lys Val Ile Arg Val Leu Arg Cys Leu Gly Phe Lys Lys
65 70 75 80

Tyr Val Ala Thr Asp Thr Tyr Asn Asn Pro Ile Ala Ile Phe Tyr Asp
85 90 95

Gly Tyr Trp Gln Asp Gln Lys Tyr Phe Asn Ile Ile Asp Ser Lys Leu
100 105 110

Ser Phe Lys Lys Phe Asp Leu Ser Ala Glu Asn Lys Ser Ile Leu Ser
115 120 125

Lys Ile Lys Ser Asn Ile Ser Val Ala Leu His Ile Arg Cys Gly Asp
130 135 140

Tyr Leu Ser Ser Ser Asn Val Glu Ile Tyr Gly Gly Val Cys Thr Lys
145 150 155 160

Glu Tyr Tyr Glu Lys Ala Leu Glu Leu Val Cys Lys Ile Lys Asn Val
165 170 175

Met Phe Phe Val Phe Ser Asp Asp Ile Glu Tyr Ala Lys Leu Leu Leu
180 185 190

Asn Leu Pro Asn Ala Ile Tyr Val Asn Ala Asn Val Gly Asn Ser Ser
195 200 205

Phe Ile Asp Met Tyr Leu Met Ala Asn Cys Lys Val Asn Val Ile Ala
210 215 220

Asn Ser Thr Phe Ser Tyr Trp Ala Ala Arg Leu Asn Gln Asp Asn Ile
225 230 235 240

Leu Thr Ile Tyr Pro Lys Lys Trp Tyr Asn Ser Lys Tyr Ala Val Pro
245 250 255

Asp Ile Phe Pro Ser Glu Trp Val Gly Val
260 265

37847-517 Sequence Listing_ST25 (003)

<210> 156
 <211> 298
 <212> PRT
 <213> Coraliomargarita sp. CAG:312

<400> 156

Met Ile Ile Val Lys Val Gln Gly Gly Leu Gly Asn Gln Met Phe Gln
 1 5 10 15

Tyr Ala Phe Gly Arg Ala Leu Ser Glu Lys His Ser Gln Asp Leu Tyr
 20 25 30

Leu Asp Cys Ser Glu Tyr Leu Arg Pro Ser Cys Lys Arg Glu Tyr Gly
 35 40 45

Leu Asp His Phe Asn Ile Arg Ala Lys Lys Ala Ser Cys Gly Asp Val
 50 55 60

Lys Ser Met Val Thr Pro His Phe Ala Leu Arg Lys Lys Leu Lys Lys
 65 70 75 80

Ile Phe Ala Val Pro Tyr Ser Leu Ser Pro Thr His Ile Leu Glu Arg
 85 90 95

Asn Phe Asn Phe Gln Pro Ser Ile Leu Glu Phe Asn Cys Gly Tyr Phe
 100 105 110

Asp Gly Phe Trp Gln Thr Gln Lys Tyr Phe Ser Gly Ile Ser Asp Ile
 115 120 125

Val Arg Lys Asp Leu Thr Phe Lys Asp Ala Val Lys Tyr Ser Gly Gly
 130 135 140

Glu Thr Phe Ala Lys Ile Thr Ser Leu Asn Ser Val Ser Leu His Ile
 145 150 155 160

Arg Arg Gly Asp Tyr Val Lys Val Lys Arg Thr Arg Lys Arg Phe Ser
 165 170 175

Val Ile Arg Ala Gly Tyr Phe Lys Arg Ala Val Glu Tyr Met Arg Ser

37847-517 Sequence Listing_ST25 (003)

180

185

190

Lys Leu Asp Thr Pro His Phe Phe Ile Phe Thr Asp Asp Pro Lys Trp
195 200 205

Val Ser Glu Asn Phe Pro Ala Gly Glu Asp Tyr Thr Leu Val Ser Ser
210 215 220

Ser Gly Met Tyr Glu Asp Leu Phe Leu Met Ala Gln Cys Arg His Asn
225 230 235 240

Ile Ile Phe Asn Ser Ser Phe Ser Trp Trp Gly Ala Trp Leu Asn Gly
245 250 255

Asn Pro Gly Lys Ile Val Val Ala Pro Asp Met Trp Phe Thr Pro His
260 265 270

Tyr Lys Leu Asp Tyr Ser Asp Val Val Pro Glu Glu Trp Ile Lys Leu
275 280 285

Asn Thr Gly Tyr Phe Glu Ser Lys Glu Phe
290 295

<210> 157

<211> 289

<212> PRT

<213> Pseudorhodobacter ferrugineus

<400> 157

Met Ile Val Met Gln Ile Lys Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Ala Gly Arg Ala Leu Ser Leu Gln Thr Gly Met Pro Leu His
20 25 30

Leu Asp Leu Arg Tyr Tyr Arg Arg Glu Arg Glu His Gly Tyr Gly Leu
35 40 45

Gly Ala Phe Asn Ile Glu Ala Ser Pro Leu Asp Glu Ser Leu Leu Pro
50 55 60

37847-517 Sequence Listing_ST25 (003)

Pro Leu Pro Arg Glu Ser Pro Leu Ala Trp Leu Ile Trp Arg Leu Gly
65 70 75 80

Arg Arg Gly Pro Asn Leu Val Arg Glu Asn Gly Met Gly Phe Asn Pro
85 90 95

Thr Leu Ser Asn Val Thr Lys Pro Ala Trp Ile Thr Gly Tyr Phe Gln
100 105 110

Ser Glu Arg Tyr Phe Ala Ala His Ala Ala Thr Ile Arg Ala Glu Leu
115 120 125

Thr Pro Val Ala Ala Pro Asp Leu Val Asn Ala Arg Trp Leu Ala Glu
130 135 140

Ile Ala Ala Glu Pro Arg Ala Val Ser Leu His Val Arg Arg Gly Asp
145 150 155 160

Tyr Val Arg Asp Ala Lys Ala Ala Ala Lys His Gly Ser Cys Thr Pro
165 170 175

Ala Tyr Tyr Glu Arg Ala Leu Ala His Ile Thr Ala Arg Met Gly Thr
180 185 190

Ala Pro Val Val Tyr Ala Phe Ser Asp Asp Pro Ala Trp Val Arg Glu
195 200 205

Asn Leu Arg Leu Pro Ala Glu Ile Arg Val Pro Gly His Asn Asp Thr
210 215 220

Ala Gly Asn Val Glu Asp Leu Arg Leu Met Ser Ala Cys Arg His His
225 230 235 240

Ile Val Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala Trp Leu Asn Pro
245 250 255

Arg Ala Asp Lys Ile Val Ala Ser Pro Ala Arg Trp Phe Ala Asp Pro
260 265 270

37847-517 Sequence Listing_ST25 (003)

Ala Phe Thr Asn Pro Asp Ile Trp Pro Glu Ala Trp Ala Arg Ile Glu
 275 280 285

Gly

<210> 158
 <211> 299
 <212> PRT
 <213> Escherichia coli

<400> 158

Met Met Tyr Cys Cys Leu Ser Gly Gly Leu Gly Asn Gln Met Phe Gln
 1 5 10 15

Tyr Ala Ala Ala Tyr Ile Leu Lys Gln His Phe Pro Asp Thr Ile Leu
 20 25 30

Val Leu Asp Asp Ser Tyr Tyr Phe Asn Gln Pro Gln Lys Asp Thr Ile
 35 40 45

Arg His Leu Glu Leu Asp Gln Phe Lys Ile Ile Phe Asp Arg Phe Ser
 50 55 60

Ser Lys Asp Glu Lys Val Lys Ile Asn Arg Leu Arg Lys His Lys Lys
 65 70 75 80

Ile Pro Leu Leu Asn Ser Phe Leu Gln Phe Thr Ala Ile Lys Leu Cys
 85 90 95

Asn Lys Tyr Ser Leu Asn Asp Ala Ser Tyr Tyr Asn Pro Glu Ser Ile
 100 105 110

Lys Asn Ile Asp Val Ala Cys Leu Phe Ser Phe Tyr Gln Asp Ser Lys
 115 120 125

Leu Leu Asn Glu His Arg Asp Leu Ile Leu Pro Leu Phe Glu Ile Arg
 130 135 140

37847-517 Sequence Listing_ST25 (003)

Asp Asp Leu Arg Val Leu Cys His Asn Leu Gln Ile Tyr Ser Leu Ile
145 150 155 160

Thr Asp Ser Lys Asn Ile Thr Ser Ile His Val Arg Arg Gly Asp Tyr
165 170 175

Val Asn Asn Lys His Ala Ala Lys Phe His Gly Thr Leu Ser Met Asp
180 185 190

Tyr Tyr Ile Ser Ala Met Glu Tyr Ile Glu Ser Glu Cys Gly Ser Gln
195 200 205

Thr Phe Ile Ile Phe Thr Asp Asp Val Ile Trp Ala Lys Glu Lys Phe
210 215 220

Ser Lys Tyr Ser Asn Cys Leu Val Ala Asp Ala Asp Glu Asn Lys Phe
225 230 235 240

Ser Val Ile Asp Met Tyr Leu Met Ser Leu Cys Asn Asn Asn Ile Ile
245 250 255

Ala Asn Ser Thr Tyr Ser Trp Trp Gly Ala Trp Leu Asn Arg Ser Glu
260 265 270

Asp Lys Leu Val Ile Ala Pro Lys Gln Trp Tyr Ile Ser Gly Asn Glu
275 280 285

Cys Ser Leu Lys Asn Glu Asn Trp Ile Ala Met
290 295

<210> 159
<211> 299
<212> PRT
<213> Lachnospiraceae bacterium 3_1_57FAA_CT1

<400> 159

Met Ile Ile Ile Lys Val Met Gly Gly Leu Gly Asn Gln Met Gln Gln
1 5 10 15

37847-517 Sequence Listing_ST25 (003)

Tyr	Ala	Leu	Tyr	Glu	Lys	Phe	Lys	Ser	Ile	Gly	Lys	Asn	Val	Lys	Leu	20	25	30
Asp	Ile	Ser	Trp	Phe	Glu	Asp	Ser	Ser	Val	Gln	Glu	Lys	Val	Phe	Ala	35	40	45
Arg	Arg	Ser	Leu	Glu	Leu	Arg	Gln	Phe	Lys	Asp	Leu	Gln	Phe	Asp	Thr	50	55	60
Cys	Ser	Ala	Glu	Glu	Lys	Glu	Ala	Leu	Leu	Gly	Lys	Ser	Gly	Ile	Leu	65	70	75
Gly	Lys	Leu	Glu	Arg	Lys	Leu	Ile	Pro	Ala	Arg	Asn	Lys	His	Phe	Tyr	85	90	95
Glu	Ser	Asp	Ile	Tyr	His	Ser	Glu	Val	Phe	Asn	Met	Ser	Asp	Ala	Tyr	100	105	110
Leu	Glu	Gly	His	Trp	Ala	Cys	Glu	Lys	Tyr	Tyr	His	Asp	Ile	Met	Pro	115	120	125
Leu	Leu	Gln	Glu	Lys	Ile	Gln	Phe	Pro	Glu	Ser	Ala	Asn	Ser	Gln	Asn	130	135	140
Ile	Thr	Val	Lys	Lys	Arg	Met	Lys	Ala	Glu	Asn	Ser	Val	Ser	Ile	His	145	150	155
Ile	Arg	Arg	Gly	Asp	Tyr	Leu	Asp	Pro	Glu	Asn	Glu	Ala	Met	Phe	Gly	165	170	175
Gly	Ile	Cys	Thr	Asn	Ser	Tyr	Tyr	Lys	Ala	Ala	Glu	Glu	Tyr	Ile	Lys	180	185	190
Ser	Arg	Val	Pro	Asp	Thr	His	Phe	Tyr	Leu	Phe	Ser	Asp	Asp	Thr	Ala	195	200	205
Tyr	Leu	Arg	Glu	Asn	Tyr	His	Gly	Asp	Glu	Tyr	Thr	Ile	Val	Asp	Trp	210	215	220

37847-517 Sequence Listing_ST25 (003)

Asn Lys Gly Glu Asp Ser Phe Tyr Asp Met Glu Leu Met Ser Cys Cys
225 230 235 240

Arg His Asn Ile Cys Ala Asn Ser Thr Phe Ser Phe Trp Gly Ala Arg
245 250 255

Leu Asn Arg Thr Pro Asp Lys Ile Val Ile Arg Pro Ala Lys His Lys
260 265 270

Asn Ser Gln Glu Ile Glu Pro Gln Leu Leu His Glu Leu Trp Asp Asn
275 280 285

Trp Val Ile Ile Asp Gly Asp Gly Arg Ile Val
290 295

<210> 160

<211> 733

<212> PRT

<213> Butyrivibrio fibrisolvens

<400> 160

Met Lys Pro Leu Val Ser Leu Ile Val Pro Val Leu Asn Val Glu Lys
1 5 10 15

Tyr Leu Glu Gln Cys Leu Thr Ser Ile Ser Ser Gln Thr Tyr Asp Asn
20 25 30

Phe Glu Val Ile Leu Val Val Gly Lys Cys Ile Asp Asn Ser Glu Asn
35 40 45

Ile Cys Lys Lys Trp Cys Glu Lys Asp His Arg Phe Arg Ile Glu Pro
50 55 60

Gln Leu Lys Ser Cys Leu Gly Tyr Ala Arg Asn Val Gly Ile Asp Ala
65 70 75 80

Ala Lys Gly Glu Tyr Ile Ala Phe Cys Asp Ser Asp Asp Cys Ile Thr
85 90 95

Ser Asp Phe Leu Ser Cys Phe Val Asp Thr Ala Leu Lys Asn Ser Ser

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

100

105

110

Asp Ile Val Glu Thr Gln Phe Thr Leu Cys Asp Gln Asn Leu Ser Pro
115 120 125

Ile Tyr Asp Tyr Asp Arg Asn Ile Leu Gly His Ile Leu Gly His Gly
130 135 140

Phe Leu Glu Tyr Thr Ser Ala Pro Ser Val Trp Lys Tyr Phe Val Lys
145 150 155 160

Arg Asp Ile Phe Thr Ser Asn Asn Leu His Tyr Pro Glu Ile Arg Phe
165 170 175

Gly Glu Asp Ile Ser Met Tyr Ser Leu Leu Phe Ser Tyr Cys Asn Lys
180 185 190

Ile Asp Tyr Val Glu Lys Pro Thr Tyr Leu Tyr Arg Gln Val Pro Ser
195 200 205

Ser Leu Met Asn Asn Pro Gln Gly Lys Arg Lys Arg Tyr Glu Ser Leu
210 215 220

Phe Asp Ile Ile Asp Phe Val Thr Asn Glu Phe Lys Thr Arg Leu Leu
225 230 235 240

Phe Gln Lys Ser Trp Leu Lys Leu Leu Phe Gln Leu Glu Met His Ser
245 250 255

Ala Ser Ile Ile Ser Asp Ser Ala Thr Ser Asp Asp Glu Ala Ile Ser
260 265 270

Met Arg Gln Glu Ile Ser Gly Tyr Leu Lys Lys Val Phe Pro Val Lys
275 280 285

Asn Thr Ile Phe Glu Val Thr Ala Leu Gly Trp Gly Gly Glu Ile Val
290 295 300

Ser Ser Ile Ala Ser Lys Phe Asn Thr Leu His Gly Val Ser Ser Ser

28 May 2018
2015259088

37847-517 Sequence Listing_ST25 (003)

305					310						315					320
Asn	Met	Phe	Asn	Arg	Tyr	Phe	Phe	Glu	Leu	Leu	Glu	Asp	Ser	Thr	Arg	
				325					330					335		
Lys	Lys	Leu	Glu	Glu	Met	Ile	Ile	Asn	Phe	Ser	Pro	Asp	Ile	Phe	Leu	
			340					345					350			
Ile	Asp	Leu	Ile	Ser	Glu	Ala	Asp	Tyr	Leu	Ser	Ser	Tyr	Lys	Gly	Asn	
		355					360					365				
Leu	Gly	Thr	Phe	Val	Lys	Asn	Trp	Lys	Ile	Gly	Phe	Ser	Ile	Phe	Met	
	370					375					380					
Lys	Met	Ile	Gln	Thr	His	Ser	Asn	Asn	Ser	Ser	Ile	Phe	Leu	Leu	Glu	
385					390					395					400	
Asn	Tyr	Met	Gln	Gln	Ala	Pro	Asp	His	Val	Asp	Asn	Thr	Asn	Glu	Ile	
			405						410					415		
Leu	Lys	Met	Leu	Tyr	Asp	Asp	Ile	Lys	Ile	Asn	His	Pro	Asp	Ile	Ile	
			420					425					430			
Cys	Ile	Ser	Pro	Ala	Pro	Asp	Ile	Leu	Asn	Arg	Ser	Ser	Glu	Pro	Glu	
		435					440					445				
Leu	Pro	Cys	Ile	Tyr	Gln	Leu	Lys	Leu	Val	Ser	Asp	Lys	Leu	His	Thr	
	450					455					460					
Met	Tyr	Ser	Pro	Val	Ile	Asn	Cys	Val	Glu	Thr	Lys	Gly	Gly	Leu	Gly	
465					470					475					480	
Asn	Gln	Ile	Phe	Gln	Tyr	Val	Phe	Ser	Lys	Tyr	Ile	Glu	Lys	Met	Thr	
			485						490					495		
Gly	Tyr	Arg	Pro	Leu	Leu	His	Ile	Gly	Phe	Phe	Asp	Tyr	Val	Lys	Ala	
			500					505					510			
Ile	Pro	Gly	Gly	Thr	Lys	Arg	Ile	Phe	Ser	Leu	Asp	Lys	Leu	Phe	Pro	

37847-517 Sequence Listing_ST25 (003)

515	520	525
Asp Ile Glu Thr Thr Ser Gly Lys Ile Pro Cys Ser His Val Val Glu 530 535 540		
Glu Lys Ser Phe Ile Ser Asn Pro Gly Ser Asp Ile Phe Tyr Arg Gly 545 550 555 560		
Tyr Trp Gln Asp Ile Arg Tyr Phe Ser Asp Val Lys Asp Glu Val Leu 565 570 575		
Glu Ser Phe Asn Val Asp Thr Ser Ser Met Ser Lys Asp Val Ile Asp 580 585 590		
Phe Ala Asp Thr Ile Arg Asn Ala Asn Ser Ile Ala Met His Ile Arg 595 600 605		
Arg Gln Asp Tyr Leu Asn Glu Asn Asn Val Ser Leu Phe Glu Gln Leu 610 615 620		
Ser Ile Asp Tyr Tyr Lys Ser Ala Val Asp Met Ile Arg Lys Glu Tyr 625 630 635 640		
Ala Asp Asp Leu Val Leu Phe Ile Phe Ser Asp Asp Pro Glu Tyr Ala 645 650 655		
Asn Ser Ile Ala Asp Ser Phe Asp Ile Glu Gly Phe Val Met Pro Leu 660 665 670		
His Lys Asp Tyr Glu Asp Leu Tyr Leu Ile Thr Leu Ala His His His 675 680 685		
Ile Ile Ala Asn Ser Thr Phe Ser Leu Trp Gly Ala Leu Leu Ser Ala 690 695 700		
Arg Lys Asp Gly Ile Arg Ile Ala Pro Arg Asn Trp Phe Lys Gly Thr 705 710 715 720		
Pro Ala Thr Asn Leu Tyr Pro Asp Lys Trp Leu Ile Leu		

37847-517 Sequence Listing_ST25 (003)

725

730

<210> 161
<211> 286
<212> PRT
<213> Anaeromusa acidaminophila

<400> 161

Met Phe Cys Val Arg Ile Tyr Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Leu Gly Arg Ala Met Ala Lys His Tyr Ser Glu Thr Ala Ala
20 25 30

Phe Asp Leu Ser Trp Tyr Glu Gln Lys Ile Lys Pro Gly Phe Glu Ala
35 40 45

Ser Val Cys Gln Tyr Asn Ile Glu Leu Ser Arg Lys Asp Arg Pro Lys
50 55 60

Ala Trp Tyr Glu Pro Ile Leu Lys Arg Ile Ser Arg His Thr Asp Lys
65 70 75 80

Leu Glu Met Trp Phe Gly Leu Phe Phe Glu Lys Lys Tyr His Tyr Asp
85 90 95

Ser Thr Val Phe Glu Arg Gly Leu Cys Lys Lys Asn Ile Thr Leu Asp
100 105 110

Gly Tyr Trp Gln Ser Tyr Lys Tyr Phe Ser Ala Ile Glu Asp Asp Leu
115 120 125

Arg Arg Glu Leu Thr Ile Pro Lys Glu Arg Glu Glu Leu Ile Ala Ile
130 135 140

Ser Arg Ser Leu Pro Glu Asn Ser Val Ser Ile His Val Arg Arg Gly
145 150 155 160

Asp Tyr Val Ser Asn Pro Lys Ala Asn Ala Met His Gly Thr Cys Ser
165 170 175

37847-517 Sequence Listing_ST25 (003)

Trp Glu Tyr Tyr Gln Ala Ala Ile Glu Lys Met Thr Gly Leu Val Lys
180 185 190

Glu Pro Gln Tyr Val Val Phe Ser Asp Asp Ile Thr Trp Thr Lys Glu
195 200 205

Asn Leu Pro Leu Pro Asn Ala Met Tyr Ile Gly Arg Glu Leu Gly Leu
210 215 220

Phe Asp Tyr Glu Glu Leu Ile Leu Met Ser Arg Cys Lys His Asn Ile
225 230 235 240

Met Ala Asn Ser Thr Phe Ser Trp Trp Gly Ala Trp Leu Asn Ser Asn
245 250 255

Pro Asn Lys Val Val Ile Ala Pro Arg Lys Trp Phe Arg His Lys Lys
260 265 270

Ile Lys Val Asn Asp Leu Phe Pro Ser Ser Trp Val Val Leu
275 280 285

<210> 162
<211> 292
<212> PRT
<213> Bacteroides sp. 2_1_16

<400> 162

Met Asp Ile Val Val Ile Phe Asn Gly Leu Gly Asn Gln Met Ser Gln
1 5 10 15

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

Ile Phe Asp Pro Lys Ser Thr Asn Val His Asn Gly Ala Glu Leu Lys
35 40 45

Arg Val Phe Gly Ile Glu Leu Asn Arg Asn Tyr Leu Asp Lys Ile Ile
50 55 60

37847-517 Sequence Listing_ST25 (003)

Ser Tyr Phe Tyr Gly Tyr Ile Phe Asn Lys Arg Ile Val Asn Lys Leu
65 70 75 80

Phe Ser Leu Val Gly Ile Arg Met Ile Tyr Glu Pro Lys Asn Tyr Asp
85 90 95

Tyr Arg Glu Glu Leu Leu Lys Pro Ser Ser Asn Phe Ile Ser Phe Tyr
100 105 110

Trp Gly Gly Trp His Ser Glu Lys Tyr Phe Lys Asp Ile Glu Leu Glu
115 120 125

Val Lys Lys Val Phe Lys Phe Pro Glu Val Thr Asn Ser Pro Tyr Phe
130 135 140

Thr Glu Trp Phe Asn Lys Ile Phe Leu Asp Asn Asn Ser Val Ser Ile
145 150 155 160

His Ile Arg Arg Gly Asp Tyr Leu Asp Lys Pro Ser Asp Pro Tyr Tyr
165 170 175

Gln Phe Asn Gly Val Cys Thr Ile Asp Tyr Tyr Glu Lys Ala Ile Leu
180 185 190

Tyr Leu Lys Glu Arg Ile Leu Glu Pro Asn Phe Tyr Ile Phe Ser Asn
195 200 205

Asp Ile Asn Trp Cys Met Lys Thr Phe Gly Thr Glu Asn Met Tyr Tyr
210 215 220

Val Asp Cys Asn Lys Gly Lys Asp Ser Trp Arg Asp Met Tyr Leu Met
225 230 235 240

Ser Glu Cys Arg His His Ile Asn Ala Asn Ser Thr Phe Ser Trp Trp
245 250 255

Ala Ala Trp Leu Ser Pro Tyr Ser Asn Gly Ile Val Leu His Pro Lys
260 265 270

37847-517 Sequence Listing_ST25 (003)

Tyr Phe Ile Lys Asp Ile Glu Thr Lys Asp Tyr Tyr Pro Gln Lys Trp
 275 280 285

Ile Met Ile Glu
 290

<210> 163

<211> 283

<212> PRT

<213> Chlorobium phaeobacteroides

<400> 163

Met Asp Lys Val Val Val His Leu Thr Gly Gly Leu Gly Asn Gln Met
 1 5 10 15

Phe Gln Tyr Ala Leu Gly Arg Ser Ile Ser Ile Asn Arg Asn Cys Pro
 20 25 30

Leu Leu Leu Asn Thr Ser Phe Tyr Asp Thr Tyr Asp Lys Phe Ser Cys
 35 40 45

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

Tyr Tyr Asn Asn Lys Tyr Tyr Arg Tyr Val Ile Arg Leu Leu Ser Arg
 65 70 75 80

Tyr Gly Val Ala Cys Tyr Phe Gly Ser Tyr Tyr Glu Lys Lys Ile Phe
 85 90 95

Ser Tyr Asp Glu Lys Val Tyr Lys Arg Ser Cys Val Ser Tyr Tyr Gly
 100 105 110

Thr Trp Gln Ser Tyr Gly Tyr Phe Asp Ser Ile Arg Asp Ile Leu Leu
 115 120 125

Arg Asp Tyr Glu Met Val Gly Cys Leu Glu Glu Glu Val Glu Lys Tyr
 130 135 140

37847-517 Sequence Listing_ST25 (003)

Val Ser Asp Ile Lys Arg Val Asp Ser Val Ser Leu His Ile Arg Arg
 145 150 155 160

Gly Asp Tyr Phe Asp Asn Lys Arg Leu Gln Ser Ile His Gly Ile Leu
 165 170 175

Thr Met Glu Tyr Tyr Tyr Lys Ala Met Ser Leu Phe Pro Asp Ser Ser
 180 185 190

Val Phe Tyr Val Phe Ser Asp Asp Ile Glu Trp Val Arg Glu Asn Leu
 195 200 205

Ile Thr Asn Thr Asn Ile Val Tyr Val Val Leu Glu Ser Asp Asn Pro
 210 215 220

Glu Asn Glu Ile Tyr Leu Met Ser Leu Cys Lys Asn Asn Ile Ile Ser
 225 230 235 240

Asn Ser Thr Phe Ser Trp Trp Gly Ala Trp Leu Asn Lys Asn Lys Tyr
 245 250 255

Lys Lys Val Ile Ala Pro Arg Met Trp Tyr Lys Asp Asn Gln Ser Ser
 260 265 270

Ser Asp Leu Met Pro Ser Asp Trp Cys Leu Ile
 275 280

<210> 164
 <211> 297
 <212> PRT
 <213> Treponema bryantii

<400> 164

Met Ile Val Ile Ser Met Gly Gly Gly Leu Gly Asn Gln Met Phe Glu
 1 5 10 15

Tyr Ala Phe Tyr Thr Gln Leu Lys His Leu Tyr Pro Lys Ser Glu Ile
 20 25 30

Lys Val Asp Thr Lys Tyr Ala Phe Pro Tyr Ser His Asn Gly Ile Glu

2015259088 28 May 2018

40

45

Val His Ser Leu Val Lys Thr Tyr Pro Ile Glu Gly Asn Lys Ala His
65 70 75 80

Phe Ile Lys Phe Phe Leu Tyr Arg Ile Leu Arg Lys Ala Asn Leu Val
85 90 95

Glu Arg Glu Pro Thr Ser Phe Cys Lys Gln Lys Asp Phe Thr Glu Phe
100 105 110

Tyr Asn Ser Phe Phe Glu Leu Pro Gln Asn Lys Ser Phe Tyr Leu Tyr
115 120 125

Gly Pro Phe Val Asn Tyr Asn Tyr Phe Ala Ala Ile His Asn Glu Ile
130 135 140

Met Asp Leu Tyr Thr Phe Pro Glu Ile Thr Asp Val Thr Asn Ile Glu
145 150 155 160

Tyr Lys Arg Lys Ile Glu Ser Ser His Ser Ile Ser Ile His Ile Arg
165 170 175

Arg Gly Asp Tyr Ile Thr Glu Gly Val Pro Leu Val Pro Asp Ala Tyr
180 185 190

Tyr Arg Glu Ala Leu Val Tyr Ile Asn Lys Lys Ile Glu Asp Pro His
195 200 205

Phe Phe Val Phe Thr Asp Asp Lys Asp Tyr Cys Lys Ser Leu Phe Ser
210 215 220

Asp Asn Gln Asn Phe Thr Ile Val Glu Gly Asn Thr Gly Ala Asn Ser
225 230 235 240

Phe Arg Asp Met Gln Leu Met Ser Leu Cys Lys His Asn Ile Ile Ala

37847-517 Sequence Listing_ST25 (003)

245 250 255

Asn Ser Thr Phe Ser Phe Trp Gly Ala Phe Leu Asn Lys Asn Ser Glu
260 265 270

Lys Ile Val Ile Ala Pro Asn Ile Ala Phe Lys Asp Cys Ser Cys Pro
275 280 285

Tyr Ile Cys Pro Asp Trp Ile Ile Leu
290 295

<210> 165

<211> 290

<212> PRT

<213> Bacteroides fragilis

<400> 165

Met Val Ile Ala Lys Leu Phe Gly Gly Leu Gly Asn Gln Met Phe Ile
1 5 10 15

Tyr Ala Ala Ala Lys Gly Ile Ala Gln Ile Ser Asn Gln Lys Leu Thr
20 25 30

Phe Asp Ile Tyr Thr Gly Phe Glu Asp Asp Ser Arg Phe Arg Arg Val
35 40 45

Tyr Glu Leu Lys Gln Phe Asn Leu Ser Val Gln Glu Ser Arg Arg Trp
50 55 60

Met Ser Phe Arg Tyr Pro Leu Gly Arg Ile Leu Arg Lys Ile Ser Arg
65 70 75 80

Lys Ile Gly Phe Cys Ile Pro Leu Val Asn Phe Lys Phe Ile Val Glu
85 90 95

Lys Lys Pro Tyr His Phe Gln Asn Glu Ile Met Arg Ile Ala Ser Phe
100 105 110

Ser Ser Ile Tyr Leu Glu Gly Tyr Phe Gln Ser Tyr Lys Tyr Phe Ser
115 120 125

37847-517 Sequence Listing_ST25 (003)

Lys Ile Glu Ala Gln Ile Arg Glu Asp Phe Lys Phe Thr Lys Glu Val
130 135 140

Ile Gly Ser Val Glu Lys Glu Ala Ser Phe Ile Thr Asn Ser Arg Tyr
145 150 155 160

Thr Pro Val Ala Ile Gly Val Arg Arg Tyr Ser Glu Met Lys Gly Glu
165 170 175

Phe Gly Glu Leu Ala Val Val Glu His Asp Tyr Tyr Asp Ala Ala Ile
180 185 190

Lys Tyr Ile Ala Asn Lys Val Pro Asn Leu Ile Phe Ile Val Phe Ser
195 200 205

Glu Asp Ile Asp Trp Val Lys Lys Asn Leu Lys Leu Asp Tyr Pro Val
210 215 220

Tyr Phe Val Thr Ser Lys Lys Gly Glu Leu Ala Ala Ile Gln Asp Met
225 230 235 240

Tyr Leu Met Ser Leu Cys Asn His His Ile Ile Ser Asn Ser Ser Phe
245 250 255

Tyr Trp Trp Gly Ala Tyr Leu Ala Ser Thr Asn Asn His Ile Val Ile
260 265 270

Ala Pro Ser Val Phe Leu Asn Lys Asp Cys Thr Pro Ile Asp Trp Val
275 280 285

Ile Ile
290

<210> 166

<211> 314

<212> PRT

<213> Firmicutes bacterium CAG:534

<400> 166

37847-517 Sequence Listing_ST25 (003)

Met Ser Gly Gly Leu Gly Asn Gln Met Phe Gln Tyr Ala Leu Tyr Met
 1 5 10 15
 Lys Leu Thr Ala Met Gly Arg Glu Val Lys Phe Asp Asp Ile Asn Glu
 20 25 30
 Tyr Arg Gly Glu Lys Ala Trp Pro Ile Met Leu Ala Val Phe Gly Ile
 35 40 45
 Glu Tyr Pro Arg Ala Thr Trp Asp Glu Ile Val Ala Phe Thr Asp Gly
 50 55 60
 Ser Met Asp Phe Ser Lys Arg Leu Lys Arg Leu Phe Arg Gly Arg His
 65 70 75 80
 Pro Ile Glu Tyr Val Glu Gln Gly Phe Tyr Asp Pro Lys Val Leu Ser
 85 90 95
 Phe Glu Asn Met Tyr Leu Lys Gly Ser Phe Gln Ser Gln Arg Tyr Phe
 100 105 110
 Glu Asp Ile Leu Glu Glu Val Gln Glu Thr Phe Arg Phe Pro Glu Leu
 115 120 125
 Lys Asp Met Asn Leu Pro Ala Pro Leu Tyr Glu Thr Thr Glu Lys Tyr
 130 135 140
 Leu Leu Arg Ile Glu Gly Cys Asn Ala Val Gly Leu His Met Tyr Arg
 145 150 155 160
 Gly Asp Ser Arg Ser Asn Glu Glu Leu Tyr Asp Gly Ile Cys Thr Glu
 165 170 175
 Lys Tyr Tyr Glu Gly Ala Val Arg Phe Ile Gln Asp Lys Cys Pro Asp
 180 185 190
 Ala Lys Phe Phe Ile Phe Ser Asn Glu Pro Lys Trp Val Lys Gly Trp
 195 200 205

37847-517 Sequence Listing_ST25 (003)

Val Ile Ser Leu Met Lys Ser Gln Ile Arg Glu Asp Met Ser Arg Glu
210 215 220

Glu Ile Arg Ala Leu Glu Asp His Phe Val Leu Ile Glu Asn Asn Thr
225 230 235 240

Glu Tyr Thr Gly Tyr Leu Asp Met Phe Leu Met Ser Arg Cys Arg His
245 250 255

Asn Ile Ile Ser Asn Ser Ser Phe Ser Trp Trp Ala Ala Phe Ile Asn
260 265 270

Glu Asn Pro Asp Lys Leu Val Thr Ala Pro Ser Arg Trp Val Asn Gly
275 280 285

Val Pro Ser Glu Asp Val Tyr Val Lys Gly Met Thr Leu Ile Asp Glu
290 295 300

Lys Gly Arg Val Glu Arg Thr Ile Lys Glu
305 310

<210> 167

<211> 279

<212> PRT

<213> Firmicutes bacterium CAG:882

<400> 167

Met Val Ile Val Lys Ile Gly Asp Gly Leu Gly Asn Gln Met Phe Asn
1 5 10 15

Tyr Val Cys Gly Tyr Ser Val Ala Lys His Asp Asn Asp Thr Leu Leu
20 25 30

Leu Asp Thr Ser Asp Val Asp Asn Ser Thr Leu Arg Thr Tyr Asp Leu
35 40 45

Asp Lys Phe Asn Ile Asp Phe Thr Asp Arg Glu Ser Phe Thr Asn Lys
50 55 60

37847-517 Sequence Listing_ST25 (003)

Gly Phe Phe His Lys Val Tyr Lys Arg Leu Arg Arg Ser Leu Lys Tyr
65 70 75 80

Asn Val Ile Tyr Glu Ser Arg Thr Glu Asn Cys Pro Cys Val Leu Asp
85 90 95

Val Tyr Arg Arg Lys Phe Ile Arg Asp Lys Tyr Leu His Gly Tyr Phe
100 105 110

Gln Asn Leu Cys Tyr Phe Lys Thr Cys Lys Glu Asp Ile Met Arg Gln
115 120 125

Phe Thr Pro Lys Glu Pro Phe Ser Ala Lys Ala Asp Glu Leu Ile His
130 135 140

Arg Phe Ala Thr Glu Asn Thr Cys Ser Val His Val Arg Gly Gly Asp
145 150 155 160

Ile Lys Pro Leu Ser Ile Lys Tyr Tyr Lys Asp Ala Leu Asp Lys Ile
165 170 175

Gly Glu Ala Lys Lys Asp Met Arg Phe Ile Val Phe Ser Asn Val Arg
180 185 190

Asn Leu Ala Glu Glu Tyr Ile Lys Glu Leu Gly Val Asp Ala Glu Phe
195 200 205

Ile Trp Asp Leu Gly Glu Phe Thr Asp Ile Glu Glu Leu Phe Leu Met
210 215 220

Lys Ala Cys Arg Arg His Ile Leu Ser Asp Ser Thr Phe Ser Arg Trp
225 230 235 240

Ala Ala Leu Leu Asp Glu Lys Ser Glu Glu Val Phe Val Pro Phe Ser
245 250 255

Pro Asp Ala Asp Lys Ile Tyr Met Pro Glu Trp Ile Met Glu Glu Tyr
260 265 270

37847-517 Sequence Listing_ST25 (003)

Asp Gly Asn Glu Glu Lys Arg
275

<210> 168

<211> 287

<212> PRT

<213> *Vibrio parahaemolyticus*

<400> 168

Met Val Ile Val Lys Val Ser Gly Gly Leu Gly Asn Gln Leu Phe Gln
1 5 10 15

Tyr Ala Ile Gly Cys Ala Ile Ser Asn Arg Leu Ser Cys Glu Leu Leu
20 25 30

Leu Asp Thr Ser Phe Tyr Pro Lys Gln Ser Leu Arg Lys Tyr Glu Leu
35 40 45

Asp Lys Phe Asn Ile Lys Ala Lys Val Ala Thr Gln Lys Glu Val Phe
50 55 60

Ser Cys Gly Gly Gly Asp Asp Leu Leu Ser Arg Phe Leu Arg Lys Leu
65 70 75 80

Asn Leu Ser Ser Leu Phe Phe Pro Asn Tyr Ile Lys Glu Lys Glu Ser
85 90 95

Leu Val Tyr Leu Ala Glu Ile Ser His Cys Lys Ser Gly Ser Phe Leu
100 105 110

Asp Gly Tyr Trp Gln Asn Pro Gln Tyr Phe Ser Asp Ile Lys Asp Glu
115 120 125

Leu Val Lys Gln Ile Val Pro Ile Met Pro Leu Ser Ser Pro Ala Leu
130 135 140

Glu Trp Gln Asn Ile Ile Ile Asn Thr Lys Asn Cys Val Ser Leu His
145 150 155 160

Val Arg Arg Gly Asp Tyr Val Asn Asn Ala His Thr Asn Ser Val His

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

165 170 175

Gly Val Cys Asp Leu Ser Tyr Tyr Arg Glu Ala Ile Thr Asn Ile His
180 185 190

Glu Thr Val Glu Lys Pro Lys Phe Phe Val Phe Ser Asp Asp Ile Ser
195 200 205

Trp Cys Lys Asp Asn Leu Gly Ser Leu Gly His Phe Thr Tyr Val Asp
210 215 220

Asn Thr Leu Ser Ala Ile Asp Asp Leu Met Leu Met Ser Phe Cys Glu
225 230 235 240

His His Ile Ile Ala Asn Ser Thr Phe Ser Trp Trp Gly Ala Trp Leu
245 250 255

Asn Asp His Gly Ile Thr Ile Ala Pro Lys Arg Trp Phe Ser Ser Val
260 265 270

Glu Arg Asn Asn Lys Asp Leu Phe Pro Glu Lys Trp Leu Ile Leu
275 280 285

<210> 169
<211> 286
<212> PRT
<213> Herbaspirillum frisingense

<400> 169

Met Ile Val Ser Arg Leu Ile Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Ala Gly Arg Ala Leu Ala Leu Arg Arg Gly Val Pro Phe Ala
20 25 30

Ile Asp Ser Arg Ala Phe Ala Asp Tyr Lys Thr His Ala Phe Gly Met
35 40 45

Gln Cys Phe Cys Ala Asp Gln Thr Glu Ala Pro Ser Arg Leu Leu Pro
50 55 60

37847-517 Sequence Listing_ST25 (003)

Asn Pro Pro Ala Glu Gly Arg Leu Gln Arg Leu Leu Arg Arg Phe Leu
65 70 75 80

Pro Asn Pro Leu Arg Val Tyr Thr Glu Lys Thr Phe Thr Phe Asp Glu
85 90 95

Ala Val Leu Ser Leu Pro Asp Gly Ile Tyr Leu Asp Gly Tyr Trp Gln
100 105 110

Ser Glu Lys Tyr Phe Ala Asp Phe Ala Asp Asp Ile Arg Lys Asp Phe
115 120 125

Ala Val Lys Ala Ala Pro Ser Ala Pro Asn Gln Ala Trp Leu Glu Leu
130 135 140

Ile Gly Arg Thr His Ser Val Ser Leu His Ile Arg Arg Gly Asp Tyr
145 150 155 160

Val Ser Asn Ala Ala Ala Ala Val His Gly Thr Cys Asp Leu Gly
165 170 175

Tyr Tyr Glu Arg Ala Val Ala His Leu His Gln Val Thr Gly Gln Ala
180 185 190

Pro Glu Leu Phe Val Phe Ser Asp Asp Leu Asp Trp Val Ala Thr Asn
195 200 205

Leu Gln Leu Pro Tyr Thr Met His Leu Val Arg Asp Asn Asp Ala Ala
210 215 220

Thr Asn Phe Glu Asp Leu Arg Leu Met Thr Ala Cys Arg His His Ile
225 230 235 240

Val Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala Trp Leu Asp Gly Arg
245 250 255

Ser Glu Ser Ile Thr Ile Ala Pro Ala Arg Trp Phe Val Ala Asp Thr
260 265 270

37847-517 Sequence Listing_ST25 (003)

Pro Asp Ala Arg Asp Leu Val Pro Gln Arg Trp Val Arg Leu
 275 280 285

<210> 170
 <211> 288
 <212> PRT
 <213> Rhizobium sp. CF080

<400> 170

Met Ile Ile Thr Arg Ile Leu Gly Gly Leu Gly Asn Gln Met Phe Gln
 1 5 10 15

Tyr Ala Ala Gly Arg Ala Leu Ala Ile Ala Asn Glu Ala Glu Leu Lys
 20 25 30

Leu Asp Leu Ile Glu Met Gly Ala Tyr Lys Leu Arg Pro Phe Ala Leu
 35 40 45

Asp Gln Phe Asn Ile Lys Ala Ala Ile Ala Gln Pro Asp Glu Val Pro
 50 55 60

Ala Lys Pro Lys Arg Gly Leu Leu Arg Lys Phe Thr Ser Ala Phe Lys
 65 70 75 80

Pro Asp Arg Ser Ser Cys Glu Arg Ile Val Glu Asn Gly Leu Thr Phe
 85 90 95

Asp Ser Arg Val Pro Ala Leu Arg Gly Ser Leu His Leu Ser Gly Tyr
 100 105 110

Trp Gln Ser Glu Gln Tyr Phe Ala Ser Ser Ala Asp Ala Ile Arg Ser
 115 120 125

Asp Phe Ser Leu Lys Ser Pro Leu Gly Pro Ala Arg Gln Asp Val Leu
 130 135 140

Ala Arg Ile Gly Ala Ala Thr Thr Pro Val Ser Ile His Val Arg Arg
 145 150 155 160

37847-517 Sequence Listing_ST25 (003)

Gly Asp Tyr Val Thr Asn Pro Ser Ala Asn Ala Val His Gly Thr Cys
165 170 175

Glu Pro Pro Trp Tyr His Glu Ala Met Arg Arg Met Leu Asp Arg Ala
180 185 190

Gly Asp Ala Ser Phe Phe Val Phe Ser Asp Glu Pro Gln Trp Ala Arg
195 200 205

Asp Asn Leu Gln Ser Ser Arg Pro Met Val Phe Ile Glu Pro Gln Asn
210 215 220

Asn Gly Arg Asp Gly Glu Asp Met His Leu Met Ala Ala Cys His Ala
225 230 235 240

His Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala Trp Leu Asn
245 250 255

Pro Arg Pro Asn Lys His Val Ile Ala Pro Arg Gln Trp Phe Arg Ala
260 265 270

Pro Asp Lys Asp Asp Arg Asp Ile Val Pro Ala Thr Trp Glu Arg Leu
275 280 285

<210> 171

<211> 328

<212> PRT

<213> Verrucomicrobium spinosum

<400> 171

Met Val Ile Ser His Ile Ser Glu Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Ala Gly Arg Arg Leu Ser Tyr His Leu Gly Thr Thr Leu Lys
20 25 30

Leu Asp Asp Tyr His Tyr Arg Leu His Pro Phe Arg Ser Phe Gln Leu
35 40 45

37847-517 Sequence Listing_ST25 (003)

Asp	Arg	Phe	Leu	Ile	Thr	Ser	Pro	Ile	Ala	Thr	Asp	Ala	Glu	Ile	Ser	50	55	60	
His	Leu	Cys	Pro	Leu	Glu	Gly	Leu	Ala	Arg	Ala	Ile	Arg	Ala	Arg	Leu	65	70	75	80
Pro	Gly	Lys	Leu	Arg	Gly	Ala	Thr	Leu	Arg	Leu	Leu	Gly	Asn	Leu	Gly	85	90	95	
Leu	Gly	Ser	Pro	Tyr	Gln	Pro	Arg	Leu	His	Ser	Phe	Lys	Glu	Glu	Thr	100	105	110	
Pro	Lys	Gln	Pro	Leu	Leu	Ile	Gly	Lys	Val	Val	Ser	Glu	Arg	His	Phe	115	120	125	
His	Phe	Asp	Pro	Asp	Val	Leu	Glu	Cys	Pro	Asp	Asn	Val	Cys	Leu	Val	130	135	140	
Gly	Tyr	Trp	Gln	Asp	Glu	Arg	Tyr	Phe	Gly	Glu	Ile	Arg	Asp	Ile	Leu	145	150	155	160
Leu	Arg	Glu	Leu	Thr	Leu	Lys	Ser	Pro	Pro	Ala	Gly	Ala	Thr	Lys	Ala	165	170	175	
Val	Leu	Glu	Arg	Ile	Gln	Arg	Ser	Ser	Ser	Val	Ser	Leu	His	Val	Arg	180	185	190	
Arg	Gly	Asp	Lys	Thr	Lys	Ser	Ser	Ser	Tyr	His	Cys	Thr	Ser	Leu	Glu	195	200	205	
Tyr	Cys	Leu	Ala	Ala	Met	Ser	Glu	Met	Arg	Ala	Arg	Leu	Gln	Ala	Pro	210	215	220	
Thr	Phe	Phe	Val	Phe	Ser	Asp	Asp	Trp	Asp	Trp	Val	Arg	Glu	Gln	Ile	225	230	235	240
Pro	Cys	Ser	Ser	Ser	Val	Ile	His	Val	Asp	His	Asn	Arg	Ala	Glu	Asp	245	250	255	

37847-517 Sequence Listing_ST25 (003)

Val Ser Glu Asp Phe Arg Leu Met Lys Ser Cys Asp His His Ile Ile
260 265 270

Ala Ser Ser Ser Leu Ser Trp Trp Ala Ala Trp Leu Gly Thr Asn Glu
275 280 285

Asn Ser Phe Val Phe Ser Pro Pro Ala Asp Arg Trp Leu Asn Phe Ser
290 295 300

Asn His Phe Thr Ala Asp Val Leu Pro Pro His Trp Ile Gln Leu Asp
305 310 315 320

Gly Ser Ser Leu Leu Pro Ala Gln
325

<210> 172

<211> 300

<212> PRT

<213> Fibrella aestuarina

<400> 172

Met Thr Ala Asn Arg Val Leu Val Asn Ser Pro Met Val Ile Ala Lys
1 5 10 15

Ile Thr Ser Gly Leu Gly Asn Gln Leu Phe Gln Tyr Ala Leu Gly Arg
20 25 30

His Leu Ala Leu Gln Gly Asn Thr Ser Leu Trp Phe Asp Leu Arg Tyr
35 40 45

Phe His Gln Glu Tyr Ala Thr Asp Thr Pro Arg Lys Phe Lys Leu Asp
50 55 60

Arg Phe Asn Val Arg Tyr Asn Leu Leu Asp Ser Ser Pro Trp Leu Tyr
65 70 75 80

Ala Ser Lys Ala Thr Arg Leu Leu Pro Gly Arg Ser Leu Arg Pro Leu
85 90 95

Ile Asp Thr Arg Phe Glu Ala Asp Phe His Phe Asp Pro Thr Val Ile

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

100

105

110

Arg Pro Ala Ala Pro Leu Thr Ile Leu Trp Gly Phe Trp Gln Ser Glu
115 120 125

Lys Tyr Phe Ala Gln Ser Thr Pro Gln Ile Arg Gln Glu Leu Thr Phe
130 135 140

Asn Arg Pro Leu Ser Asp Thr Phe Val Gly Tyr Gln Gln Gln Ile Glu
145 150 155 160

Gln Ala Glu Val Pro Ile Ser Val His Val Arg Arg Gly Asp Tyr Val
165 170 175

Thr His Pro Glu Phe Ser Gln Ser Phe Gly Phe Val Gly Leu Ala Tyr
180 185 190

Tyr Gln Lys Ala Leu Ala His Leu Gln Asp Leu Phe Pro Asn Ala Thr
195 200 205

Leu Phe Phe Phe Ser Asp Asp Pro Asp Trp Val Arg Ala Asn Ile Val
210 215 220

Thr Glu Gln Pro His Val Phe Val Gln Asn Ser Gly Pro Asp Ala Asp
225 230 235 240

Val Asp Asp Leu Gln Leu Met Ser Leu Cys His His His Val Ile Ala
245 250 255

Asn Ser Ser Phe Ser Trp Trp Gly Ala Trp Leu Asn Pro Arg Pro Asp
260 265 270

Lys Val Val Ile Gly Pro Gln Arg Trp Phe Ala Asn Lys Pro Trp Asp
275 280 285

Thr Lys Asp Leu Leu Pro Ser Gly Trp Leu Arg Leu
290 295 300

<210> 173

37847-517 Sequence Listing_ST25 (003)

<211> 287
 <212> PRT
 <213> Rhodobacter sp. CACIA14H1

<400> 173

Met Ile His Met Arg Leu Val Gly Gly Leu Gly Asn Gln Leu Phe Gln
 1 5 10 15

Tyr Ala Cys Gly Arg Ala Val Ala Leu Arg His Gly Thr Glu Leu Val
 20 25 30

Leu Asp Thr Arg Glu Leu Ser Arg Gly Ala Ala His Ala Val Phe Gly
 35 40 45

Leu Asp His Phe Ala Ile Arg Ala Arg Met Gly Ala Ser Ala Asp Leu
 50 55 60

Pro Pro Pro Arg Ser Arg Val Leu Ala Tyr Gly Leu Trp Arg Ala Gly
 65 70 75 80

Phe Met Ala Pro Arg Phe Leu Arg Glu Arg Gly Leu Gly Val Asn Pro
 85 90 95

Ala Val Leu Ala Ala Gly Asp Gly Thr Tyr Leu His Gly Tyr Phe Gln
 100 105 110

Ser Glu Ala Tyr Phe Arg Asp Val Val Pro Gln Ile Arg Pro Glu Leu
 115 120 125

Glu Ile Val Thr Pro Pro Ser Asp Asp Asn Leu Arg Trp Ala Ser Arg
 130 135 140

Ile Ala Gly Asp Asp Arg Ala Val Ser Leu His Val Arg Arg Gly Asp
 145 150 155 160

Tyr Val Ala Ser Ala Lys Gly Gln Gln Val His Gly Thr Cys Asp Ala
 165 170 175

Asp Tyr Tyr Ala Arg Ala Val Ala Ala Ile Arg Ala Arg Ala Gly Ile
 180 185 190

37847-517 Sequence Listing_ST25 (003)

Asp Pro Arg Leu Tyr Val Phe Ser Asp Asp Pro His Trp Ala Arg Asp
195 200 205

Asn Leu Ala Leu Asp Ala Glu Thr Val Val Leu Asp His Asn Pro Pro
210 215 220

Gly Ala Ala Val Glu Asp Met Arg Leu Met Gly Val Cys Arg His His
225 230 235 240

Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala Trp Arg Asn Pro
245 250 255

Ser Ala Gly Lys Val Val Val Ala Pro Val Arg Trp Phe Ala Asp Pro
260 265 270

Lys Leu His Asn Pro Asp Ile Cys Pro Pro Glu Trp Leu Arg Val
275 280 285

<210> 174
<211> 316
<212> PRT
<213> Rhodopirellula baltica

<400> 174

Met Ala Thr Ser Ala His Leu His Leu Ser Asp Glu Lys Gln Thr Leu
1 5 10 15

Asp Ser Lys Ala Ser Asp Arg Asp Cys Ala Thr Thr Glu Ala Ser Ala
20 25 30

Ser Asp Lys Thr Cys Thr Ile Ser Ile Ser Gly Gly Leu Gly Asn Gln
35 40 45

Met Leu Gln Tyr Ala Ala Gly Arg Ala Leu Ser Ile His His Asp Cys
50 55 60

Ser Leu Gln Leu Asp Leu Lys Phe Tyr Ser Ser Lys Arg His Arg Ser
65 70 75 80

37847-517 Sequence Listing_ST25 (003)

Tyr Glu Leu Asp Ala Phe Pro Ile Gln Ala His Arg Ser Ile Lys Pro
 85 90 95
 Ser Phe Phe Ser Gln Ile Leu Ser Lys Ile Gln Ser Glu Ser Lys His
 100 105 110
 Val Pro Thr Tyr Gln Glu Gln Ser Lys Arg Phe Asp Pro Ala Phe Phe
 115 120 125
 Asn Thr Glu Pro Pro Val Lys Ile Arg Gly Tyr Phe Phe Ser Glu Lys
 130 135 140
 Tyr Phe Ser Pro Tyr Ala Asp Gln Ile Arg Thr Glu Leu Thr Pro Pro
 145 150 155 160
 Ile Pro Pro Asp Gln Pro Ala Arg Asp Met Ala Ile Arg Leu Lys Glu
 165 170 175
 Cys Val Ser Thr Ser Leu His Val Arg Arg Gly Asp Tyr Val Thr Asn
 180 185 190
 Ala Asn Ala Arg Gln Arg Phe Trp Cys Cys Thr Ser Glu Tyr Phe Glu
 195 200 205
 Ala Ala Ile Glu Arg Leu Pro Thr Asp Ser Thr Val Phe Val Phe Ser
 210 215 220
 Asp Asp Ile Glu Trp Ala Lys Gln Asn Ile Arg Ser Ser Arg Thr Thr
 225 230 235 240
 Val Tyr Val Asn Asp Glu Leu Lys Lys Ala Gly Ser Pro Glu Thr Gly
 245 250 255
 Leu Arg Asp Leu Trp Leu Met Thr His Ala Lys Ser His Ile Ile Ala
 260 265 270
 Asn Ser Ser Phe Ser Trp Trp Gly Ala Trp Leu Ala Asn Ser Glu Ala
 275 280 285

37847-517 Sequence Listing_ST25 (003)

Asn Leu Thr Ile Ala Pro Lys Lys Trp Phe Asn Asp Pro Glu Ile Asp
290 295 300

Asp Ser Asp Ile Val Pro Ser Ser Trp His Arg Ile
305 310 315

<210> 175

<211> 286

<212> PRT

<213> Spirosoma spitsbergense

<400> 175

Met Val Val Val Glu Leu Met Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Phe Gly Met Gln Leu Ala His Gln Arg Gln Asp Thr Leu Thr
20 25 30

Val Ser Thr Phe Leu Leu Ser Asn Lys Leu Leu Ala Asn Leu Arg Asn
35 40 45

Tyr Thr Tyr Arg Pro Phe Glu Leu Cys Ile Phe Gly Ile Asp Lys Pro
50 55 60

Lys Ala Ser Pro Phe Asn Leu Leu Arg Ala Leu Leu Pro Phe Asp Leu
65 70 75 80

Asn Thr Ser Leu Leu Arg Glu Thr Asp Asp Pro Glu Ala Val Ile Pro
85 90 95

Ala Ala Ser Ala Arg Ile Val Cys Val Gly Tyr Trp Gln Ser Glu His
100 105 110

Tyr Phe Glu Glu Val Thr Val His Val Arg Glu Lys Phe Ile Phe Arg
115 120 125

Gln Pro Phe Asn Ser Phe Thr Ser Arg Leu Ala Asn Asn Leu Asn Gly
130 135 140

37847-517 Sequence Listing_ST25 (003)

Ile Pro Asn Ser Val Phe Val His Ile Arg Arg Gly Asp Tyr Val Thr
145 150 155 160

Asn Lys Gly Ala Asn Ala His His Gly Leu Cys Asp Arg Thr Tyr Tyr
165 170 175

Glu Arg Ala Val Thr Phe Met Arg Glu His Leu Glu Asn Pro Leu Phe
180 185 190

Phe Ile Phe Ser Asp Asp Leu Glu Trp Val Ser Gln Glu Leu Gly Pro
195 200 205

Ile Leu Glu Pro Ala Thr Tyr Val Gly Gly Asn Gln Lys Asn Asp Ser
210 215 220

Trp Gln Asp Met Tyr Leu Met Ser Leu Cys Arg His Ala Ile Val Ala
225 230 235 240

Asn Ser Ser Phe Ser Trp Trp Gly Ala Trp Leu Ser Pro His Ala Ser
245 250 255

Lys Ile Val Val Ala Pro Lys Glu Trp Phe Gly Lys Pro Leu Leu Pro
260 265 270

Val Lys Thr Asn Asp Leu Ile Pro Asn Ser Trp Ile Arg Ile
275 280 285

<210> 176
<211> 295
<212> PRT
<213> Unknown

<220>
<223> uncultured bacterium

<400> 176

Met Asn Ala Ile Ile Pro Arg Leu Thr Gly Gly Ile Gly Asn Gln Leu
1 5 10 15

Phe Ile Tyr Ala Ala Ala Arg Arg Met Ala Ile Ala Asn Ser Met Asn
20 25 30

37847-517 Sequence Listing_ST25 (003)

Leu Val Ile Asp Asp Thr Ser Gly Phe Lys Tyr Asp Val Leu Tyr Lys
35 40 45

Arg Phe Tyr Gln Leu Glu Lys Phe Asn Ile Thr Ser Arg Met Ala Thr
50 55 60

Pro Thr Glu Arg Leu Glu Pro Phe Ser Lys Ile Arg Arg Tyr Leu Lys
65 70 75 80

Arg Lys Ile Asn Lys Thr Tyr Pro Phe Ala Gln Arg Ala Tyr Ile Thr
85 90 95

Gln Glu Lys Ser Gly Phe Asp Pro Arg Leu Leu Val Phe Arg Pro Lys
100 105 110

Gly Asn Val Tyr Leu Asp Gly Tyr Trp Gln Ser Glu Asn Tyr Phe Lys
115 120 125

Asp Ile Glu Gly Ile Ile Arg Gln Asp Leu Ile Ile Lys Ser Pro Ser
130 135 140

Asp Ser Leu Asn Ile Ala Thr Ala Glu Arg Ile Lys Asn Thr Leu Ala
145 150 155 160

Ile Ala Val His Val Arg Phe Phe Asp Met Val Asp Ile Ser Asp Ser
165 170 175

Ser Asn Cys Gln Ser Asn Tyr Tyr His Thr Ala Ile Ala Lys Met Glu
180 185 190

Glu Lys Ile Pro Asn Ala His Tyr Phe Ile Phe Ser Asp Lys Pro Val
195 200 205

Leu Ala Arg Leu Ala Met Pro Leu Pro Asp Asp Arg Ile Thr Ile Ile
210 215 220

Asp His Asn Ile Gly Asp Met Asn Ala Tyr Ala Asp Leu Trp Leu Met
225 230 235 240

37847-517 Sequence Listing_ST25 (003)

Ser Leu Cys Lys His Phe Val Ile Ala Asn Ser Thr Phe Ser Trp Trp
245 250 255

Gly Ala Trp Leu Ser Asp Asn Lys Glu Lys Ile Val Ile Ala Pro Asp
260 265 270

Ile Lys Ile Thr Ser Gly Val Thr Gln Trp Gly Phe Asp Gly Leu Ile
275 280 285

Pro Asp Glu Trp Ile Lys Leu
290 295

<210> 177

<211> 289

<212> PRT

<213> Prevotella micans

<400> 177

Met Asp Val Ile Val Ile Phe Asn Gly Leu Gly Asn Gln Met Ser Gln
1 5 10 15

Tyr Ala Phe Tyr Leu Glu Lys Arg Leu Arg Asn Arg Gln Thr Thr Tyr
20 25 30

Phe Val Leu Asn Pro Arg Ser Thr Tyr Glu Leu Glu Arg Leu Phe Gly
35 40 45

Ile Pro Tyr Arg Ser Asn Leu Met Cys Arg Met Ile Tyr Lys Leu Leu
50 55 60

Asp Lys Ala Tyr Phe Ser Asn His Ile Arg Leu Lys Lys Ile Leu Arg
65 70 75 80

Thr Ala Leu Asn Ala Val Gly Ile Arg Leu Ile Val Glu Pro Ile Thr
85 90 95

Arg Asn Tyr Ser Leu Ser Asn Phe Thr His His Pro Gly Leu Thr Phe
100 105 110

37847-517 Sequence Listing_ST25 (003)

Tyr Arg Gly Gly Trp His Ser Glu Leu Asn Phe Thr Ser Val Val Thr
 115 120 125

Glu Leu Arg Arg Lys Phe Ile Phe Pro Pro Ser Asp Asp Glu Glu Phe
 130 135 140

Lys Arg Ile Ser Ala Leu Ile Ile Arg Thr Gln Ser Ile Ser Leu His
 145 150 155 160

Ile Arg Arg Gly Asp Tyr Leu Asp Tyr Ser Glu Tyr Gln Gly Val Cys
 165 170 175

Thr Glu Glu Tyr Tyr Glu Arg Ala Ile Glu Tyr Ile Arg Ser His Val
 180 185 190

Glu Asn Pro Val Phe Phe Val Phe Ser Asp Asp Lys Glu Tyr Ala Ile
 195 200 205

Asn Lys Phe Ser Gly Asp Asp Ser Phe Arg Ile Val Asp Phe Asn Thr
 210 215 220

Gly Glu Asn Ser Trp Arg Asp Met Gln Leu Met Ser Leu Cys Arg His
 225 230 235 240

His Ile Leu Ala Asn Ser Thr Phe Ser Trp Trp Gly Ala Trp Leu Asp
 245 250 255

Ser Ala Pro Glu Lys Ile Val Leu His Pro Ile Tyr His Met Arg Asp
 260 265 270

Val Pro Thr Arg Asp Phe Tyr Pro His Asn Trp Ile Gly Ile Ser Gly
 275 280 285

Glu

<210> 178

<211> 293

<212> PRT

37847-517 Sequence Listing_ST25 (003)

<213> Thermosynechococcus sp. NK55

<400> 178

Met Ile Ile Val Arg Leu Tyr Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Ala Gly Leu Ala Leu Ser Leu Arg His Ala Val Pro Leu Arg
20 25 30

Phe Asp Leu Asp Trp Phe Asp Gly Val Arg Leu His Gln Gly Leu Glu
35 40 45

Leu His Arg Val Phe Asp Leu Asp Leu Pro Arg Ala Ala Pro Ser Glu
50 55 60

Met Arg Gln Val Leu Gly Ser Phe Ser His Pro Leu Val Arg Arg Leu
65 70 75 80

Leu Val Arg Arg Arg Leu Arg Trp Leu Leu Pro Gln Gly Tyr Ala Leu
85 90 95

Glu Pro His Phe His Tyr Trp Pro Gly Phe Glu Ala Leu Gly Pro Lys
100 105 110

Ala Tyr Leu Asp Gly Tyr Trp Gln Ser Glu Arg Tyr Phe Ser Glu Tyr
115 120 125

Gln Asp Ala Val Arg Ala Ala Phe Arg Phe Ala Gln Pro Leu Asp Glu
130 135 140

Arg Asn Arg Gln Ile Val Glu Glu Met Ala Ala Cys Glu Ser Val Ser
145 150 155 160

Leu His Val Arg Arg Gly Asp Phe Val Gln Asp Pro Val Val Arg Arg
165 170 175

Val His Gly Val Asp Leu Ser Ala Tyr Tyr Pro Arg Ala Val Ala Leu
180 185 190

37847-517 Sequence Listing_ST25 (003)

Leu Met Glu Arg Met Arg Glu Pro Arg Phe Tyr Val Phe Ser Asp Asp
195 200 205

Pro Asp Trp Val Arg Ala Asn Leu Lys Leu Pro Ala Pro Met Ile Val
210 215 220

Ile Asp His Asn Arg Gly Glu His Ser Phe Arg Asp Met Gln Leu Met
225 230 235 240

Ser Ala Cys Arg His His Ile Leu Ala Asn Ser Ser Phe Ser Trp Trp
245 250 255

Gly Ala Trp Leu Asn Ser Gln Pro His Lys Leu Val Ile Ala Pro Lys
260 265 270

Arg Trp Phe Asn Val Asp Asp Phe Asp Thr Arg Asp Leu Tyr Cys Ser
275 280 285

Gly Trp Thr Val Leu
290

<210> 179

<211> 309

<212> PRT

<213> Coleofasciculus chthonoplastes

<400> 179

Met Leu Ser Leu Asn Lys Asn Phe Leu Phe Val His Ile Pro Lys Ser
1 5 10 15

Cys Ile Leu Lys Glu Val Tyr Ile Tyr Met Ile Ser Phe Pro Asn Leu
20 25 30

Gly Lys Gly Val Arg Leu Gly Asn Gln Met Phe Gln Tyr Ala Phe Leu
35 40 45

Arg Ser Thr Ala Arg Arg Leu Gly Val Lys Phe Tyr Cys Pro Ala Trp
50 55 60

Ser Gly Asp Ser Leu Phe Thr Leu Asn Asp Gln Glu Glu Arg Val Ser

2015259088 28 May 2018

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

275

280

285

Asp Tyr Arg Leu Val Ser Arg Leu Glu Lys Arg Leu Pro Lys Ser Leu
 290 295 300

Met Asn Phe Phe Tyr
 305

<210> 180

<211> 281

<212> PRT

<213> Bacteroides gallinarum

<400> 180

Met Arg Leu Ile Lys Met Thr Gly Gly Leu Gly Asn Gln Met Phe Ile
 1 5 10 15

Tyr Ala Phe Tyr Leu Arg Met Lys Lys Arg His Thr Asn Thr Arg Ile
 20 25 30

Asp Leu Ser Asp Met Met His Tyr Asn Val His His Gly Tyr Glu Met
 35 40 45

His His Val Phe Asn Leu Pro Lys Thr Glu Phe Cys Ile Asn Gln Pro
 50 55 60

Leu Lys Lys Val Ile Glu Phe Leu Phe Phe Lys Lys Ile Tyr Glu Arg
 65 70 75 80

Lys Gln Asp Ser Ser Asn Leu Leu Pro Phe Asp Lys Lys Tyr Phe Trp
 85 90 95

Pro Leu Leu Tyr Phe Lys Gly Phe Tyr Gln Ser Glu Arg Phe Phe Ala
 100 105 110

Asp Met Glu Asn Asp Ile Arg Lys Ala Phe Thr Phe Asn Ser Gly Leu
 115 120 125

Phe Asn Glu Lys Thr Gln Thr Met Leu Lys Gln Ile Glu His Asn Glu
 130 135 140

37847-517 Sequence Listing_ST25 (003)

His Ala Val Ser Leu His Val Arg Arg Gly Asp Tyr Leu Glu Pro Lys
145 150 155 160

His Trp Lys Thr Thr Gly Ser Val Cys Gln Leu Pro Tyr Tyr Ile Asn
165 170 175

Ala Ile Ala Glu Met Asn Arg Arg Ile Glu Gln Pro Phe Tyr Tyr Val
180 185 190

Phe Ser Asp Asp Ile Ala Trp Val Lys Glu Asn Leu Pro Leu Pro Gln
195 200 205

Ala Val Phe Ile Asp Trp Asn Lys Gly Val Glu Ser Trp Gln Asp Met
210 215 220

Met Leu Met Ser His Cys Arg His His Ile Ile Cys Asn Ser Thr Phe
225 230 235 240

Ser Trp Trp Gly Ala Trp Leu Asn Pro Lys Glu Asn Lys Thr Val Ile
245 250 255

Met Pro Glu Arg Trp Phe Gln His Cys Glu Thr Pro Asn Ile Tyr Pro
260 265 270

Ala Gly Trp Ile Lys Val Pro Ile Asn
275 280

<210> 181

<211> 268

<212> PRT

<213> Firmicutes bacterium CAG:882

<400> 181

Met Asn Asn Val Glu Ile Met Gly Gly Leu Gly Lys Gln Leu Phe Gln
1 5 10 15

Tyr Ala Phe Ser Arg Tyr Leu Gln Lys Leu Gly Val Lys Asn Val Val
20 25 30

37847-517 Sequence Listing_ST25 (003)

Leu Arg Lys Asp Phe Phe Thr Ile Gln Phe Pro Glu Asn Asn Gly Ile
35 40 45

Thr Lys Arg Glu Phe Val Leu Asp Lys Tyr Asn Thr Arg Tyr Val Ala
50 55 60

Ala Ala Gly Glu Lys Thr Tyr Arg Asp Tyr Cys Asp Glu Asn Asp Tyr
65 70 75 80

Arg Asp Asp Tyr Ala Ile Gly Ser Asp Glu Val Leu Tyr Glu Gly Tyr
85 90 95

Trp Gln Asn Ile Asp Phe Tyr Asn Val Val Arg Lys Glu Met Gln Glu
100 105 110

Glu Leu Lys Leu Lys Pro Glu Phe Ile Asp Asn Ser Met Ala Ala Val
115 120 125

Glu Lys Asp Met Ser Ser Cys Asn Ser Val Ala Leu His Ile Arg Arg
130 135 140

Ser Asp Tyr Leu Thr Gln Val Asn Ala Gln Ile Phe Glu Gln Leu Thr
145 150 155 160

Gln Asp Tyr Tyr Ala Ser Ala Val Ser Ile Ile Glu Gln Tyr Thr His
165 170 175

Glu Lys Pro Val Leu Tyr Ile Phe Ser Asp Asp Pro Glu Tyr Ala Ala
180 185 190

Glu Asn Met Lys Asp Phe Met Gly Cys Arg Thr Val Ile Met Pro Pro
195 200 205

Cys Glu Pro Tyr Gln Asp Met Tyr Leu Met Thr Arg Ala Lys His Asn
210 215 220

Ile Ile Ala Asn Ser Thr Phe Ser Trp Trp Gly Ala Thr Leu Asn Ala
225 230 235 240

37847-517 Sequence Listing_ST25 (003)

Asn Pro Asp Asn Ile Thr Val Ala Pro Ser Arg Trp Met Lys Gly Arg
245 250 255

Thr Val Asn Leu Tyr His Lys Asp Trp Ile Thr Leu
260 265

<210> 182

<211> 301

<212> PRT

<213> Bacteroides xylanisolvens

<400> 182

Met Ile Ala Val Asn Val Asn Ala Gly Leu Ala Asn Gln Met Phe His
1 5 10 15

Tyr Ala Phe Gly Arg Gly Leu Met Ala Lys Gly Leu Asp Val Cys Phe
20 25 30

Asp Gln Ser Asn Phe Lys Pro Arg Ser Gln Trp Ala Phe Glu Leu Val
35 40 45

Arg Leu Gln Asp Ala Phe Pro Ser Ile Asp Ile Lys Val Met Pro Glu
50 55 60

Gly His Phe Lys Trp Val Phe Pro Ser Leu Pro Arg Asn Gly Leu Glu
65 70 75 80

Arg Arg Phe Gln Glu Phe Met Lys Lys Trp His Asn Phe Ile Gly Asp
85 90 95

Glu Val Tyr Ile Asp Glu Pro Met Tyr Gly Tyr Val Pro Asp Met Glu
100 105 110

Lys Cys Ala Thr Arg Asn Cys Ile Tyr Lys Gly Phe Trp Gln Ser Glu
115 120 125

Lys Tyr Phe Arg His Cys Glu Asp Asp Ile Arg Lys Gln Phe Thr Phe
130 135 140

37847-517 Sequence Listing_ST25 (003)

Leu Pro Phe Asp Glu Leu Lys Asn Ile Glu Val Ala Ala Lys Met Ser
145 150 155 160

Gln Glu Asn Ser Val Ala Ile His Leu Arg Lys Gly Asp Asp Tyr Met
165 170 175

Gln Ser Glu Leu Met Gly Lys Gly Leu Cys Thr Val Asp Tyr Tyr Met
180 185 190

Lys Ala Ile Asp Tyr Met Arg Lys His Ile Asn Asn Pro His Phe Tyr
195 200 205

Val Phe Thr Asp Asn Pro Cys Trp Val Lys Asp Asn Leu Pro Glu Phe
210 215 220

Glu Tyr Ile Leu Val Asp Trp Asn Glu Val Ser Gly Lys Arg Asn Phe
225 230 235 240

Arg Asp Met Gln Leu Met Ser Cys Ala Lys His Asn Ile Ile Gly Asn
245 250 255

Ser Thr Tyr Ser Trp Trp Ala Ala Trp Leu Asn Ala Asn Gln Asp Lys
260 265 270

Ile Val Val Gly Pro Lys Arg Phe Phe Asn Pro Ile Asn Ser Phe Phe
275 280 285

Ser Thr Cys Asp Ile Met Cys Glu Asp Trp Ile Ser Leu
290 295 300

<210> 183
<211> 281
<212> PRT
<213> Geobacter sp. M18

<400> 183

Met Ile Gly Met Val Ile Phe Arg Ala Tyr Asn Gly Leu Gly Asn Gln
1 5 10 15

Met Phe Gln Tyr Ala Leu Gly Arg His Leu Ala Leu Leu Asn Glu Ala

37847-517 Sequence Listing_ST25 (003)

20

25

30

Glu Leu Lys Ile Asp Thr Thr Ala Phe Ala Asp Asp Pro Leu Arg Glu
35 40 45

Tyr Glu Leu His Arg Leu Lys Val Gln Gly Ser Ile Ala Thr Pro Asp
50 55 60

Glu Ile Ala Phe Phe Arg Glu Met Glu Asn Thr His Pro Gln Ala Tyr
65 70 75 80

Leu Arg Leu Thr Gln Lys Ser Arg Leu Phe Asp Pro Ala Ile Leu Ser
85 90 95

Ala Arg Gly Asn Ile Tyr Leu His Gly Phe Trp Gln Thr Glu Lys Tyr
100 105 110

Phe Ala Asp Ile Arg Glu Ile Leu Leu Asp Glu Phe Glu Pro Ile Val
115 120 125

Pro Ala Gly Glu Asp Ser Ile Lys Val Leu Ser His Met Lys Ala Thr
130 135 140

Asn Ala Val Ala Leu His Val Arg Arg Ser Asp Tyr Val Ser Asn Pro
145 150 155 160

Met Thr Leu Arg His His Gly Val Leu Pro Leu Asp Tyr Tyr Arg Glu
165 170 175

Ala Val Arg Arg Ile Ala Gly Met Val Pro Asp Pro Val Phe Phe Ile
180 185 190

Phe Ser Asp Asp Pro Gln Trp Ala Lys Asp Asn Ile Arg Leu Glu Tyr
195 200 205

Pro Ala Phe Cys Val Asp Ala His Asp Ala Ser Asn Gly His Glu Asp
210 215 220

Leu Arg Leu Met Arg Asn Cys Lys His Phe Ile Ile Ala Asn Ser Ser

37847-517 Sequence Listing_ST25 (003)

225 230 235 240

Phe Ser Trp Trp Gly Ala Trp Leu Ser Gln Asn Thr Gly Lys Lys Val
245 250 255

Val Ala Pro Leu Lys Trp Phe Ala Lys Pro Glu Ile Asp Thr Arg Asp
260 265 270

Ile Val Pro Leu Gln Trp Ile Arg Ile
275 280

<210> 184

<211> 281

<212> PRT

<213> Ruegeria pomeroyi DSS-3

<400> 184

Met Ile Thr Thr Arg Leu His Gly Arg Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Ala Ala Arg Gly Leu Ala Ala Arg Leu Gly Thr Gln Val Ala
20 25 30

Leu Asp Thr Arg Leu Ala Glu Ser Arg Gly Glu Gly Val Leu Thr Arg
35 40 45

Val Phe Asp Leu Asp Leu Ala Gln Pro Asp Gln Leu Pro Pro Leu Lys
50 55 60

Gly Asp Gly Leu Leu Arg His Gly Ala Trp Arg Leu Leu Gly Leu Ala
65 70 75 80

Pro Arg Phe Arg Arg Glu His Gly Leu Gly Tyr Asn Ala Ala Ile Glu
85 90 95

Thr Trp Asp Asp Gly Thr Tyr Leu His Gly Tyr Trp Gln Ser Glu Arg
100 105 110

Tyr Phe Ala His Ile Ala Ala Arg Ile Arg Ala Asp Phe Ala Phe Pro
115 120 125

37847-517 Sequence Listing_ST25 (003)

Ala Phe Ser Asn Ser Gln Asn Ala Glu Met Ala Ala Arg Ile Gly Asp
130 135 140

Thr Asp Ala Ile Ser Leu His Val Arg Arg Gly Asp Tyr Val Ala Leu
145 150 155 160

Ala Ala His Thr Leu Cys Asp Gln Arg Tyr Tyr Ala Ala Ala Leu Thr
165 170 175

Arg Leu Leu Glu Gly Val Ala Gly Asp Pro Val Val Tyr Leu Phe Ser
180 185 190

Asp Asp Pro Ala Trp Ala Arg Asp Asn Leu Ala Leu Pro Val Gln Lys
195 200 205

Val Val Val Asp Phe Asn Gly Pro Glu Thr Asp Phe Glu Asp Met Arg
210 215 220

Leu Met Ser Leu Cys Arg His Asn Ile Ile Gly Asn Ser Ser Phe Ser
225 230 235 240

Trp Trp Ala Ala Trp Leu Asn Ala His Pro Gly Lys Arg Ile Ala Ala
245 250 255

Pro Ala Ser Trp Phe Gly Asp Ala Lys Leu His Asn Pro Asp Leu Leu
260 265 270

Pro Pro Asp Trp Leu Lys Ile Glu Val
275 280

<210> 185

<211> 299

<212> PRT

<213> Lachnospiraceae bacterium 28-4

<400> 185

Met Ile Ile Ile Gln Leu Ala Gly Gly Leu Gly Asn Gln Met Gln Gln
1 5 10 15

37847-517 Sequence Listing_ST25 (003)

Tyr Ala Met Tyr Gln Lys Leu Leu Ser Leu Gly Lys Lys Val Lys Leu
 20 25 30
 Asp Ile Ser Trp Phe Glu Glu Lys Asn Arg Gln Lys Asn Val Tyr Ala
 35 40 45
 Arg Arg Glu Leu Glu Leu Asn Tyr Phe Lys Lys Ala Glu Tyr Glu Ala
 50 55 60
 Cys Thr Glu Glu Glu Arg Lys Ala Leu Val Gly Glu Gly Gly Phe Ala
 65 70 75 80
 Gly Lys Ile Lys Gly Lys Leu Phe Pro Gly Thr Arg Lys Ile Phe Arg
 85 90 95
 Glu Thr Glu Met Tyr His Pro Glu Ile Phe Asp Phe Glu Asp Arg Tyr
 100 105 110
 Leu Tyr Gly Tyr Phe Ala Cys Glu Lys Tyr Tyr Ala Asp Ile Met Glu
 115 120 125
 Ile Leu Gln Glu Gln Phe Val Phe Pro Pro Ser Gly Asn Pro Glu Asn
 130 135 140
 Gln Lys Met Ala Glu Arg Ile Ala Asp Gly Glu Ser Val Ser Leu His
 145 150 155 160
 Ile Arg Arg Gly Asp Tyr Leu Asp Ala Glu Asn Met Ala Met Phe Gly
 165 170 175
 Asn Ile Cys Thr Glu Glu Tyr Tyr Ala Gly Ala Ile Arg Glu Met Lys
 180 185 190
 Lys Ile Tyr Pro Ser Ala His Phe Phe Val Phe Ser Asp Asp Ile Pro
 195 200 205
 Tyr Ala Lys Glu Thr Tyr Ser Gly Glu Glu Phe Thr Val Val Asp Ile
 210 215 220

37847-517 Sequence Listing_ST25 (003)

Asn Arg Gly Lys Asp Ser Phe Phe Asp Ile Trp Leu Met Ser Gly Cys
225 230 235 240

Arg His Asn Ile Cys Ala Asn Ser Thr Phe Ser Phe Trp Gly Ala Arg
245 250 255

Leu Asn Arg Asn Lys Gly Lys Val Val Met Arg Pro Phe Ile His Lys
260 265 270

Asn Ser Gln Lys Phe Glu Pro Glu Leu Met His Glu Leu Trp Lys Gly
275 280 285

Trp Val Phe Ile Asp Asn Arg Gly Asn Ile Cys
290 295

<210> 186

<211> 273

<212> PRT

<213> Prevotella sp. CAG:1092

<400> 186

Met Arg Ile Leu Val Phe Thr Gly Gly Leu Gly Asn Gln Met Phe Glu
1 5 10 15

Tyr Ala Phe Tyr Lys His Leu Lys Ser Cys Phe Pro Lys Glu Ser Phe
20 25 30

Tyr Gly His Tyr Gly Val Lys Leu Lys Glu His Tyr Gly Leu Glu Ile
35 40 45

Asn Lys Trp Phe Asp Val Thr Leu Pro Pro Ala Lys Trp Trp Thr Leu
50 55 60

Pro Val Val Gly Leu Phe Tyr Leu Tyr Lys Lys Leu Val Pro Asn Ser
65 70 75 80

Lys Trp Leu Asp Leu Phe Gln Arg Glu Trp Lys His Lys Asp Ala Lys
85 90 95

37847-517 Sequence Listing_ST25 (003)

Val Phe Phe Pro Phe Lys Phe Thr Lys Gln Tyr Phe Pro Lys Glu Asn
100 105 110

Gly Trp Leu Lys Trp Lys Val Asp Glu Ala Ser Leu Cys Glu Lys Asn
115 120 125

Lys Lys Leu Leu Gln Val Ile His Asp Glu Glu Thr Cys Phe Val His
130 135 140

Val Arg Arg Gly Asp Tyr Leu Ala Ser Asn Phe Lys Ser Ile Phe Glu
145 150 155 160

Gly Cys Cys Thr Leu Asp Tyr Tyr Lys Arg Ala Leu Glu Tyr Met Asn
165 170 175

Lys Asn Asn Pro Lys Val Arg Phe Ile Cys Phe Ser Asp Asp Leu Glu
180 185 190

Trp Met Arg Lys Asn Leu Pro Met Asp Asp Ser Ala Ile Tyr Val Asp
195 200 205

Trp Asn Thr Gly Thr Asp Ser Pro Leu Asp Met Tyr Met Met Ser Gln
210 215 220

Cys Asp Asn Gly Ile Ile Ala Asn Ser Ser Phe Ser Tyr Trp Gly Ala
225 230 235 240

Tyr Leu Gly Gly Lys Lys Thr Thr Val Ile Tyr Pro Gln Lys Trp Trp
245 250 255

Asn Met Glu Gly Gly Asn Pro Asn Ile Phe Met Asp Glu Trp Leu Gly
260 265 270

Met

<210> 187
<211> 289
<212> PRT
<213> Spirosoma luteum

37847-517 Sequence Listing_ST25 (003)

<400> 187

Met Val Ile Ser Val Leu Ser Gly Gly Leu Gly Asn Gln Leu Phe Gln
 1 5 10 15

Tyr Ala Phe Gly Leu Lys Leu Ala Ala Gln Leu Gln Thr Glu Leu Arg
 20 25 30

Leu Glu Arg His Leu Leu Glu Ser Lys Ala Ile Ala Arg Leu Arg Gln
 35 40 45

Tyr Thr Pro Arg Thr Tyr Glu Leu Asp Thr Phe Gly Val Glu Ala Pro
 50 55 60

Ala Ala Ser Leu Met Asp Thr Val Ser Cys Leu Ser Arg Val Ala Leu
 65 70 75 80

Ser Asp Lys Thr Ala Leu Leu Leu Arg Glu Ser Thr Leu Thr Pro Asn
 85 90 95

Ala Ile Asn Asn Leu Asn Asn Arg Val Arg Asp Val Val Cys Leu Gly
 100 105 110

Tyr Trp Gln Ser Glu Glu Tyr Phe Arg Pro Ala Thr Glu Gln Leu Arg
 115 120 125

Lys His Leu Val Phe Arg Lys Asn Pro Ala Gln Ser Arg Ser Met Ala
 130 135 140

Asp Thr Ile Leu Ser Cys Gln Asn Ala Ala Phe Val His Ile Arg Arg
 145 150 155 160

Gly Asp Tyr Val Thr Asn Thr His Ala Asn Gln His His Gly Leu Cys
 165 170 175

Asp Val Ser Tyr Tyr Arg Arg Ala Cys Glu Tyr Val Lys Glu Cys Ile
 180 185 190

Pro Asp Val Gln Phe Phe Val Phe Ser Asp Asp Pro Asp Trp Ala Lys

37847-517 Sequence Listing_ST25 (003)

195

200

205

Arg Glu Leu Gly Ile His Leu Gln Pro Ala Arg Phe Ile Asp His Asn
210 215 220

Arg Gly Ala Asp Ser Trp Gln Asp Met Tyr Leu Met Ser Leu Cys Arg
225 230 235 240

His Ala Ile Val Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala Trp Leu
245 250 255

Asn Pro Val Ala Glu Arg Leu Val Val Ala Pro Gly Gln Trp Phe Val
260 265 270

Asn Gln Pro Val Leu Ser Gln Gln Ile Ile Pro Pro His Trp His Cys
275 280 285

Leu

<210> 188
<211> 296
<212> PRT
<213> Marinomonas posidonica

<400> 188

Met Ile Ile Val Asp Leu Ser Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Cys Ala Arg Ser Leu Ser Ile Glu Leu Asn Leu Pro Leu Lys
20 25 30

Val Val Tyr Gly Ser Leu Ala Ser Gln Thr Val His Asn Gly Tyr Glu
35 40 45

Leu Asn Arg Val Phe Gly Leu Asp Leu Glu Phe Ala Thr Glu Asn Asp
50 55 60

Met Gln Lys Asn Leu Gly Phe Phe Leu Ser Lys Pro Ile Leu Arg Lys
65 70 75 80

37847-517 Sequence Listing_ST25 (003)

Ile Phe Ser Lys Lys Pro Leu Asn Asn Leu Lys Phe Gln Asn Phe Phe
85 90 95

Pro Glu Asn Ser Phe Asn Tyr Asn Ser Ser Leu Phe Ser Tyr Ile Lys
100 105 110

Asp Ser Gly Phe Leu Gln Gly Tyr Trp Gln Thr Glu Lys Tyr Phe Leu
115 120 125

Asn His Lys Ser Gln Ile Leu Lys Asp Phe Cys Phe Val Asn Met Asp
130 135 140

Asp Glu Thr Asn Ile Ser Ile Ala Asn Asp Ile Gln Ser Gly His Ser
145 150 155 160

Ile Ser Ile His Val Arg Arg Gly Asp Tyr Leu Thr Asn Leu Lys Ala
165 170 175

Lys Ala Ile His Gly His Cys Ser Leu Asp Tyr Tyr Leu Lys Ala Ile
180 185 190

Glu Phe Leu Gln Glu Lys Ile Gly Glu Ser Arg Leu Phe Ile Phe Ser
195 200 205

Asp Asp Pro Glu Trp Val Ser Glu Asn Ile Ala Thr Arg Phe Ser Asp
210 215 220

Val Ser Val Ile Gln His Asn Arg Gly Val Lys Ser Phe Asn Asp Met
225 230 235 240

Arg Leu Met Ser Met Cys Asp His His Ile Ile Ala Asn Ser Ser Phe
245 250 255

Ser Trp Trp Gly Ala Trp Leu Asn Pro Ser Gln Asn Lys Lys Ile Ile
260 265 270

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

37847-517 Sequence Listing_ST25 (003)

Ile Pro Ser Ser Trp Ile Leu Lys
290 295

<210> 189

<211> 263

<212> PRT

<213> Bacteroides vulgatus

<400> 189

Met Lys Ile Val Val Phe Lys Gly Gly Leu Gly Asn Gln Leu Phe Gln
1 5 10 15

Tyr Ala Phe Tyr Lys Tyr Leu Ser Arg Lys Asp Glu Thr Phe Tyr Phe
20 25 30

Tyr Asn Asp Ala Trp Tyr Asn Val Ser His Asn Gly Phe Glu Leu Asp
35 40 45

Lys Tyr Phe Lys Thr Asp Asp Leu Lys Lys Cys Ser Arg Phe Trp Ile
50 55 60

Ile Leu Phe Lys Thr Ile Leu Ser Lys Leu Tyr His Trp Lys Ile Tyr
65 70 75 80

Val Val Gly Ser Val Glu Tyr Gln Tyr Pro Asn His Leu Phe Gln Ala
85 90 95

Gly Tyr Phe Leu Asp Lys Lys Tyr Tyr Asp Glu Asn Thr Ile Asp Phe
100 105 110

Lys His Leu Leu Leu Ser Glu Lys Asn Gln Ser Leu Leu Lys Asp Ile
115 120 125

Gln Asn Ser Asn Ser Val Gly Val His Ile Arg Arg Gly Asp Tyr Met
130 135 140

Thr Lys Gln Asn Leu Val Ile Phe Gly Asn Ile Cys Thr Gln Lys Tyr
145 150 155 160

37847-517 Sequence Listing_ST25 (003)

Tyr His Asp Ala Ile Arg Ile Ile Thr Glu Lys Val Asn Asp Ala Val
165 170 175

Phe Tyr Val Phe Ser Asp Asp Ile Ser Trp Val Gln Thr His Leu Asp
180 185 190

Ile Pro Asn Ala Val Tyr Val Asn Trp Asn Thr Gly Glu Ser Ser Ile
195 200 205

Tyr Asp Met Tyr Leu Met Ser Ser Cys Lys Tyr Asn Ile Ile Ala Asn
210 215 220

Ser Thr Phe Ser Tyr Trp Ala Ala Arg Leu Asn Lys Lys Thr Asn Met
225 230 235 240

Val Ile Tyr Pro Ser Lys Trp Tyr Asn Thr Phe Thr Pro Asp Ile Phe
245 250 255

Pro Glu Ser Trp Cys Gly Ile
260

<210> 190
<211> 285
<212> PRT
<213> Candidatus Pelagibacter ubique

<400> 190

Met Thr Ile Arg Ile Lys Leu Thr Gly Gly Leu Gly Asn Gln Met Phe
1 5 10 15

Gln Phe Ala Thr Gly Phe Ala Ile Ala Lys Lys Lys Asn Val Arg Leu
20 25 30

Ser Leu Asp Leu Lys Tyr Ile Asn Lys Arg Lys Leu Phe Asn Gly Phe
35 40 45

Glu Leu Gln Lys Ile Phe Asn Ile Tyr Ser Lys Val Ser Phe Leu Asn
50 55 60

37847-517 Sequence Listing_ST25 (003)

Lys	Thr	Leu	Ser	Phe	Lys	Ser	Ile	Asn	Phe	Thr	Glu	Ile	Leu	Asn	Arg	65	70	75	80
Ile	Asp	Thr	Thr	Phe	Tyr	Asn	Phe	Lys	Glu	Pro	His	Phe	His	Tyr	Thr	85	90	95	
Ser	Asn	Ile	Leu	Asn	Leu	Pro	Lys	His	Ser	Phe	Leu	Asp	Gly	Tyr	Trp	100	105	110	
Gln	Ser	Glu	Leu	Tyr	Phe	Asn	Glu	Phe	Ala	Thr	Glu	Ile	Lys	Arg	Ile	115	120	125	
Phe	Asn	Phe	Ser	Gly	Lys	Leu	Asp	Lys	Ser	Asn	Leu	Leu	Val	Ala	Asp	130	135	140	
Asp	Ile	Asn	Arg	Asn	Asn	Ser	Ile	Ser	Ile	His	Ile	Arg	Arg	Gly	Asp	145	150	155	160
Phe	Leu	Leu	Lys	Gln	Asn	Asn	Asn	His	His	Thr	Asp	Leu	Lys	Glu	Tyr	165	170	175	
Tyr	Leu	Lys	Ala	Ile	Asn	Glu	Thr	Ser	Lys	Ile	Phe	Lys	Asn	Pro	Lys	180	185	190	
Tyr	Phe	Ile	Phe	Ser	Asp	Asp	Thr	Ser	Trp	Thr	Val	Asp	Asn	Phe	Val	195	200	205	
Ile	Asp	His	Pro	Tyr	Ile	Ile	Val	Asp	Ile	Asn	Phe	Gly	Ala	Arg	Ser	210	215	220	
Phe	Leu	Asp	Met	Tyr	Leu	Met	Ser	Leu	Cys	Lys	Ser	Asn	Ile	Ile	Ala	225	230	235	240
Asn	Ser	Ser	Phe	Ser	Trp	Trp	Ser	Ala	Trp	Leu	Asn	Asn	Asn	Lys	Asp	245	250	255	
Lys	Ile	Ile	Tyr	Ala	Pro	Lys	Asn	Trp	Phe	Asn	Asp	Lys	Ser	Ile	Cys	260	265	270	

37847-517 Sequence Listing_ST25 (003)

Thr Asp Asp Leu Ile Pro Glu Ser Trp Asn Ile Ile Leu
 275 280 285

<210> 191
 <211> 289
 <212> PRT
 <213> Bacteroides sp. CAG:875

<400> 191

Met Ser Val Ile Ile Asn Met Ala Cys Gly Leu Ala Asn Arg Met Phe
 1 5 10 15

Gln Tyr Ala Phe Tyr Leu Tyr Leu Gln Lys Glu Gly Tyr Asp Ala Tyr
 20 25 30

Val Asp Tyr Phe Thr Arg Ala Asp Leu Val His Glu Asn Val Asp Trp
 35 40 45

Leu Arg Ile Phe Pro Glu Ala Thr Phe Arg Arg Ala Thr Ala Arg Asp
 50 55 60

Ile Arg Lys Met Gly Gly Gly His Asp Cys Phe Ser Arg Leu Arg Arg
 65 70 75 80

Lys Leu Leu Pro Met Thr Thr Lys Val Leu Glu Thr Ser Gly Ala Phe
 85 90 95

Glu Ile Ile Leu Pro Pro Lys Asn Arg Asp Ser Tyr Leu Leu Gly Ala
 100 105 110

Phe Gln Ser Ala Lys Met Val Glu Ser Val Asp Ala Glu Val Arg Arg
 115 120 125

Ile Phe Thr Phe Pro Glu Phe Glu Ser Gly Lys Asn Gln Tyr Phe Gln
 130 135 140

Thr Arg Leu Ala Gln Glu Asn Ser Val Gly Leu His Ile Arg Lys Gly
 145 150 155 160

Lys Asp Tyr Gln Glu Arg Ile Trp Tyr Lys Asn Thr Cys Gly Val Glu

37847-517 Sequence Listing_ST25 (003)

165 170 175

Tyr Tyr Arg Lys Ala Val Asp Leu Met Lys Glu Lys Val Asp Ser Pro
180 185 190

Ser Phe Tyr Val Phe Thr Asp Asn Pro Ala Trp Val Lys Glu Asn Leu
195 200 205

Ser Trp Leu Glu Tyr Lys Leu Val Asp Gly Asn Pro Gly Ser Gly Trp
210 215 220

Gly Ser His Cys Asp Met Gln Leu Met Ser Leu Cys Lys His Asn Ile
225 230 235 240

Ile Ser Asn Ser Ser Tyr Ser Trp Trp Gly Ala Tyr Leu Asn Asn Thr
245 250 255

Leu Asn Lys Ile Val Val Cys Pro Arg Ile Trp Phe Asn Pro Glu Ser
260 265 270

Thr Lys Asp Phe Ser Ser Asn Pro Leu Leu Ala Glu Gly Trp Ile Ser
275 280 285

Leu

<210> 192
<211> 309
<212> PRT
<213> Butyrivibrio fibrisolvens

<400> 192

Met Ile Ile Ile Lys Leu Gln Gly Gly Leu Gly Asn Gln Leu Phe Leu
1 5 10 15

Tyr Gly Leu Tyr Lys Asn Leu Lys His Leu Lys Arg Asp Val Lys Met
20 25 30

Asp Ile Glu Ser Gly Phe Glu Gly Asp Glu Leu Arg Lys Pro Cys Leu
35 40 45

37847-517 Sequence Listing_ST25 (003)

Asp Cys Met Asn Leu Glu Tyr Ala Ile Ala Thr Arg Asp Glu Val Thr
50 55 60

Asp Ile Arg Asp Ser Tyr Met Asp Ile Phe Ser Arg Ile Arg Arg Lys
65 70 75 80

Ile Thr Gly Arg Lys Thr Phe Asp Tyr Tyr Glu Pro Glu Asp Gly Asn
85 90 95

Tyr Asp Pro Lys Val Leu Glu Met Thr Lys Ala Tyr Leu Asn Gly Tyr
100 105 110

Phe Gln Ser Glu Lys Tyr Phe Gly Asp Glu Glu Ser Val Lys Ala Leu
115 120 125

Lys Asp Glu Leu Thr Lys Gly Lys Glu Asp Ile Leu Thr Ser Thr Asp
130 135 140

Leu Ile Thr Lys Ile Tyr His Asp Ile Lys Asn Ser Glu Ser Val Ser
145 150 155 160

Leu His Ile Arg Arg Gly Asp Tyr Leu Thr Pro Gly Ile Ile Glu Thr
165 170 175

Tyr Gly Gly Ile Cys Thr Asp Glu Tyr Tyr Asp Lys Ala Ile Ala Met
180 185 190

Ile Arg Glu Thr Phe Pro Glu Ala Arg Phe Phe Ile Phe Ser Asn Asp
195 200 205

Ile Glu Trp Cys Lys Glu Lys Phe Ala Gly Asp Lys Asn Ile Leu Phe
210 215 220

Val Asn Thr Ile Gly Ile Asn Leu Asp Ser Glu Asp Asn Ile Lys Ile
225 230 235 240

Gly Lys Ser Asp Lys Asp Ile Ser Glu Tyr Arg Asp Leu Ala Glu Leu
245 250 255

37847-517 Sequence Listing_ST25 (003)

Tyr Leu Met Ser Ala Cys Lys His His Ile Leu Ala Asn Ser Ser Phe
260 265 270

Ser Trp Trp Gly Ala Trp Leu Ser Asp His Glu Gly Met Thr Ile Ala
275 280 285

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

Asp Met Leu Leu Ile
305

<210> 193

<211> 282

<212> PRT

<213> Roseburia hominis

<400> 193

Met Val Thr Val Lys Ile Gly Asp Gly Met Gly Asn Gln Met Tyr Asn
1 5 10 15

Tyr Ala Cys Gly Tyr Ala Ala Ala Lys Arg Ser Gly Glu Lys Leu Arg
20 25 30

Leu Asp Ile Ser Glu Cys Asp Asn Ser Thr Leu Arg Asp Tyr Glu Leu
35 40 45

Asp His Phe Arg Val Val Tyr Asp Glu Lys Glu Ser Phe Pro Asn Arg
50 55 60

Thr Phe Trp Gln Lys Leu Tyr Lys Arg Leu Arg Arg Asp Ile Arg Tyr
65 70 75 80

His Val Ile Arg Glu Arg Asp Met Tyr Ala Val Asp Ala Arg Val Phe
85 90 95

Val Pro Ala Arg Arg Gly Arg Tyr Leu His Gly Tyr Trp Gln Cys Leu
100 105 110

37847-517 Sequence Listing_ST25 (003)

Gly Tyr Phe Glu Glu Tyr Leu Asp Asp Leu Arg Glu Met Phe Thr Pro
 115 120 125

Ala Tyr Glu Gln Thr Asp Ala Val Arg Glu Leu Met Gln Gln Phe Thr
 130 135 140

Gln Thr Pro Thr Cys Ala Leu His Val Arg Gly Gly Asp Leu Gly Gly
 145 150 155 160

Pro Asn Arg Ala Tyr Phe Gln Gln Ala Ile Ala Arg Met Gln Lys Glu
 165 170 175

Lys Pro Asp Val Thr Phe Ile Val Phe Thr Asn Asp Leu Pro Lys Ala
 180 185 190

Lys Glu Cys Leu Asp Asp Gly Glu Ala Arg Met Arg Tyr Ile Ala Glu
 195 200 205

Phe Gly Glu Ala Leu Ser Asp Ile Asp Glu Phe Phe Leu Met Ser Ala
 210 215 220

Cys Gln Asn Gln Ile Ile Ser Asn Ser Thr Tyr Ser Thr Trp Ala Ala
 225 230 235 240

Tyr Leu Asn Thr Leu Pro Gly Arg Ile Val Ile Val Pro Lys Phe His
 245 250 255

Gly Val Glu Gln Met Ala Leu Pro Asp Trp Ile Val Leu Asp Gly Gly
 260 265 270

Ala Cys Gln Lys Gly Glu Ile Asp Ala Val
 275 280

<210> 194

<211> 316

<212> PRT

<213> Rhodopirellula europaea

<400> 194

37847-517 Sequence Listing_ST25 (003)

Met	Ala	Thr	Ser	Val	His	Pro	His	Leu	Ser	Asp	Gly	Lys	Gln	Ala	Leu	1	5	10	15
Asp	Ser	Lys	Ala	Ala	Gln	Gln	Val	Cys	Ser	Thr	Gln	Ala	Ala	Ser	Ala	20	25	30	
Ser	Asp	Arg	Ala	Cys	Thr	Ile	Ser	Ile	Ser	Gly	Gly	Leu	Gly	Asn	Gln	35	40	45	
Met	Leu	Gln	Tyr	Ala	Ala	Gly	Arg	Ala	Leu	Ser	Ile	His	His	Asp	Cys	50	55	60	
Pro	Leu	Gln	Leu	Asp	Leu	Lys	Phe	Tyr	Ser	Ser	Lys	Arg	His	Arg	Ser	65	70	75	80
Tyr	Glu	Leu	Asp	Ala	Phe	Pro	Ile	Gln	Ala	Gln	Arg	Trp	Ile	Lys	Pro	85	90	95	
Ser	Phe	Phe	Ser	Gln	Val	Leu	Asp	Lys	Ile	Gln	Gly	Glu	Ser	Lys	Ser	100	105	110	
Ala	Pro	Thr	Tyr	Glu	Glu	Gln	Ser	Lys	Arg	Phe	Asp	Arg	Ala	Phe	Phe	115	120	125	
Asp	Ile	Glu	Leu	Pro	Ala	Arg	Ile	Arg	Gly	Tyr	Phe	Phe	Ser	Glu	Lys	130	135	140	
Tyr	Phe	Leu	Pro	Tyr	Ala	Asp	Gln	Ile	Arg	Thr	Glu	Leu	Thr	Pro	Pro	145	150	155	160
Val	Pro	Leu	Asp	Gln	Pro	Ala	Arg	Asp	Met	Ala	Gln	Arg	Leu	Ser	Glu	165	170	175	
Gly	Met	Ser	Thr	Ser	Leu	His	Val	Arg	Arg	Gly	Asp	Tyr	Val	Ser	Asn	180	185	190	
Ala	Asn	Ala	Arg	Gln	Arg	Phe	Trp	Ser	Cys	Thr	Ser	Glu	Tyr	Phe	Glu	195	200	205	

37847-517 Sequence Listing_ST25 (003)

Ala Ala Ile Glu Gln Met Pro Ala Asp Ser Thr Val Phe Val Phe Ser
210 215 220

Asp Asp Ile Glu Trp Ala Lys Gln Asn Ile Arg Ser Ser Arg Pro Thr
225 230 235 240

Val Tyr Val Asn Asp Glu Leu Lys Leu Ala Gly Ser Pro Glu Thr Gly
245 250 255

Leu Arg Asp Leu Trp Leu Met Thr His Ala Lys Ser His Ile Ile Ala
260 265 270

Asn Ser Ser Phe Ser Trp Trp Gly Ala Trp Leu Ser Gly Ser Glu Ala
275 280 285

Asn Leu Thr Ile Ala Pro Lys Lys Trp Phe Asn Asp Pro Glu Ile Asp
290 295 300

Asp Ser Asp Ile Val Pro Thr Ser Trp Arg Arg Ile
305 310 315

<210> 195
<211> 289
<212> PRT
<213> Rudanella lutea

<400> 195

Met Val Ile Ala Lys Ile Thr Ser Gly Leu Gly Asn Gln Leu Phe Gln
1 5 10 15

Tyr Ala Leu Gly Arg His Leu Ala Ile Gln Asn Gln Thr Arg Leu Trp
20 25 30

Phe Asp Leu Arg Tyr Tyr His Arg Thr Tyr Glu Thr Asp Thr Pro Arg
35 40 45

Gln Phe Lys Leu Asp Arg Phe Ser Ile Asp Tyr Asp Leu Leu Asp Tyr
50 55 60

Ser Pro Trp Leu Tyr Val Ser Lys Ala Thr Arg Leu Leu Pro Gly Arg

2015259088 28 May 2018

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

275

280

285

Leu

<210> 196

<211> 280

<212> PRT

<213> Paraprevotella clara YIT 11840

<400> 196

Met Arg Leu Ile Lys Met Thr Gly Gly Leu Gly Asn Gln Met Phe Ile
 1 5 10 15

Tyr Ala Met Tyr Leu Lys Met Lys Thr Ile Phe Pro Asp Val Arg Ile
 20 25 30

Asp Leu Ser Asp Met Val His Tyr Gln Val His Tyr Gly Tyr Glu Met
 35 40 45

Asn Lys Val Phe His Leu Pro Arg Thr Glu Phe Cys Ile Asn Arg Ser
 50 55 60

Leu Lys Lys Ile Ile Glu Phe Leu Leu Phe Lys Thr Ile Leu Glu Arg
 65 70 75 80

Lys Gln Gly Gly Ser Leu Val Pro Tyr Thr Arg Lys Tyr His Trp Pro
 85 90 95

Trp Ile Tyr Phe Lys Gly Phe Tyr Gln Ser Glu Lys Tyr Phe Ala Gly
 100 105 110

Ile Glu Lys Glu Val Arg Glu Ala Phe Val Phe Asp Ile Arg Arg Ala
 115 120 125

Ser Arg Arg Ser Leu Arg Ala Met Gln Glu Ile Lys Ala Asp Pro His
 130 135 140

Ala Val Ser Ile His Val Arg Arg Gly Asp Tyr Leu Leu Glu Lys His
 145 150 155 160

37847-517 Sequence Listing_ST25 (003)

Trp Lys Ala Leu Gly Cys Ile Cys Gln Ser Ser Tyr Tyr Leu Asn Ala
165 170 175

Leu Ala Glu Leu Glu Lys Arg Val Lys His Pro His Tyr Tyr Val Phe
180 185 190

Ser Glu Asp Leu Asn Trp Val Arg Gln Asn Leu Pro Leu Ile Lys Ala
195 200 205

Glu Phe Ile Asp Trp Asn Lys Gly Glu Asp Ser Trp Gln Asp Met Met
210 215 220

Leu Met Ser His Cys Arg His His Ile Ile Cys Asn Ser Thr Phe Ser
225 230 235 240

Trp Trp Gly Ala Trp Leu Asn Pro Leu Pro Asp Lys Ile Val Ile Ala
245 250 255

Pro Glu Arg Trp Thr Gln Thr Thr Asp Ser Ala Asp Val Val Pro Glu
260 265 270

Ser Trp Leu Lys Val Ser Ile Gly
275 280

<210> 197
<211> 288
<212> PRT
<213> *Smaragdicoccus niigatensis*

<400> 197

Met Ala Asp Val Val Val Thr Leu Ala Gly Gly Leu Gly Asn Gln Leu
1 5 10 15

Phe Gln Thr Ala Tyr Ala Lys Asn Leu Glu Ala Arg Gly His Arg Val
20 25 30

Thr Leu Asp Gly Thr Val Val Arg Trp Thr Arg Gly Leu His Ile Asp
35 40 45

37847-517 Sequence Listing_ST25 (003)

Pro Gln Ile Cys Gly Leu Lys Ile Leu Asn Ala Thr Pro Pro Ala Pro
50 55 60

Val Pro Gly Arg Leu Ala Ala Thr Val Leu Arg Arg Ala Leu Ala Thr
65 70 75 80

Arg Leu Arg Phe Gly Pro Asp Gly Arg Ile Val Arg Thr Gln Arg Thr
85 90 95

Leu Glu Phe Asp Glu Gln Tyr Leu Asn Leu Asn Ser Pro Gly Arg Tyr
100 105 110

Arg Val Glu Gly Tyr Trp Gln Cys Glu Arg Tyr Phe Ser Asp Val Gly
115 120 125

Gln Thr Val Arg Lys Val Phe Leu Asp Met Leu Gly Arg His Val Ser
130 135 140

Tyr Asn Gly Leu Ser Arg Leu Pro Ala Met Ala Asp Pro Ser Ser Ile
145 150 155 160

Ser Leu His Val Arg Arg Gly Asp Tyr Val Thr Ala Asn Phe Ile Asp
165 170 175

Pro Leu Ala Leu Glu Tyr Tyr Glu Arg Ala Leu Glu Glu Leu Ala Val
180 185 190

Pro Ser Pro Arg Ile Phe Val Phe Ser Asp Asp Leu Asp Trp Ala Thr
195 200 205

Arg Glu Leu Gly Arg Ile Cys Asp Val Ile Pro Val Glu Pro Asp Trp
210 215 220

Thr Ser His Pro Gly Gly Glu Ile Phe Leu Met Ser Gln Cys Ser His
225 230 235 240

His Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala Trp Leu Asp
245 250 255

37847-517 Sequence Listing_ST25 (003)

Gly Arg Thr Ser Ser Arg Val Val Ala Pro Arg Gln Trp Phe Ser Leu
260 265 270

Glu Thr Tyr Ser Ala Arg Asp Ile Val Pro Asp Arg Trp Thr Lys Val
275 280 285

<210> 198

<211> 304

<212> PRT

<213> Bacteroides fragilis CAG:558

<400> 198

Met Ile His Leu Ile Leu Gly Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Phe Ala Arg Ser Leu Ala Leu Gln Tyr Asn Glu Asn Ile Ser
20 25 30

Phe Asn Thr Ile Leu Tyr Lys Glu Leu Lys Asn Glu Glu Arg Ser Phe
35 40 45

Ser Leu Gly His Leu Asn Ile Asn Thr Met Cys Ile Val Glu Thr Pro
50 55 60

Asp Glu Asn Lys Arg Ile Trp Glu Leu Phe Asn Lys Gln Ile Phe His
65 70 75 80

Gln Lys Ile Ala Arg Lys Ile Leu Pro Ala Ser Ile Arg Trp Trp Trp
85 90 95

Met Ser Asn Arg Asn Ile Tyr Ala Asn Val Cys Gly Pro Tyr Lys Tyr
100 105 110

Tyr His Pro Arg His Arg Ser Gln Asn Thr Thr Ile Ile His Gly Gly
115 120 125

Phe Gln Ser Trp Lys Tyr Phe Lys Glu His Gln Ser Met Ile Lys Ala
130 135 140

37847-517 Sequence Listing_ST25 (003)

Glu Leu Lys Val Ile Thr Pro Ile Ser Glu Pro Asn Lys Lys Ile Leu
145 150 155 160

Lys Glu Ile Gln Asn Ser Asn Ser Ile Cys Val His Ile Arg Arg Gly
165 170 175

Asp Phe Leu Ser Ala Gln Phe Ser Pro His Leu Glu Val Cys Asn Lys
180 185 190

Asp Tyr Tyr Glu Lys Ala Ile Lys Met Ile Ser Ser Gln Ile Glu Asn
195 200 205

Pro Thr Phe Phe Ile Phe Ser Asn Thr His Glu Asp Leu Val Trp Ile
210 215 220

Arg Lys Asn Tyr Asn Ile Pro Gln Asn Ser Val Tyr Val Asp Leu Asn
225 230 235 240

Asn Pro Asp Tyr Glu Glu Leu Arg Leu Met Tyr Asn Cys Lys His Phe
245 250 255

Ile Leu Ser Asn Ser Ser Phe Ser Trp Trp Ala Gln Tyr Leu Ser Glu
260 265 270

Ser Lys Asn Lys Ile Ile Ile Ala Pro Lys Ile Trp Asp Lys Arg Lys
275 280 285

Gly Ile Asp Phe Ser Asp Ile Tyr Met Pro Glu Trp Ile Ile Ile Lys
290 295 300

<210> 199
<211> 295
<212> PRT
<213> Desulfovibrio desulfuricans

<400> 199

Met Ser Phe Ser Ile Asp Val Ala Ala Ile Gln Arg Met Ala Leu Val
1 5 10 15

Lys Val Asp Gly Gly Leu Gly Ser Gln Met Trp Gln Tyr Ala Leu Ser

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

20

25

30

Leu Ala Val Gly Lys Ser Ser Ser Phe Thr Val Lys His Asp Leu Ser
35 40 45

Trp Phe Arg His Tyr Ala Lys Asp Ile Arg Gly Ile Glu Asn Arg Phe
50 55 60

Phe Ile Leu Asn Ser Val Phe Thr Asn Ile Asn Leu Arg Leu Ala Ser
65 70 75 80

Glu Asn Glu Arg Leu Phe Phe His Ile Ala Leu Asn Arg Tyr Pro Asp
85 90 95

Ser Ile Cys Asn Phe Asp Pro Asp Ile Leu Ala Leu Lys Gln Pro Thr
100 105 110

Tyr Leu Gly Gly Tyr Tyr Val Asn Ala Gln Tyr Val Thr Ser Ala Glu
115 120 125

Lys Glu Ile Arg Glu Ala Tyr Val Phe Ala Pro Ala Val Glu Glu Ser
130 135 140

Asn Gln Ala Met Leu Gln Thr Ile His Ala Ala Pro Met Pro Val Ala
145 150 155 160

Val His Val Arg Arg Gly Asp Tyr Ile Gly Ser Met His Glu Val Leu
165 170 175

Thr Pro Arg Tyr Phe Glu Arg Ala Phe Lys Ile Leu Ala Ala Ala Leu
180 185 190

Gln Pro Lys Pro Thr Phe Phe Val Phe Ser Asn Gly Met Glu Trp Thr
195 200 205

Lys Lys Ala Phe Ala Gly Leu Pro Tyr Asp Phe Val Tyr Val Asp Ala
210 215 220

Asn Asp Asn Asp Asn Val Ala Gly Asp Leu Phe Leu Met Thr Gln Cys

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

225 230 235 240

Lys His Phe Thr Ile Ser Asn Ser Ser Leu Ser Trp Trp Gly Ala Trp
245 250 255

Leu Ser Gln Arg Ala Glu Asn Lys Thr Val Ile Met Pro Ser Lys Trp
260 265 270

Arg Gly Gly Lys Ser Pro Ile Pro Gly Glu Cys Met Arg Val Glu Gly
275 280 285

Trp His Met Cys Pro Val Glu
290 295

<210> 200
<211> 282
<212> PRT
<213> Hoeflea phototrophica

<400> 200

Met His Gly Gly Leu Gly Asn Gln Leu Phe Gln Tyr Ala Val Gly Arg
1 5 10 15

Ala Val Ala Leu Arg Thr Gly Ser Glu Leu Leu Leu Asp Thr Arg Glu
20 25 30

Phe Thr Ser Ser Asn Pro Phe Gln Tyr Asp Leu Gly His Phe Ser Ile
35 40 45

Gln Ala Lys Val Ala Asn Ser Ser Glu Leu Pro Pro Gly Lys Asn Arg
50 55 60

Pro Leu Ala Tyr Ala Trp Trp Arg Lys Phe Gly Arg Ser Pro Arg Phe
65 70 75 80

Val Arg Glu Gln Asp Leu Gly Tyr Asn Ala Arg Ile Glu Thr Ile Glu
85 90 95

Ala Asp Cys Tyr Leu His Gly Tyr Phe Gln Ser Gln Lys Tyr Phe Glu
100 105 110

37847-517 Sequence Listing_ST25 (003)

Asp Ile Ala Ser Ile Leu Trp Lys Asp Leu Ser Phe Arg Gln Ala Ile
115 120 125

Ser Gly Glu Asn Ala Ser Met Ala Glu Arg Ile Gln Ser Ala Pro Ser
130 135 140

Val Ser Met His Ile Arg Arg Gly Asp Tyr Leu Thr Ser Ala Lys Ala
145 150 155 160

Arg Ser Thr His Gly Ala Pro Asp Leu Gly Tyr Tyr Gly Arg Ala Leu
165 170 175

Gly Glu Ile Arg Ala Arg Ser Gly Ser Asp Pro Val Val Tyr Leu Phe
180 185 190

Ser Asp Asp Pro Asp Trp Val Arg Asn Asn Met Arg Met Asp Ala Asn
195 200 205

Leu Val Thr Val Ala Ile Asn Asp Gly Lys Thr Ala Phe Glu Asp Leu
210 215 220

Arg Leu Met Ser Leu Cys Asp His Asn Ile Ile Val Asn Ser Thr Phe
225 230 235 240

Ser Trp Trp Gly Ala Trp Leu Asn Pro Ser Leu Asp Lys Ile Val Val
245 250 255

Ala Pro Lys Arg Trp Phe Ala Asp Pro Lys Leu Ser Asn Pro Asp Ile
260 265 270

Thr Pro Pro Gly Trp Leu Arg Leu Gly Asp
275 280

<210> 201

<211> 286

<212> PRT

<213> Vibrio cholerae

<400> 201

37847-517 Sequence Listing_ST25 (003)

Met Lys Ile Ile Ser Phe Ser Gly Gly Leu Gly Asn Gln Leu Phe Gln
 1 5 10 15
 Tyr Ala Phe Tyr Leu Tyr Leu Lys Asp Asn Ser Asp Phe Gly Asn Ile
 20 25 30
 Phe Leu Asp Phe Ser Phe Tyr Glu Ser Gln Asn Lys Arg Asp Ala Val
 35 40 45
 Ile Arg Asn Phe Tyr Gly Val Asp Ser Leu Asp Ile Ile Lys Gln Ser
 50 55 60
 Ser Tyr Val Arg Gly Lys Phe Leu Ile Leu Lys Leu Ile Asn Lys Phe
 65 70 75 80
 Arg Phe Phe Asn Asn Leu Leu Glu Phe Val Asp Lys Glu Asn Gly Leu
 85 90 95
 Asp Glu Thr Leu Leu Ser Thr Asn Lys Val Phe Phe Asp Gly Tyr Trp
 100 105 110
 Gln Ser Tyr Arg Tyr Val Lys Asp Tyr Lys Ser Asn Ile Lys Glu Leu
 115 120 125
 Phe Ser Phe Tyr Asp Phe Lys Gly Asn Ile Leu Glu Val Arg Lys Lys
 130 135 140
 Ile Cys Gln Ser Asn Ser Val Cys Met His Val Arg Arg Gly Asp Tyr
 145 150 155 160
 Val Ala Glu Lys Asn Thr Lys Leu Val His Gly Val Cys Ser Leu Gln
 165 170 175
 Tyr Tyr Arg Asp Ala Leu Asn Asn Ile Lys Asn Val Asp Asn Ser Ile
 180 185 190
 Asp His Ile Phe Ile Phe Ser Asp Asp Ile Asp Trp Val Lys Asn Asn
 195 200 205

37847-517 Sequence Listing_ST25 (003)

Ile Ser Phe Asp Ile Pro Val Thr Val Val Asp Phe Val Gly Gln Ser
210 215 220

Val Pro Asp Tyr Ala Glu Met Leu Leu Phe Ser Cys Gly Lys His Lys
225 230 235 240

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

Arg Asn Gly Val Ile Val Ser Pro Lys Lys Trp Phe Ala Lys Glu Glu
260 265 270

Lys Asn Tyr Asp Glu Ile Phe Ile Glu Gly Ser Leu Arg Leu
275 280 285

<210> 202

<211> 295

<212> PRT

<213> Lachnospiraceae bacterium NK4A179

<400> 202

Met Ile Ile Val Arg Phe Arg Gly Gly Met Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Phe Leu Arg Tyr Leu Glu Met Lys Gly Ala Thr Leu Lys Ala
20 25 30

Asp Leu Ser Glu Phe Lys Cys Met Lys Thr His Ala Gly Tyr Glu Leu
35 40 45

Asp Lys Ala Phe Asp Leu His Pro Ala Glu Ala Ser Tyr Lys Glu Ile
50 55 60

Arg Ala Val Ala Asp Tyr Ile Pro Val Met His Arg Phe Pro Phe Ser
65 70 75 80

Arg Lys Val Phe Glu Ile Leu Tyr Lys Lys Glu Thr Lys Arg Val Glu
85 90 95

2015259088 28 May 2018

Met Ser Glu Asp Glu Arg Leu His Leu Ala Ser Ser Ser Glu Asp Leu
115 120 125

Leu Lys Cys Phe Thr Phe Ser Lys Thr Leu Asp Glu Lys Tyr Lys Gly
145 150 155 160

Tyr Thr Gly Thr Gly Leu Asp Ile Leu Gly Lys Glu Tyr Tyr Glu Lys
180 185 190

Ala Ala Glu Lys Ile Leu Ser Glu Asp Ala Asp Val Lys Phe Tyr Val
195 200 205

Phe Ser Asp Asp Arg Glu Lys Ala Glu Lys Leu Leu Ser Pro Phe Met
210 215 220

Lys Lys Met Val Phe Cys Asp Thr Pro Ala Ser His Ala Tyr Asp Asp
225 230 235 240

Met Tyr Leu Met Ser Arg Cys Arg His His Ile Ile Ala Asn Ser Thr
245 250 255

Phe Ser Phe Trp Gly Ala Arg Leu Ser Ala Asp Lys Ser Gly Ile Thr
260 265 270

Ile Cys Pro Lys Tyr Glu Asp Lys Asn Asn Thr Ala Asn Arg Leu Val
275 280 285

His Glu Gly Trp Gln Met Leu
290 295

37847-517 Sequence Listing_ST25 (003)

<210> 203
 <211> 289
 <212> PRT
 <213> Cecembia lonarensis

<400> 203

Met Ile Ile Met Lys Phe Met Gly Gly Leu Gly Asn Gln Ile Tyr Gln
 1 5 10 15

Tyr Ala Leu Gly Arg Lys Leu Ser Glu Leu His Asn Ser Phe Leu Ala
 20 25 30

Ser Asp Ile His Ile Tyr Lys Asn Asp Pro Asp Arg Glu Phe Val Leu
 35 40 45

Asp Lys Phe Asn Ile Lys Val Lys His Leu Pro Trp Lys Val Ile Lys
 50 55 60

Leu Leu Asn Ser Asp Tyr Ala Leu Lys Phe Asp Lys Val Phe His Thr
 65 70 75 80

Glu Phe Tyr His Glu Leu Val Leu Glu Lys Ala Leu Glu Ser Lys Asp
 85 90 95

Ile Pro Arg Lys Asn Asn Leu Tyr Leu Arg Gly Ser Trp Gly Asn Arg
 100 105 110

Lys Tyr Tyr Glu Asp Tyr Ile Asp Lys Ile Ser Asp Glu Ile Thr Leu
 115 120 125

Lys Glu Lys Phe Lys Thr Lys Asp Phe Asn Thr Val Asn Lys Lys Val
 130 135 140

Lys Asn Ser Asp Ser Val Gly Ile His Ile Arg Arg Gly Asp Tyr Glu
 145 150 155 160

Lys Val Ala His Phe Lys Asn Phe Tyr Gly Leu Leu Pro Pro Ser Tyr
 165 170 175

Tyr Ser Ala Ala Val Asp Phe Ile Gly Asn Arg Ile Glu Lys Ser Asn

37847-517 Sequence Listing_ST25 (003)

180

185

190

Phe Phe Ile Phe Ser Asp Asp Thr Asp Trp Val Lys Glu Asn Leu Pro
195 200 205

Phe Leu Lys Asp Ser Phe Phe Val Ser Asp Ile Ile Gly Ser Val Asp
210 215 220

Tyr Leu Glu Phe Glu Leu Leu Lys Asn Cys Lys His Gln Ile Ile Ala
225 230 235 240

Asn Ser Thr Phe Ser Trp Trp Ala Ala Arg Leu Asn Ser Asn Pro Ala
245 250 255

Lys Ile Val Ile Lys Pro Lys Arg Trp Phe Ala Asp Asp Arg Gln Gln
260 265 270

Ala Val Tyr Glu Ile Glu Asp Ser Tyr Tyr Ile Lys Glu Ala Ile Lys
275 280 285

Leu

<210> 204
<211> 292
<212> PRT
<213> Bacteroides ovatus

<400> 204

Met Lys Ile Val Asn Ile Leu Gly Gly Leu Gly Asn Gln Met Phe Val
1 5 10 15

Tyr Ala Met Tyr Leu Ala Leu Lys Glu Ala His Pro Glu Glu Glu Ile
20 25 30

Leu Leu Cys Arg Arg Ser Tyr Lys Gly Tyr Pro Leu His Asn Gly Tyr
35 40 45

Glu Leu Glu Arg Ile Phe Gly Val Glu Ala Pro Glu Ala Ala Leu Ser
50 55 60

37847-517 Sequence Listing_ST25 (003)

Gln Leu Ala Arg Val Ala Tyr Pro Phe Phe Asn Tyr Lys Ser Trp Gln
65 70 75 80

Leu Met Arg His Phe Leu Pro Leu Arg Lys Ser Met Ala Ser Gly Thr
85 90 95

Thr Gln Ile Pro Phe Asp Tyr Ser Glu Val Thr Arg Asn Asp Asn Val
100 105 110

Tyr Tyr Asp Gly Tyr Trp Gln Asn Glu Lys Asn Phe Leu Ser Ile Arg
115 120 125

Asp Lys Val Ile Lys Ala Phe Thr Phe Pro Glu Phe Arg Asp Glu Lys
130 135 140

Asn Lys Ala Leu Ser Asp Lys Leu Lys Ser Val Lys Thr Ala Ser Cys
145 150 155 160

His Ile Arg Arg Gly Asp Tyr Leu Lys Asp Pro Ile Tyr Gly Val Cys
165 170 175

Asn Ser Asp Tyr Tyr Thr Arg Ala Ile Thr Glu Leu Asn Gln Ser Val
180 185 190

Asn Pro Asp Met Tyr Cys Ile Phe Ser Asp Asp Ile Gly Trp Cys Lys
195 200 205

Glu Asn Phe Lys Phe Leu Ile Gly Asp Lys Glu Val Val Phe Val Asp
210 215 220

Trp Asn Lys Gly Gln Glu Ser Phe Tyr Asp Met Gln Leu Met Ser Leu
225 230 235 240

Cys His Tyr Asn Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala
245 250 255

Trp Leu Asn Asn Asn Asp Asp Lys Val Val Val Ala Pro Glu Arg Trp
260 265 270

37847-517 Sequence Listing_ST25 (003)

Met Asn Lys Thr Leu Glu Asn Asp Pro Ile Cys Asp Asn Trp Lys Arg
 275 280 285

Ile Lys Val Glu
 290

<210> 205
 <211> 281
 <212> PRT
 <213> Bacteroides coprocola CAG:162

<400> 205

Met Arg Leu Ile Lys Met Thr Gly Gly Leu Gly Asn Gln Met Phe Ile
 1 5 10 15

Tyr Ala Phe Tyr Leu Lys Met Lys Lys Leu Phe Pro His Thr Lys Ile
 20 25 30

Asp Leu Ser Asp Met Met His Tyr His Val His His Gly Tyr Glu Met
 35 40 45

Asn Arg Val Phe Ala Leu Pro His Thr Glu Phe Cys Ile Asn Arg Thr
 50 55 60

Leu Lys Lys Leu Met Glu Phe Leu Leu Cys Lys Val Val Tyr Glu Arg
 65 70 75 80

Lys Gln Lys Asn Gly Ser Met Glu Ala Phe Glu Lys Lys Tyr Ala Trp
 85 90 95

Pro Leu Ile Tyr Phe Lys Gly Phe Tyr Gln Ser Glu Arg Phe Phe Ala
 100 105 110

Asp Ile Glu Asp Asp Val Arg Lys Thr Phe Cys Phe Asn Met Glu Leu
 115 120 125

Ile Asn Ser Arg Ser Arg Glu Met Met Lys Ile Ile Asp Ala Asp Glu
 130 135 140

37847-517 Sequence Listing_ST25 (003)

His Ala Val Ser Ile His Ile Arg Arg Gly Asp Tyr Leu Leu Pro Lys
145 150 155 160

Phe Trp Ala Asn Ala Gly Cys Val Cys Gln Leu Pro Tyr Tyr Lys Asn
165 170 175

Ala Ile Thr Glu Leu Glu Lys His Glu Ser Thr Pro Ser Phe Tyr Val
180 185 190

Phe Ser Asp Asp Ile Glu Trp Val Lys Gln Asn Leu Ser Leu Pro Asn
195 200 205

Ala His Tyr Ile Asp Trp Asn Gln Gly Asn Asp Ser Trp Gln Asp Met
210 215 220

Met Leu Met Ser His Cys Arg Asn His Ile Ile Cys Asn Ser Thr Phe
225 230 235 240

Ser Trp Trp Gly Ala Trp Leu Asn Pro Arg Lys Asn Lys Thr Val Ile
245 250 255

Val Pro Ser Arg Trp Phe Met Lys Glu Glu Thr Pro Tyr Ile Tyr Pro
260 265 270

Val Ser Trp Ile Lys Val Pro Ile Asn
275 280

<210> 206

<211> 281

<212> PRT

<213> Bacteroides dorei

<400> 206

Met Arg Leu Ile Lys Val Thr Gly Gly Leu Gly Asn Gln Met Phe Ile
1 5 10 15

Tyr Ala Phe Tyr Leu Arg Met Lys Lys Tyr Tyr Pro Lys Val Arg Ile
20 25 30

2015259088 28 May 2018

His Arg Val Phe Lys Leu Pro His Thr Glu Phe Cys Ile Asn Gln Pro
50 55 60

Lys Gln Ala Pro Asn Ser Leu Arg Ala Phe Glu Lys Lys Tyr Phe Trp
85 90 95

Asp Ile Lys Asp Glu Val Arg Glu Ala Phe Thr Phe Asp Arg Ser Lys
115 120 125

Asn Ala Val Ser Leu His Ile Arg Arg Gly Asp Tyr Leu Gln Pro Lys
145 150 155 160

Ala Ile Ala Glu Met Ser Lys Arg Val Thr Ser Pro Ser Tyr Tyr Ile
180 185 190

Ala Val Tyr Ile Asp Trp Asn Thr Gly Glu Asp Ser Trp Gln Asp Met
210 215 220

Met Leu Met Ser His Cys Lys His His Ile Ile Cys Asn Ser Thr Phe
225 230 235 240

37847-517 Sequence Listing_ST25 (003)

Ser Trp Trp Gly Ala Trp Leu Asn Pro Ser Ile Asp Lys Thr Val Ile
245 250 255

Val Pro Ser Arg Trp Phe Gln Tyr Ser Glu Thr Pro Asp Ile Tyr Pro
260 265 270

Thr Gly Trp Ile Lys Val Pro Val Asp
275 280

<210> 207

<211> 276

<212> PRT

<213> Bacteroides intestinalis CAG:564

<400> 207

Met Ile Ile Val Arg Leu Trp Gly Gly Leu Gly Asn Gln Leu Phe Gln
1 5 10 15

Tyr Ser Phe Gly Gln Tyr Leu Glu Ile Glu Thr Asp Lys Lys Val Phe
20 25 30

Tyr Asp Val Ala Ser Phe Gly Thr Ser Asp Gln Leu Arg Lys Leu Glu
35 40 45

Leu Cys Ser Phe Ile Pro Asp Ile Pro Leu Tyr Asn Ala Tyr Phe Thr
50 55 60

Arg Tyr Thr Gly Val Lys Asn Arg Leu Phe Lys Ala Leu Phe Gln Trp
65 70 75 80

Ser Asn Thr Tyr Leu Ser Glu Ser Met Phe Asp Ile Cys Leu Leu Glu
85 90 95

Lys Ala Arg Gly Lys Ile Phe Leu Gln Gly Tyr Trp Gln Glu Glu Lys
100 105 110

Tyr Ala Thr Tyr Phe Pro Met Gln Lys Val Leu Ser Glu Trp Lys Asn
115 120 125

Pro Asn Val Leu Ser Glu Ile Glu Glu Asn Ile Arg Ser Ala Lys Ile

37847-517 Sequence Listing_ST25 (003)

130 135 140

Ser Val Ser Leu His Val Arg Arg Gly Asp Tyr Phe Ser Pro Lys Asn
145 150 155 160

Ile Asn Val Tyr Gly Val Cys Thr Glu Lys Tyr Tyr Glu Gln Ala Ile
165 170 175

Asp Arg Ala Asn Ser Glu Ile Glu Glu Asp Lys Gln Phe Phe Val Phe
180 185 190

Ser Asp Asp Ile Leu Trp Val Lys Asn His Val Ser Leu Pro Glu Ser
195 200 205

Thr Val Phe Val Pro Asn His Glu Ile Ser Gln Phe Ala Tyr Ile Tyr
210 215 220

Leu Met Ser Leu Cys Lys Val Asn Ile Ile Ser Asn Ser Thr Phe Ser
225 230 235 240

Trp Trp Gly Ala Tyr Leu Asn Gln His Lys Asn Gln Leu Val Ile Ala
245 250 255

Pro Ser Arg Trp Thr Phe Thr Ser Asn Lys Thr Leu Ala Leu Asp Ser
260 265 270

Trp Thr Lys Ile
275

<210> 208
<211> 297
<212> PRT
<213> Lachnospiraceae bacterium A4

<400> 208

Met Ile Val Ile His Val Met Gly Gly Leu Gly Asn Gln Leu Tyr Gln
1 5 10 15

Tyr Ala Leu Tyr Glu Lys Leu Arg Ala Leu Gly Arg Glu Val Lys Leu
20 25 30

37847-517 Sequence Listing_ST25 (003)

Asp Val Tyr Ala Tyr Arg Gln Ala Glu Gly Ala Glu Arg Glu Trp Arg
35 40 45

Ala Leu Glu Leu Glu Trp Leu Glu Gly Ile Arg Tyr Glu Val Cys Thr
50 55 60

Ala Ala Glu Arg Gln Gln Leu Leu Asp Asn Ser Met Arg Leu Ala Asp
65 70 75 80

Arg Val Arg Arg Arg Leu Thr Gly Arg Arg Asp Lys Thr Val Arg Glu
85 90 95

Cys Ala Ala Tyr Met Pro Glu Ile Phe Glu Met Asp Asp Val Tyr Leu
100 105 110

Tyr Gly Phe Trp Gly Cys Glu Lys Tyr Tyr Glu Asp Ile Ile Pro Leu
115 120 125

Leu Gln Glu Lys Ile Val Phe Pro Glu Ser Ser Asn Pro Lys Asn Ala
130 135 140

Asp Val Leu Arg Ala Met Ala Gly Glu Asn Ala Val Ser Val His Ile
145 150 155 160

Arg Arg Lys Asp Tyr Leu Thr Val Ala Asp Gly Lys Arg Tyr Met Gly
165 170 175

Ile Cys Thr Asp Ala Tyr Tyr Lys Gly Ala Phe Arg Tyr Ile Thr Glu
180 185 190

Arg Val Glu Arg Pro Val Phe Tyr Ile Phe Ser Asp Asp Pro Ala Phe
195 200 205

Ala Lys Thr Gln Phe Cys Glu Glu Asn Met His Val Val Asp Trp Asn
210 215 220

Thr Gly Arg Glu Ser Leu Gln Asp Met Ala Leu Met Ser Arg Cys Arg
225 230 235 240

37847-517 Sequence Listing_ST25 (003)

His Asn Ile Cys Ala Asn Ser Thr Phe Ser Ile Trp Gly Ala Arg Leu
245 250 255

Asn Arg His Pro Asp Lys Ile Met Ile Arg Pro Leu His His Asp Asn
260 265 270

Tyr Glu Ala Leu Asp Ala Arg Thr Val His Glu Tyr Trp Lys Gly Trp
275 280 285

Val Leu Ile Asp Ala Asp Gly Lys Val
290 295

<210> 209

<211> 285

<212> PRT

<213> Phaeobacter gallaeciensis

<400> 209

Met Ile Ile Thr Arg Leu His Gly Arg Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Ala Gly Arg Ala Leu Ala Asp Arg Leu Gly Val Ser Val Ala
20 25 30

Leu Asp Ser Arg Gly Ala Glu Leu Arg Gly Glu Gly Val Leu Thr Arg
35 40 45

Val Phe Asp Leu Asp Leu Ala Thr Pro Asp Ile Leu Pro Pro Leu Arg
50 55 60

Gln Arg Ala Pro Leu Gly Tyr Ala Leu Trp Arg Gly Leu Gly Gln His
65 70 75 80

Leu Gly Thr Gly Pro Lys Leu Arg Arg Glu Val Gly Leu Gly Tyr Asn
85 90 95

Pro Asp Phe Val Asp Trp Ser Asp Asn Ser Tyr Leu His Gly Tyr Trp
100 105 110

37847-517 Sequence Listing_ST25 (003)

Gln Ser Glu Arg Tyr Phe Ala Gln Ser Ala Glu Arg Ile Arg Arg Asp
 115 120 125

Phe Thr Phe Pro Glu Tyr Ser Asn Gln Gln Asn Ala Glu Met Ala Ala
 130 135 140

Arg Ile Gly Glu Thr Asn Ala Ile Ser Leu His Val Arg Arg Gly Asp
 145 150 155 160

Tyr Leu Thr Leu Ala Ala His Val Leu Cys Asp Gln Ala Tyr Tyr Glu
 165 170 175

Ala Ala Leu Ala Gln Val Leu Asp Gly Leu Glu Gly Gln Pro Thr Val
 180 185 190

Tyr Val Phe Ser Asp Asp Pro Gln Trp Ala Lys Glu Asn Leu Pro Leu
 195 200 205

Pro Cys Asp Lys Val Val Val Asp Phe Asn Gly Ala Asp Thr Asp Tyr
 210 215 220

Glu Asp Met Arg Leu Met Ser Leu Cys Lys His Asn Ile Ile Gly Asn
 225 230 235 240

Ser Ser Phe Ser Trp Trp Ala Ala Trp Leu Asn Gln Thr Pro Asp Arg
 245 250 255

Arg Val Ala Gly Pro Thr Lys Trp Phe Gly Asp Pro Lys Leu Asn Asn
 260 265 270

Pro Asp Ile Leu Pro Pro Asp Trp Leu Arg Ile Ser Val
 275 280 285

<210> 210

<211> 686

<212> PRT

<213> Firmicutes bacterium CAG:791

<400> 210

37847-517 Sequence Listing_ST25 (003)

Met	Ser	Gly	Gly	Leu	Gly	Asn	Gln	Met	Phe	Gln	Tyr	Ala	Leu	Tyr	Leu	1	5	10	15
Lys	Leu	Arg	Ser	Leu	Gly	Arg	Glu	Val	Cys	Phe	Asp	Asp	Lys	Ser	Gln	20	25	30	
Tyr	Asp	Glu	Glu	Thr	Phe	Arg	Asn	Ser	Ser	Gln	Lys	Arg	Arg	Pro	Lys	35	40	45	
His	Leu	Asp	Ile	Phe	Gly	Ile	Thr	Tyr	Pro	Ser	Ala	Gly	Lys	Glu	Glu	50	55	60	
Leu	Glu	Lys	Leu	Thr	Asp	Gly	Ala	Met	Asp	Leu	Pro	Ser	Arg	Ile	Arg	65	70	75	80
Arg	Lys	Ile	Leu	Gly	Arg	Lys	Ser	Leu	Glu	Lys	Asn	Asp	Arg	Asp	Phe	85	90	95	
Met	Phe	Asp	Pro	Ser	Phe	Leu	Glu	Glu	Thr	Glu	Gly	Tyr	Phe	Cys	Gly	100	105	110	
Gly	Phe	Gln	Ser	Pro	Arg	Tyr	Phe	Ala	Gly	Ala	Glu	Glu	Glu	Val	Arg	115	120	125	
Lys	Ala	Phe	Thr	Phe	Pro	Glu	Glu	Leu	Leu	Cys	Pro	Lys	Glu	Gly	Cys	130	135	140	
Ser	Arg	Gln	Glu	Gln	Lys	Met	Leu	Glu	Gln	Ser	Ala	Ser	Tyr	Ala	Glu	145	150	155	160
Arg	Ile	Arg	Lys	Ala	Asn	Cys	Glu	Ala	Ala	Asp	Arg	Gly	Val	Pro	Gly	165	170	175	
Gly	Gly	Ser	Ala	Ser	Ile	His	Leu	Arg	Phe	Gly	Asp	Tyr	Val	Asp	Lys	180	185	190	
Gly	Asp	Ile	Tyr	Gly	Gly	Ile	Cys	Thr	Asp	Ala	Tyr	Tyr	Asp	Thr	Ala	195	200	205	

37847-517 Sequence Listing_ST25 (003)

Ile	Arg	Cys	Leu	Lys	Glu	Arg	Asp	Pro	Gly	Met	Ile	Phe	Phe	Val	Phe	210	215	220
Ser	Asn	Asp	Glu	Glu	Lys	Ala	Gly	Glu	Trp	Ile	Arg	Tyr	Gln	Ala	Glu	225	230	235
Arg	Ser	Glu	Asn	Leu	Gly	Arg	Gly	His	Phe	Val	Leu	Val	Lys	Gly	Cys	245	250	255
Asp	Glu	Asp	His	Gly	Tyr	Leu	Asp	Leu	Tyr	Leu	Met	Thr	Leu	Cys	Arg	260	265	270
Asn	His	Val	Ile	Ala	Asn	Ser	Ser	Phe	Ser	Trp	Trp	Ala	Ser	Phe	Met	275	280	285
Cys	Asp	Ala	Pro	Asp	Lys	Met	Val	Phe	Ala	Pro	Ser	Ile	Trp	Asn	Asn	290	295	300
Gln	Lys	Asp	Gly	Ser	Glu	Leu	Ala	Arg	Thr	Asp	Ile	Tyr	Ala	Asp	Phe	305	310	315
Met	Gln	Arg	Ile	Ser	Pro	Arg	Gly	Thr	Arg	Leu	Ser	Asp	Arg	Pro	Leu	325	330	335
Ile	Ser	Val	Ile	Val	Thr	Ala	Tyr	Asn	Val	Ala	Pro	Tyr	Ile	Gly	Arg	340	345	350
Ala	Leu	Asp	Ser	Val	Cys	Gly	Gln	Thr	Trp	Lys	Asn	Leu	Glu	Ile	Ile	355	360	365
Ala	Val	Asp	Asp	Gly	Ser	Ser	Asp	Glu	Thr	Gly	Ala	Ile	Leu	Asp	Arg	370	375	380
Tyr	Ala	Ala	Gly	Asp	Ser	Arg	Ile	Gln	Val	Val	His	Thr	Glu	Asn	Arg	385	390	395
Gly	Val	Ser	Ala	Ala	Arg	Asn	Glu	Gly	Ile	Ala	His	Ala	Arg	Gly	Glu	405	410	415

37847-517 Sequence Listing_ST25 (003)

Tyr Ile Gly Phe Val Asp Gly Asp Asp Arg Ala His Pro Ala Met Tyr
420 425 430

Glu Ala Met Ile Arg Gly Ile Leu Ser Ser Gly Ala Asp Met Ala Val
435 440 445

Val Arg Tyr Arg Glu Val Ser Ala Glu Glu Thr Leu Thr Asp Ala Glu
450 455 460

Glu Gln Val Ala Ser Phe Asp Pro Val Leu Arg Ala Ser Val Leu Leu
465 470 475 480

Gln Gln Arg Asp Ala Val Gln Cys Phe Ile Arg Ala Gly Met Ala Glu
485 490 495

Glu Glu Gly Lys Ile Val Leu Arg Ser Ala Val Trp Asn Lys Leu Phe
500 505 510

His Arg Arg Leu Leu Arg Asp Asn Arg Phe Pro Glu Gly Thr Ser Ala
515 520 525

Glu Asp Ile Pro Phe Thr Thr Arg Ala Leu Cys Leu Ser Lys Lys Val
530 535 540

Leu Cys Val Pro Glu Ile Leu Tyr Asp Tyr Val Val Asn Arg Gln Glu
545 550 555 560

Ser Ile Met Asn Thr Gly Arg Ala Glu Arg Thr Leu Thr Gln Glu Ile
565 570 575

Pro Ala Trp Arg Thr His Leu Glu Leu Leu Lys Glu Ser Gly Leu Ser
580 585 590

Asp Leu Ala Glu Glu Ser Glu Tyr Trp Phe Tyr Arg Arg Met Leu Ser
595 600 605

Tyr Glu Glu Glu Tyr Arg Arg Cys Ser Glu Thr Ala Lys Glu Ala Lys
610 615 620

37847-517 Sequence Listing_ST25 (003)

Glu Leu Gln Glu Arg Ile Leu Lys His Arg Asp Arg Ile Leu Glu Leu
625 630 635 640

Ala Glu Glu His Ser Phe Gly Arg Arg Gly Asp Arg Glu Arg Leu Lys
645 650 655

Leu Tyr Val Asn Ser Pro Arg Gln Tyr Phe Leu Leu Ser Asp Leu Tyr
660 665 670

Glu Lys Thr Val Val Asn Trp Lys Asn Arg Pro Asp Lys Thr
675 680 685

<210> 211

<211> 290

<212> PRT

<213> Butyrivibrio proteoclasticus

<400> 211

Met Arg Lys Arg Ile Ile Ala Leu Asn Gly Gly Leu Gly Asn Gln Met
1 5 10 15

Phe Gln Tyr Ala Phe Ala Arg Met Leu Glu Asp Arg Lys His Cys Leu
20 25 30

Ile Glu Phe Asp Thr Gly Phe Tyr Ser Thr Val Asn Asp Arg Lys Leu
35 40 45

Ala Ile Gln Asn Tyr Asn Ile His Lys Tyr Asp Phe Cys Asn His Glu
50 55 60

Tyr Tyr Asn Lys Ile Arg Leu Leu Phe Gln Lys Ile Pro Phe Val Ala
65 70 75 80

Trp Leu Ala Gly Thr Tyr Lys Glu Tyr Ser Glu Tyr Gln Leu Asp Pro
85 90 95

Arg Val Phe Leu Phe Asn Tyr Arg Phe Tyr Tyr Gly Tyr Trp Gln Asn
100 105 110

Lys Gln Tyr Phe Glu Asn Ile Ser Asn Asp Ile Arg Asn Glu Leu Ser

2015259088 28 May 2018

120

Tyr Ile Gly Asn Val Ser Glu Lys Glu Asn Ala Leu Leu Asn Met Leu
130 135 140

Glu Ala His Asn Ala Ile Ala Ile His Val Arg Arg Gly Asp Tyr Thr
145 150 155 160

Gln Glu Gly Tyr Asn Lys Ile Tyr Ile Ser Leu Ser Lys Glu Tyr Tyr
165 170 175

Lys Arg Ala Val Ser Ile Ala Cys Lys Glu Leu Gly Asp Asn Asn Ile
180 185 190

Pro Leu Tyr Val Phe Ser Asp Asp Ile Asp Trp Cys Lys Ala Asn Leu
195 200 205

Ala Asp Ile Gly Asn Val Thr Phe Val Asp Asn Thr Ile Ser Ser Ser
210 215 220

Ala Asp Ile Asp Met Leu Met Met Lys Lys Ser Arg Cys Leu Ile Thr
225 230 235 240

Ala Asn Ser Thr Phe Ser Trp Trp Ser Ala Trp Leu Ser Asp Arg Asp
245 250 255

Asp Lys Ile Val Leu Val Pro Asp Lys Trp Leu Gln Asp Glu Glu Lys
260 265 270

Asn Thr Lys Leu Met Lys Ala Phe Ile Cys Asp Lys Trp Lys Ile Val
275 280 285

Pro Val
290

```
<210> 212
<211> 286
<212> PRT
<213> Bacteroides sp. 2 1 16
```

37847-517 Sequence Listing_ST25 (003)

<400> 212

Met Gln Val Val Ala Arg Ile Ile Gly Gly Leu Gly Asn Gln Met Phe
1 5 10 15

Ile Tyr Ala Thr Ala Arg Ala Leu Ala Leu Arg Ile Asp Ala Asp Leu
20 25 30

Ile Leu Asp Thr Gln Ser Gly Tyr Lys Asn Asp Leu Phe Lys Arg Asn
35 40 45

Phe Leu Leu Asp Ser Phe Cys Ile Ser Tyr Arg Lys Ala Asn Cys Phe
50 55 60

Gln Lys Tyr Asp Tyr Tyr Leu Gly Glu Lys Val Lys Ser Leu Gly Lys
65 70 75 80

Lys Thr His Phe Ser Val Ile Pro Phe Met Lys Tyr Ile Ser Glu Asn
85 90 95

Thr Ser Cys Asp Phe Val Asp Gly Leu Leu Lys Lys His Ile Leu Ser
100 105 110

Val Tyr Leu Asp Gly Tyr Trp Gln Asn Glu Ala Tyr Phe Lys Asp Tyr
115 120 125

Ala Ser Ile Ile Lys Lys Asp Phe Gln Phe Cys Gln Val Asn Asp Leu
130 135 140

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

Ala Ile Gly Val Arg Arg Tyr Gln Glu Leu Asn Ser His Gln Asn Thr
165 170 175

Lys Val Thr Asp Leu Asp Phe Tyr Gln Lys Ala Ile Asn Tyr Ile Glu
180 185 190

Ser Lys Val Asp Asn Pro Thr Phe Phe Ile Phe Ser Glu Asp Gln Glu
195 200 205

37847-517 Sequence Listing_ST25 (003)

Trp Val Lys Asn Asn Leu Glu Gln Lys Ser Asn Phe Ile Met Ile Ser
210 215 220

Pro Lys Glu Gly Asn Tyr Ser Ala Leu Asn Asp Met Tyr Leu Ile Ser
225 230 235 240

Leu Cys Lys His His Ile Val Ser Asn Ser Ser Phe Tyr Trp Trp Gly
245 250 255

Ala Trp Leu Ala Asn Asn Lys Asn Lys Ile Val Val Ala Ser Asp Cys
260 265 270

Phe Leu Asn Pro Gln Ser Ile Pro Asp Ser Trp Ile Lys Phe
275 280 285

<210> 213

<211> 300

<212> PRT

<213> Desulfomicrobium baculatum

<400> 213

Met Ala Lys Ile Val Thr Arg Ile Met Gly Gly Ile Gly Asn Gln Leu
1 5 10 15

Phe Cys Tyr Ala Ala Ala Arg Arg Leu Ala Leu Val Asn His Ala Glu
20 25 30

Leu Val Ile Asp Asp Val Thr Gly Phe Ser Arg Asp Arg Val Tyr Arg
35 40 45

Arg Arg Tyr Met Leu Asp His Phe Asn Ile Ser Ala Arg Lys Ala Thr
50 55 60

Asn Tyr Glu Arg Met Glu Pro Phe Glu Arg Tyr Arg Arg Gly Leu Ala
65 70 75 80

Lys Tyr Ile Ser Lys Lys Leu Pro Phe Phe Glu Arg Glu Tyr Ile Glu
85 90 95

37847-517 Sequence Listing_ST25 (003)

Gln Glu Arg Ile Glu Phe Asp Pro Arg Phe Leu Glu Tyr Arg Thr Tyr
 100 105 110

Asn Asn Ile Tyr Ile Asp Gly Leu Trp Gln Ser Glu Asn Tyr Phe Lys
 115 120 125

Asp Val Glu Asp Ile Ile Arg Asp Asp Leu Lys Ile Ile Pro Pro Thr
 130 135 140

Asp Leu Glu Asn Ile Asn Ile Ala Lys Lys Ile Lys Asn Ile Gln Asn
 145 150 155 160

Thr Ile Ala Met His Val Arg Trp Phe Asp Leu Pro Gly Ile Asn Leu
 165 170 175

Gly Asn Asn Val Ser Thr Tyr Tyr Tyr His Arg Ala Ile Ala Met Met
 180 185 190

Glu Gln Arg Ile Asn Ala Pro His Tyr Phe Leu Phe Ser Asp Asn Leu
 195 200 205

Glu Ala Val His Ser Lys Leu Asp Leu Pro Glu Gly Arg Val Thr Phe
 210 215 220

Val Ser Asn Asn Asp Gly Asp Asp Asn Ala Tyr Ala Asp Leu Trp Leu
 225 230 235 240

Met Ser Gln Cys Lys His Phe Ile Thr Ala Asn Ser Thr Phe Ser Trp
 245 250 255

Trp Gly Ala Trp Leu Gly Glu Ser Arg Asp Ser Val Val Leu Val Pro
 260 265 270

Arg Phe Ser Pro Asp Gly Gly Val Thr Ser Trp Cys Phe Thr Gly Leu
 275 280 285

Ile Pro Glu Arg Trp Glu Gln Val Ser Ser Ile Arg
 290 295 300

37847-517 Sequence Listing_ST25 (003)

<210> 214
 <211> 287
 <212> PRT
 <213> Prevotella pleuritidis

<400> 214

Met Asp Ile Val Leu Ile Phe Asn Gly Leu Gly Asn Gln Met Ser Gln
 1 5 10 15

Tyr Ala Phe Tyr Leu Ala Lys Arg Gln Arg Asn Asn His Thr Val Tyr
 20 25 30

Cys Val Phe Gly Pro Arg Thr Gln Tyr Ser Leu Asp Lys Leu Phe Asp
 35 40 45

Ile Pro Tyr Arg His Asn Ala Val Leu Val Leu Leu Tyr Arg Ala Leu
 50 55 60

Asp Lys Ala His Phe Ser Asn His Arg Trp Leu Arg Arg Leu Leu Arg
 65 70 75 80

Pro Thr Leu Gln Leu Leu Gly Val Lys Met Ile Val Glu Pro Leu Ser
 85 90 95

Arg Asp Phe Asp Met Arg His Phe Thr His Gln Lys Gly Ile Val Phe
 100 105 110

Tyr Arg Gly Gly Trp His Ser Glu Leu Asn Phe Thr Ala Val Ala Asp
 115 120 125

Ala Val Lys Arg Arg Phe Arg Phe Pro Glu Ile Gln Asp Ala Ala Val
 130 135 140

Leu Ala Val Ile Asp Arg Ile Lys Ser Cys Gln Ser Val Ser Leu His
 145 150 155 160

Leu Arg Arg Gly Asp Tyr Leu Gly Leu Ser Glu Phe Gln Gly Val Cys
 165 170 175

37847-517 Sequence Listing_ST25 (003)

Thr Glu Ala Tyr Tyr Glu His Ala Ile Ala Tyr Phe Glu Ser Gln Ile
180 185 190

Glu Ser Pro Glu Tyr Phe Val Phe Ser Asp Asp Pro Thr Tyr Ala Arg
195 200 205

Glu Gln Phe Gly Ala Asp Pro Asn Phe His Ile Ile Asp Leu Asn His
210 215 220

Gly Glu Asp Ala Trp Cys Asp Leu Leu Met Met Thr Gln Cys Arg Tyr
225 230 235 240

Asn Ile Ile Ala Asn Ser Thr Phe Ser Trp Trp Gly Ala Trp Leu Asn
245 250 255

Asp Asn Pro Ser Lys Ile Val Val His Pro Arg Tyr His Leu Asn Gly
260 265 270

Val Glu Thr Arg Asp Phe Tyr Pro Arg Asn Trp Ile Cys Ile Glu
275 280 285

<210> 215
<211> 277
<212> PRT
<213> Bacteroides sp. 1_1_14

<400> 215

Met Lys Val Ile Trp Phe Asn Gly Asn Leu Gly Asn Gln Val Phe Tyr
1 5 10 15

Cys Lys Tyr Lys Glu Phe Leu His Asn Lys Tyr Pro Asn Glu Thr Ile
20 25 30

Lys Tyr Tyr Ser Asn Ser Arg Ser Pro Lys Ile Cys Val Glu Gln Tyr
35 40 45

Phe Arg Leu Ser Leu Pro Asp Arg Ile Asp Ser Phe Lys Val Arg Phe
50 55 60

Val Phe Glu Phe Leu Gly Lys Phe Phe Arg Arg Ile Pro Leu Lys Phe

37847-517 Sequence Listing_ST25 (003)

65		70		75		80
Val	Pro	Lys	Trp	Tyr	Cys	Thr
			85			90
						95
Tyr	Phe	Glu	His	Tyr	Leu	Gln
			100			105
						110
Ser	Ser	Trp	Leu	Lys	Ala	Lys
			115			120
						125
Leu	Ile	Phe	Glu	Asn	Leu	Ile
			130			135
						140
Ile	Arg	Arg	Gly	Asp	Tyr	Ile
						150
						155
						160
Ser	Ala	Thr	Asp	Tyr	Tyr	Glu
						165
						170
						175
Tyr	Leu	Asp	Ser	Gln	Phe	Phe
						180
						185
						190
Lys	Asn	Asn	Phe	Lys	Gly	Asp
						195
						200
						205
Gly	Ala	Asp	Ser	Tyr	Leu	Asp
						210
						215
						220
Asn	Ile	Ile	Ala	Asn	Ser	Thr
						225
						230
						235
						240
His	Glu	Lys	Lys	Lys	Val	Met
						245
						250
						255
Ser	Gly	Arg	Gln	Met	Pro	Asn
						260
						265
						270
Glu	Thr	Lys	Arg	Lys		

37847-517 Sequence Listing_ST25 (003)

275

<210> 216
 <211> 314
 <212> PRT
 <213> Agromyces subbeticus

<400> 216

Met Val Gly Arg Val Gly Ile Ala Arg Arg Gln Ala Ala Asp Val Ser
 1 5 10 15

Cys Thr Asp Gly Glu Gly Leu Val Ala Trp Arg Ile Arg Thr Gly Glu
 20 25 30

Ile Val Leu Gly Leu Gln Gly Gly Ile Gly Asn Gln Leu Phe Glu Trp
 35 40 45

Ala Phe Ala Met Ala Leu Arg Ser Ile Gly Arg Arg Val Leu Phe Asp
 50 55 60

Ala Val Arg Cys Arg Gly Asp Arg Pro Leu Met Ile Gly Pro Leu Leu
 65 70 75 80

Pro Ala Ser Asp Trp Leu Ala Ala Pro Val Gly Leu Ala Leu Ala Gly
 85 90 95

Ala Thr Lys Ala Gly Leu Leu Ser Asp Arg Ser Trp Pro Arg Leu Val
 100 105 110

Arg Gln Arg Arg Ser Gly Tyr Asp Pro Ser Val Leu Glu Arg Leu Gly
 115 120 125

Gly Thr Ser Tyr Leu Leu Gly Thr Phe Gln Ser Ala Arg Tyr Phe Asp
 130 135 140

Gly Val Glu His Glu Val Arg Ala Ala Val Arg Ala Leu Leu Glu Gly
 145 150 155 160

Met Leu Thr Pro Ser Gly Arg Arg Phe Ala Asp Glu Leu Arg Ala Asp
 165 170 175

37847-517 Sequence Listing_ST25 (003)

28 May 2018
2015259088

Pro His Arg Val Ala Val His Val Arg Arg Gly Asp Tyr Val Ser Asp
180 185 190

Pro Asn Ala Ala Val Arg His Gly Val Leu Gly Ala Gly Tyr Tyr Asp
195 200 205

Gln Ala Leu Glu His Ala Ala Ala Leu Gly His Val Arg Arg Val Trp
210 215 220

Phe Ser Asp Asp Leu Asp Trp Val Arg Glu His Leu Ala Arg Asp Asp
225 230 235 240

Asp Leu Leu Cys Pro Ala Asp Ala Thr Arg His Asp Gly Gly Glu Ile
245 250 255

Ala Leu Ile Ala Ser Cys Ala Thr Arg Ile Ile Ala Asn Ser Ser Phe
260 265 270

Ser Trp Trp Gly Gly Trp Leu Gly Ala Pro Ser Ser Pro Ala His Pro
275 280 285

Val Ile Ala Pro Ser Thr Trp Phe Ala Asp Gly His Ser Asp Ala Ala
290 295 300

Glu Leu Val Pro Arg Asp Trp Val Arg Leu
305 310

<210> 217
<211> 291
<212> PRT
<213> Prevotella salivae

<400> 217

Met Ile Ala Thr Thr Leu Phe Gly Gly Leu Gly Asn Gln Met Phe Ile
1 5 10 15

Tyr Ala Thr Ala Lys Ala Leu Ser Leu His Tyr Arg Thr Pro Met Ala
20 25 30

37847-517 Sequence Listing_ST25 (003)

Phe Asn Leu Arg Gln Gly Phe Glu Gln Asp Tyr Lys Tyr Gln Arg His
35 40 45

Leu Glu Leu Asn His Phe Lys Cys Gln Leu Pro Thr Ala Lys Trp Ile
50 55 60

Thr Phe Asn Tyr Lys Gly Glu Leu Asn Ile Lys Arg Ile Ser Arg Arg
65 70 75 80

Ile Gly Arg Asn Leu Leu Cys Pro His Tyr Gln Phe Ile Lys Glu Lys
85 90 95

Glu Pro Phe His Tyr Glu Lys Arg Leu Phe Glu Phe Thr Asn Lys Asn
100 105 110

Ile Phe Leu Glu Gly Tyr Trp Gln Ser Pro Arg Tyr Phe Glu Asn Tyr
115 120 125

Ser Asp Glu Ile Arg Arg Asp Phe Gln Leu Lys Ser Ile Leu Pro His
130 135 140

Thr Ile Thr Asp Glu Leu Gln Met Leu Lys Gly Thr Gly Lys Pro Leu
145 150 155 160

Val Met Leu Gly Ile Arg Arg Tyr Gln Glu Val Lys Asp Lys Lys Asp
165 170 175

Ser Pro Tyr Pro Leu Cys Asn Lys Asp Tyr Tyr Ala Lys Ala Ile Ser
180 185 190

His Val Gln Glu Gln Leu Pro Ala Pro Leu Phe Val Val Phe Thr Gln
195 200 205

Glu Gln Ala Trp Ala Met Asn Asn Leu Pro Thr Asn Ala Asn Leu Tyr
210 215 220

Phe Val Lys Glu Lys Asp Asn Ala Trp Ala Thr Ile Ala Asp Met Tyr
225 230 235 240

37847-517 Sequence Listing_ST25 (003)

Leu Met Thr Gln Cys Gln His Ala Ile Ile Ser Asn Ser Thr Phe Tyr
245 250 255

Trp Trp Gly Ala Trp Leu Gln His Pro Ile Glu Asn His Ile Val Val
260 265 270

Ala Pro Asn Asn Phe Ile Asn Arg Asp Cys Val Cys Asp Asn Trp Ile
275 280 285

Ile Leu Asp
290

<210> 218

<211> 304

<212> PRT

<213> Carnobacterium sp. WN1359

<400> 218

Met Ile Phe Val Asp Leu Ser Glu Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Tyr Ser Arg Tyr Leu Gln Glu Leu Tyr Gly Gly Thr Leu Tyr
20 25 30

Leu Asn Thr Ser Ser Phe Lys Arg Lys Asn Ser Thr Arg Ser Tyr Ser
35 40 45

Leu Asn Asn Phe Tyr Leu Tyr Glu Asn Val Lys Leu Pro Ser Lys Phe
50 55 60

Arg Arg Val Ile Tyr Asn Phe Tyr Ser Lys Thr Ile Arg Met Phe Ile
65 70 75 80

Lys Lys Val Ile Arg Met Asn Pro Tyr Ser Asp Lys Tyr Tyr Phe Ser
85 90 95

Met Ile Pro Tyr Gly Phe Tyr Val Ser Ser Gln Val Phe Lys Tyr Leu
100 105 110

37847-517 Sequence Listing_ST25 (003)

Thr Val Pro Thr Thr Lys Arg His Asn Ile Phe Val Met Gly Thr Trp
 115 120 125

Gln Thr Asn Lys Tyr Phe Gln Ser Ile Asn Asp Lys Ile Lys Asp Glu
 130 135 140

Leu Lys Val Lys Thr Glu Pro Asn Glu Leu Asn Lys Lys Leu Ile Thr
 145 150 155 160

Glu Ile Asn Ser Asn Gln Ser Val Cys Val His Ile Arg Leu Gly Asp
 165 170 175

Tyr Thr Asn Pro Glu Phe Asp Tyr Leu His Val Cys Thr Ser Asp Tyr
 180 185 190

Tyr Leu Lys Gly Met Asp Tyr Ile Val Ser Lys Val Lys Glu Pro Asn
 195 200 205

Phe Tyr Ile Phe Ser Asn Ser Ser Ser Asp Ile Glu Trp Ile Lys Asn
 210 215 220

Asn Tyr Asn Phe Lys Tyr Lys Val Lys Tyr Ile Asp Leu Asn Asn Pro
 225 230 235 240

Asp Phe Glu Asp Phe Arg Leu Met Tyr Asn Cys Lys His Phe Ile Ile
 245 250 255

Ser Asn Ser Thr Phe Ser Trp Trp Ala Gln Phe Leu Ser Asn Asn Asp
 260 265 270

Lys Lys Ile Ile Val Ala Pro Ser Lys Trp Gln Lys Ser Asn Glu Asn
 275 280 285

Glu Ala Lys Asp Ile Tyr Leu Asp His Trp Lys Leu Ile Glu Ile Glu
 290 295 300

<210> 219

<211> 304

<212> PRT

<213> Butyrivibrio sp. AD3002

37847-517 Sequence Listing_ST25 (003)

<400> 219

Met Leu Ile Ile Gln Ile Ala Gly Gly Leu Gly Asn Gln Met Gln Gln
1 5 10 15

Tyr Ala Met Tyr Arg Lys Leu Leu Lys Ala Gly Ala Asp Arg Asn Ile
20 25 30

Lys Leu Asp Thr Lys Trp Phe Asp Glu Asp Lys Gln Ser Gly Val Leu
35 40 45

Ala Lys Arg Lys Leu Glu Leu Glu Tyr Phe Thr Gly Leu Pro Leu Pro
50 55 60

Val Cys Ser Glu Ser Glu Arg Ala Arg Phe Thr Asp Arg Ser Val Ala
65 70 75 80

Arg Lys Val Val Glu Lys Leu Val Pro Gly Met Gly Ser Arg Phe Thr
85 90 95

Glu Ser Cys Met Tyr His Pro Glu Ile Phe Glu Leu Lys Asp Lys Tyr
100 105 110

Ile Glu Gly Tyr Phe Ala Cys Gln Lys Tyr Tyr Asp Asp Ile Met Gly
115 120 125

Glu Leu Gln Glu Leu Phe Val Phe Pro Thr His Pro Asp Glu Glu Ile
130 135 140

Asn Ile Lys Asn Met Asn Leu Met Asn Glu Met Glu Met Val Pro Ser
145 150 155 160

Val Ser Val His Ile Arg Arg Gly Asp Tyr Leu Asp Pro Glu Asn Ala
165 170 175

Ala Leu Phe Gly Asn Ile Ala Thr Asp Ala Tyr Tyr Asp Ser Ala Met
180 185 190

Glu Tyr Phe Lys Ala Ile Asp Pro Asp Thr His Phe Tyr Ile Phe Thr

37847-517 Sequence Listing_ST25 (003)

195

200

205

Asn Asp Pro Glu Tyr Ala Arg Glu Lys Tyr Ala Asp Pro Gly Arg Tyr
 210 215 220

Thr Ile Val Asp His Asn Thr Gly Lys Tyr Ser Leu Leu Asp Ile Gln
 225 230 235 240

Leu Met Ser His Cys Arg Gly Asn Ile Cys Ala Asn Ser Thr Phe Ser
 245 250 255

Phe Trp Gly Ala Arg Leu Asn Arg Arg Lys Asp Lys Ile Pro Val Arg
 260 265 270

Thr Leu Val Met Arg Asn Asn Gln Pro Val Thr Pro Glu Leu Met His
 275 280 285

Glu Tyr Trp Pro Gly Trp Val Leu Val Asp Lys Asp Gly Lys Val Arg
 290 295 300

<210> 220

<211> 299

<212> PRT

<213> Clostridium sp. KLE 1755

<400> 220

Met Ile Val Ile Arg Val Met Gly Gly Leu Gly Asn Gln Met Gln Gln
 1 5 10 15

Tyr Ala Leu Tyr Glu Lys Phe Lys Ala Leu Gly Lys Glu Thr Arg Leu
 20 25 30

Asp Thr Ser Trp Phe Asp Asn Ala Ser Met Gln Glu Asn Val Leu Ala
 35 40 45

Arg Arg Ser Leu Glu Leu Arg Phe Phe Asp Asn Leu Thr Tyr Glu Ala
 50 55 60

Cys Thr Pro Gln Glu Arg Glu Ala Leu Leu Gly Lys Glu Gly Phe Phe
 65 70 75 80

37847-517 Sequence Listing_ST25 (003)

Asn Lys Leu Glu Arg Lys Leu Phe Pro Ser Lys Asn Lys His Phe Tyr
85 90 95

Glu Ser Glu Met Phe His Pro Glu Ile Phe Lys Leu Asp Asn Val Tyr
100 105 110

Leu Glu Gly His Trp Ala Cys Glu Lys Tyr Tyr His Asp Ile Met Pro
115 120 125

Leu Leu Gln Ser Lys Ile Ile Phe Pro Lys Thr Asp Asn Ile Gln Asn
130 135 140

Asn Met Leu Lys Asn Lys Met Asn Ser Glu Asn Ser Val Ser Ile His
145 150 155 160

Ile Arg Arg Gly Asp Tyr Leu Asp Pro Glu Asn Ala Ala Met Phe Gly
165 170 175

Gly Ile Cys Thr Asp Ser Tyr Tyr Lys Ser Ala Glu Gly Tyr Ile Arg
180 185 190

Asn Arg Val Thr Asn Pro His Phe Tyr Leu Phe Ser Asp Asp Pro Ala
195 200 205

Tyr Leu Arg Glu His Tyr Lys Gly Glu Glu Tyr Thr Val Val Asp Trp
210 215 220

Asn His Gly Ala Asp Ser Phe Tyr Asp Met Glu Leu Met Ser Cys Cys
225 230 235 240

Lys His Asn Val Cys Ala Asn Ser Thr Phe Ser Phe Trp Gly Ala Arg
245 250 255

Leu Asn Arg Thr Glu Lys Lys Ile Val Ile Arg Pro Ala Lys His Lys
260 265 270

Asn Ser Gln Gln Ala Glu Pro Glu Arg Met His Glu Leu Trp Glu Asn
275 280 285

37847-517 Sequence Listing_ST25 (003)

Trp Val Ile Ile Asp Glu Glu Gly Arg Ile Val
290 295

<210> 221

<211> 275

<212> PRT

<213> Bacteroides vulgatus

<400> 221

Met Lys Phe Phe Val Phe Gly Gly Gly Leu Gly Asn Gln Leu Phe Gln
1 5 10 15

Tyr Ser Tyr Tyr Arg Tyr Leu Lys Lys Lys Tyr Pro Ser Glu Arg Ile
20 25 30

Leu Gly Ile Tyr Pro Asp Ser Leu Lys Ala His Asn Gly Ile Glu Ile
35 40 45

Asp Lys Trp Phe Asp Ile Glu Leu Pro Pro Thr Ser Tyr Leu Tyr Asn
50 55 60

Lys Leu Gly Ile Leu Leu Tyr Arg Val Asn Arg Phe Leu Tyr Asn His
65 70 75 80

Gly Tyr Arg Leu Leu Phe Cys Asn Arg Val Tyr Pro Gln Ser Met Lys
85 90 95

His Phe Phe Gln Trp Gly Asp Trp Gln Asp Tyr Ser Ile Ile Lys Gln
100 105 110

Ile Asn Ile Phe Glu Phe Arg Ser Glu Leu Pro Ile Gly Lys Glu Asn
115 120 125

Met Glu Phe Leu Lys Lys Met Glu Thr Cys Asn Ser Ile Ser Val His
130 135 140

Ile Arg Arg Gly Asp Tyr Leu Lys Thr Asp Leu Ile His Ile Tyr Gly
145 150 155 160

37847-517 Sequence Listing_ST25 (003)

Gly Ile Cys Thr Ser Lys Tyr Tyr Arg Glu Ala Ile Lys Phe Met Glu
165 170 175

Gln Glu Val Glu Glu Pro Phe Phe Phe Phe Ser Asp Asp Cys Leu
180 185 190

Tyr Val Glu Thr Glu Phe Ala Asp Ile Arg Asn Lys Ile Ile Ile Ser
195 200 205

His Asn Arg Asp Asp Arg Ser Phe Phe Asp Met Tyr Leu Met Ala His
210 215 220

Ala Lys Asn Met Ile Leu Ala Asn Ser Thr Phe Ser Cys Trp Ala Ala
225 230 235 240

Tyr Leu Asn Arg Thr Ala Lys Ile Ile Ile Thr Pro Asp Arg Trp Val
245 250 255

Asn Thr Asp Phe Ser Lys Leu Glu Ala Leu Pro Asn Glu Trp Ile Lys
260 265 270

Ile Arg Val
275

<210> 222
<211> 280
<212> PRT
<213> Paraprevotella xylaniphila

<400> 222

Met Arg Leu Ile Lys Met Thr Gly Gly Leu Gly Asn Gln Met Phe Ile
1 5 10 15

Tyr Ala Met Tyr Leu Lys Met Arg Ala Val Phe Pro Asp Thr Arg Ile
20 25 30

Asp Leu Ser Asp Met Val His Tyr Arg Val His Tyr Gly Tyr Glu Met
35 40 45

37847-517 Sequence Listing_ST25 (003)

Asn Lys Val Phe Asn Leu Pro Arg Thr Glu Phe Arg Ile Asn Arg Ser
50 55 60

Leu Lys Lys Ile Ile Glu Phe Leu Leu Phe Lys Thr Ile Leu Glu Arg
65 70 75 80

Lys Gln Gly Gly Ser Leu Val Pro Tyr Ile Arg Lys Tyr His Trp Pro
85 90 95

Trp Ile Tyr Phe Lys Gly Phe Tyr Gln Ser Glu Glu Tyr Phe Ala Gly
100 105 110

Val Glu Lys Glu Val Arg Glu Ala Phe Val Phe Asp Val Arg Arg Val
115 120 125

Asn Arg Lys Ser Leu Cys Ala Met Gln Glu Ile Met Ala Asp Pro Asp
130 135 140

Ala Val Ser Ile His Val Arg Arg Gly Asp Tyr Leu Gln Gly Lys His
145 150 155 160

Trp Lys Ser Leu Gly Cys Ile Cys Gln Arg Ser Tyr Tyr Leu Asn Ala
165 170 175

Leu Ser Glu Leu Glu Lys Arg Ile Val His Pro His Tyr Tyr Val Phe
180 185 190

Ser Glu Asp Leu Asp Trp Val Arg Gln Tyr Leu Pro Leu Glu Asn Ala
195 200 205

Val Phe Ile Asp Trp Asn Lys Gly Glu Asp Ser Trp Gln Asp Met Met
210 215 220

Leu Met Ser His Cys Arg His His Ile Ile Cys Asn Ser Thr Phe Ser
225 230 235 240

Trp Trp Gly Ala Trp Leu Asn Pro Ser Pro Asp Lys Ile Val Ile Ala
245 250 255

37847-517 Sequence Listing_ST25 (003)

Pro Glu Arg Trp Thr Gln Thr Thr Asn Ser Ala Asp Val Val Pro Glu
260 265 270

Ser Trp Leu Lys Val Ser Ile Gly
275 280

<210> 223
<211> 290
<212> PRT
<213> Thauera sp. 28

<400> 223

Met Thr Asp Arg Ala Leu Ile Ala Ile Val Lys Gly Gly Leu Gly Asn
1 5 10 15

Gln Leu Phe Ile Tyr Ala Ala Ala Arg Ala Met Ala Leu Arg Thr Gly
20 25 30

Arg Gln Leu Tyr Leu Asp Ala Val Arg Gly Tyr Leu Ala Asp Asp Tyr
35 40 45

Gly Arg Ser Phe Arg Leu Asn Arg Phe Pro Ile Glu Ala Glu Leu Met
50 55 60

Pro Glu Gln Trp Arg Val Ala Ser Thr Leu Arg His Pro Arg Ala Lys
65 70 75 80

Leu Val Arg Ala Leu Asn Lys Tyr Leu Pro Glu Ala Trp Arg Phe Tyr
85 90 95

Val Ala Glu Arg Gly Asp Thr Arg Pro Gly Ala Leu Trp Asn His Gly
100 105 110

Arg Asn Val Lys Arg Val Thr Leu Met Gly Tyr Trp Gln Asp Glu Ala
115 120 125

Tyr Phe Leu Asp Tyr Ala Glu Leu Leu Arg Arg Glu Leu Gly Pro Pro
130 135 140

Met Pro Asp Ala Pro Glu Val Arg Ala Arg Gly Glu Arg Phe Ala Gly

28 May 2018
2015259088

37847-517 Sequence Listing_ST25 (003)

145 150 155 160

Thr Glu Ser Val Phe Leu His Val Arg Arg Cys Arg Tyr Ser Pro Leu
165 170 175

Leu Asp Ala Gly Tyr Tyr Gln Lys Ala Val Asp Leu Ala Cys Ala Glu
180 185 190

Leu Asn Lys Pro Val Phe Met Ile Phe Gly Asp Asp Ile Glu Trp Val
195 200 205

Val Asn Asn Ile Asp Phe Arg Gly Ala Gly Tyr Glu Arg Gln Asp Tyr
210 215 220

Asp Glu Ser Asp Glu Leu Ala Asp Leu Trp Leu Met Thr Arg Cys Arg
225 230 235 240

His Ala Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Ala Ala Trp Leu
245 250 255

Gly Gly Ala Ala Gly Ser Gly Arg His Val Trp Ala Pro Gly Gln Ser
260 265 270

Gly Leu Ala Leu Lys Cys Ala Lys Ser Trp Glu Ala Val Asp Ala Gln
275 280 285

Pro Glu
290

<210> 224
<211> 305
<212> PRT
<213> Subdoligranulum variabile

<400> 224

Met Ile Tyr Ala Glu Leu Ala Gly Gly Leu Gly Asn Gln Met Phe Ile
1 5 10 15

Tyr Ala Phe Ala Arg Ala Leu Gly Leu Arg Cys Gly Glu Ala Val Thr
20 25 30

37847-517 Sequence Listing_ST25 (003)

Leu Leu Asp Arg Gln Asp Trp Arg Asp Gly Ala Pro Ala His Thr Ala
35 40 45

Cys Ala Leu Glu Gly Leu Asn Leu Val Pro Glu Val Lys Ile Leu Ala
50 55 60

Glu Pro Gly Phe Ala Lys Arg His Leu Pro Arg Gln Asn Thr Ala Lys
65 70 75 80

Ala Leu Met Ile Lys Tyr Glu Gln Arg Gln Gly Leu Met Ala Arg Asp
85 90 95

Trp His Asp Trp Glu Arg Arg Cys Ala Pro Val Leu Asn Leu Leu Gly
100 105 110

Leu His Phe Ala Thr Asp Gly Tyr Thr Pro Val Arg Arg Gly Pro Ala
115 120 125

Arg Asp Phe Leu Ala Trp Gly Tyr Phe Gln Ser Glu Ala Tyr Phe Ala
130 135 140

Asp Phe Ala Pro Thr Ile Arg Ala Glu Leu Arg Ala Lys Gln Ala Pro
145 150 155 160

Ala Gly Val Trp Ala Glu Lys Ile Arg Ala Ala Ala Cys Pro Val Ala
165 170 175

Leu His Leu Arg Arg Gly Asp Tyr Cys Arg Pro Glu Asn Glu Ile Leu
180 185 190

Gln Val Cys Ser Pro Ala Tyr Tyr Ala Arg Ala Ala Ala Ala Ala Ala
195 200 205

Ala Ala Tyr Pro Glu Ala Thr Leu Phe Val Phe Ser Asp Asp Ile Asp
210 215 220

Trp Ala Lys Glu His Leu Asp Thr Ala Gly Leu Pro Ala Val Trp Met
225 230 235 240

37847-517 Sequence Listing_ST25 (003)

Pro Arg Gly Asp Ala Val Gly Asp Leu Asn Leu Met Ala Leu Cys Arg
245 250 255

Gly Phe Ile Leu Ser Asn Ser Thr Tyr Ser Trp Trp Ala Gln Tyr Leu
260 265 270

Ala Gly Glu Gly Arg Thr Val Trp Ala Pro Asp Arg Trp Phe Ala His
275 280 285

Thr Lys Gln Thr Ala Leu Tyr Gln Pro Gly Trp His Leu Ile Glu Thr
290 295 300

Arg
305

<210> 225

<211> 308

<212> PRT

<213> Firmicutes bacterium CAG:24

<400> 225

Met Ile Ile Val Glu Val Met Gly Gly Leu Gly Asn Gln Met Gln Gln
1 5 10 15

Tyr Ala Leu Tyr Arg Lys Leu Glu Ser Leu Gly Lys Asp Ala Arg Leu
20 25 30

Asp Val Ser Trp Phe Leu Asp Lys Glu Arg Gln Thr Lys Val Leu Ala
35 40 45

Ser Arg Lys Leu Glu Leu Ser Trp Phe Glu Asn Leu Pro Ala Lys Tyr
50 55 60

Cys Thr Gln Glu Glu Lys Gln Ala Ile Leu Gly Lys Asn Asn Leu Ile
65 70 75 80

Gly Lys Leu Lys Lys Lys Leu Leu Gly Gly Ser Asn Arg His Phe Thr
85 90 95

37847-517 Sequence Listing_ST25 (003)

Glu Ser Asp Met Tyr His Pro Glu Ile Phe Asp Leu Glu Asp Ala Tyr
100 105 110

Leu Ser Gly Phe Trp Ala Cys Glu Ala Tyr Tyr Ala Asp Ile Leu Pro
115 120 125

Met Leu Arg Ser Gln Ile His Phe Pro Asp Pro Glu Lys Gly Glu Gly
130 135 140

Trp Asp Leu Glu Ala Ala Ala Lys Asn Lys Glu Thr Met Glu Arg Met
145 150 155 160

Lys Gln Glu Thr Ser Val Ser Ile His Ile Arg Arg Gly Asp Tyr Leu
165 170 175

Asp Ala Lys Asn Ala Glu Met Phe Gly Gly Ile Cys Thr Asp Ala Tyr
180 185 190

Tyr Glu Ala Ala Ile Ser Tyr Ile Lys Glu Gln Thr Pro Asp Ala His
195 200 205

Phe Tyr Val Phe Ser Asp Asp Ser Ala Tyr Val Lys Asn Ala Tyr Pro
210 215 220

Gly Lys Glu Phe Thr Val Val Asp Trp Asn Thr Gly Lys Asn Ser Leu
225 230 235 240

Phe Asp Met Gln Leu Met Ser Cys Cys Asn His Asn Ile Cys Ala Asn
245 250 255

Ser Thr Phe Ser Phe Trp Gly Ala Arg Leu Asn Pro Ser Pro Asp Lys
260 265 270

Val Met Ile Arg Pro Ser Lys His Lys Asn Ser Gln Asn Ile Val Pro
275 280 285

Glu Glu Met Lys Arg Leu Trp Asp Gly Trp Val Leu Ile Asp Gly Lys
290 295 300

37847-517 Sequence Listing_ST25 (003)

Gly Arg Ile Ile
305

<210> 226
<211> 283
<212> PRT
<213> Prevotella sp. CAG:474

<400> 226

Met Ile Ile Thr Lys Leu Asn Gly Gly Leu Gly Asn Gln Leu Phe Glu
1 5 10 15

Tyr Ala Cys Ala Arg Asn Leu Gln Leu Lys Tyr Asn Asp Val Leu Tyr
20 25 30

Leu Asp Ile Glu Gly Phe Lys Arg Ser Pro Arg His Tyr Ser Leu Glu
35 40 45

Lys Phe Lys Leu Ser Ser Asp Val Arg Met Leu Pro Glu Lys Asp Ser
50 55 60

Lys Ser Leu Ile Leu Leu Gln Ala Ile Ser Lys Leu Asn Arg Asn Leu
65 70 75 80

Ala Phe Lys Leu Gly Pro Leu Phe Gly Thr Tyr Ile Trp Lys Ser Ser
85 90 95

Asn Tyr Arg Pro Leu Lys Ile Lys Asn Thr Arg Gly Lys Lys Leu Tyr
100 105 110

Leu Tyr Gly Tyr Trp Gln Ser Tyr Glu Tyr Phe Lys Glu Asn Glu Ala
115 120 125

Ile Ile Lys Gln Glu Leu Asn Val Lys Thr Glu Ile Pro Ile Glu Cys
130 135 140

Ser Glu Leu Leu Lys Glu Ile Asn Lys Pro His Ser Ile Cys Val His
145 150 155 160

37847-517 Sequence Listing_ST25 (003)

Val Arg Arg Gly Asp Tyr Val Ser Cys Gly Phe Leu His Cys Asp Glu
165 170 175

Ala Tyr Tyr Asn Arg Gly Ile Asn His Ile Phe Asp Lys His Pro Asp
180 185 190

Ser Asn Val Val Val Phe Ser Asp Asp Ile Lys Trp Val Lys Ala Asn
195 200 205

Met Asn Phe Asp His Pro Val Ala Tyr Val Glu Val Asp Val Pro Asp
210 215 220

Tyr Glu Thr Leu Arg Leu Met Tyr Met Cys Lys His Phe Val Met Ser
225 230 235 240

Asn Ser Ser Phe Ser Trp Trp Ala Ser Tyr Leu Ser Asp Asn Lys Glu
245 250 255

Lys Ile Val Val Ala Pro Ser Tyr Trp Leu Pro Ala Asn Lys Asp Asn
260 265 270

Lys Ser Met Tyr Leu Asp Asn Trp Thr Ile Leu
275 280

<210> 227

<211> 298

<212> PRT

<213> Roseburia intestinalis

<400> 227

Met Arg Gly Asn Arg Gly Met Ile Ala Val Lys Ile Gly Asp Gly Met
1 5 10 15

Gly Asn Gln Leu Phe Asn Tyr Ala Cys Gly Tyr Ala Gln Ala Arg Arg
20 25 30

Asp Gly Asp Ser Leu Val Leu Asp Ile Ser Glu Cys Asp Asn Ser Thr
35 40 45

Leu Arg Asp Phe Glu Leu Asp Lys Phe His Leu Lys Tyr Asp Lys Lys

28 May 2018
2015259088

37847-517 Sequence Listing_ST25 (003)

50

55

60

Glu Ser Phe Pro Asn Arg Asn Leu Gly Gln Lys Ile Tyr Lys Asn Leu
65 70 75 80

Arg Arg Ala Leu Lys Tyr His Val Ile Lys Glu Arg Glu Val Tyr His
85 90 95

Asn Arg Asp His Arg Tyr Asp Val Asn Asp Ile Asp Pro Arg Val Tyr
100 105 110

Lys Lys Lys Gly Leu Arg Asn Lys Tyr Leu Tyr Gly Tyr Trp Gln His
115 120 125

Leu Ala Tyr Phe Glu Asp Tyr Leu Asp Glu Ile Thr Ala Met Met Thr
130 135 140

Pro Ala Tyr Glu Gln Ser Glu Thr Val Lys Lys Leu Gln Glu Glu Phe
145 150 155 160

Lys Lys Thr Pro Thr Cys Ala Val His Val Arg Gly Gly Asp Ile Met
165 170 175

Gly Pro Ala Gly Ala Tyr Phe Lys His Ala Met Glu Arg Met Glu Gln
180 185 190

Glu Lys Pro Gly Val Arg Tyr Ile Val Phe Thr Asn Asp Met Glu Arg
195 200 205

Ala Glu Glu Ala Leu Ala Pro Val Leu Glu Ser Gln Lys Lys Asp Ala
210 215 220

Val Gly Gln Ala Glu Asn Arg Leu Glu Phe Val Ser Glu Met Gly Glu
225 230 235 240

Phe Ser Asp Val Asp Glu Phe Phe Leu Met Ala Ala Cys Gln Asn Gln
245 250 255

Ile Leu Ser Asn Ser Thr Phe Ser Thr Trp Ala Ala Tyr Leu Asn Gln

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

260 265 270

Asn Pro Asp Lys Thr Val Ile Met Pro Asp Asp Leu Leu Ser Glu Arg
275 280 285

Met Arg Gln Lys Asn Trp Ile Ile Leu Lys
290 295

<210> 228
<211> 265
<212> PRT
<213> Bacteroides ovatus

<400> 228

Met Lys Ile Val Leu Phe Thr Pro Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Leu Phe Tyr Leu Tyr Leu Arg Asp Asn Tyr Pro Asn Gln Asn Ile
20 25 30

Tyr Gly Tyr Tyr Asn Arg Asn Ile Leu Asn Lys His Asn Gly Leu Glu
35 40 45

Val Asp Lys Val Phe Asp Ile Gln Leu Pro Pro His Thr Val Ile Ser
50 55 60

Asp Ala Ser Ala Phe Phe Ile Arg Ala Leu Gly Gly Leu Gly Leu Lys
65 70 75 80

Tyr Phe Ile Gly Lys Asp Gln Leu Ser Pro Trp Lys Val Tyr Phe Asp
85 90 95

Gly Tyr Trp Gln Asn Lys Glu Tyr Phe Gln Asn Asn Val Asp Lys Met
100 105 110

Arg Phe Arg Glu Gly Phe Leu Asn Lys Lys Asn Asp Asp Ile Leu Ser
115 120 125

Leu Ile Arg Asn Thr Asn Ser Val Ser Val His Val Arg Arg Gly Asp
130 135 140

37847-517 Sequence Listing_ST25 (003)

Tyr Cys Asp Ser Cys Arg Lys Asp Leu Phe Leu Gln Ser Cys Thr Pro
145 150 155 160

Gln Tyr Tyr Glu Ser Ala Ile Ser Val Met Lys Glu Lys Phe Gln Lys
165 170 175

Pro Val Phe Phe Val Phe Ser Asp Asp Ile Pro Trp Val Lys Val Asn
180 185 190

Leu Asn Ile Pro Asn Ala Tyr Tyr Ile Asp Trp Asn Lys Lys Glu Asn
195 200 205

Ser Tyr Leu Asp Met Tyr Leu Met Ser Leu Cys Thr Ala Ser Ile Ile
210 215 220

Ala Asn Ser Thr Phe Ser Phe Trp Gly Ala Met Leu Gly Asn Lys Lys
225 230 235 240

Glu Leu Val Ile Lys Pro Lys Lys Trp Ile Gly Asp Glu Ile Pro Glu
245 250 255

Ile Phe Pro Pro Ser Trp Leu Ser Leu
260 265

<210> 229

<211> 310

<212> PRT

<213> Butyrivibrio sp. AE3009

<400> 229

Met Leu Ile Ile Gln Ile Ala Gly Gly Leu Gly Asn Gln Met Gln Gln
1 5 10 15

Tyr Ala Leu Tyr Arg Lys Leu Leu Lys Tyr His Pro Asp Gly Val Arg
20 25 30

Leu Asp Leu Ser Trp Phe Asp Ser Glu Val Gln Lys Asn Met Leu Ala
35 40 45

37847-517 Sequence Listing_ST25 (003)

Lys Arg Glu Phe Glu Leu Ala Leu Phe Lys Gly Leu Pro Tyr Ile Glu
 50 55 60

Cys Lys Pro Glu Glu Arg Ala Ala Phe Leu Asp Arg Asn Ala Ala Gln
 65 70 75 80

Lys Leu Ser Gly Lys Val Leu Lys Lys Leu Gly Leu Arg Asp Asn Ala
 85 90 95

Asn Pro Asn Val Phe Glu Glu Ser Arg Met Phe His Pro Glu Ile Phe
 100 105 110

Glu Leu Asp Asn Lys Tyr Ile Ile Gly Tyr Phe Ala Cys Gln Lys Tyr
 115 120 125

Tyr Asp Asp Ile Met Gly Asp Leu Cys Asn Leu Phe Glu Phe Pro Glu
 130 135 140

His Leu Asp Pro Glu Leu Glu Lys Lys Asn Leu Glu Leu Ile Ser Lys
 145 150 155 160

Met Glu Lys Glu Asn Ser Val Ser Val His Ile Arg Arg Gly Asp Tyr
 165 170 175

Leu Asp Pro Glu Asn Phe Lys Ile Leu Gly Asn Ile Ala Thr Asp Glu
 180 185 190

Tyr Tyr Glu Ser Ala Met Lys Tyr Phe Glu Asp Arg Tyr Glu Lys Val
 195 200 205

His Phe Tyr Ile Phe Thr Ser Asp His Glu Tyr Ala Arg Glu His Phe
 210 215 220

Ala Asp Glu Ser Lys Tyr Thr Ile Val Asp Trp Asn Thr Gly Lys Asp
 225 230 235 240

Ser Leu Gln Asp Val Arg Leu Met Asn His Cys Leu Gly Asn Ile Cys
 245 250 255

37847-517 Sequence Listing_ST25 (003)

Ala Asn Ser Thr Phe Ser Phe Trp Gly Ala Arg Leu Asn Gln Arg Gln
260 265 270

Asp Lys Val Met Ile Arg Thr Tyr Lys Met Arg Asn Asn Gln Pro Val
275 280 285

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

Glu Thr Gly Arg Glu Val
305 310

<210> 230

<211> 300

<212> PRT

<213> Butyrivibrio proteoclasticus

<400> 230

Met Thr Lys Asn Glu Lys Lys Leu Ile Val Lys Phe Gln Gly Gly Leu
1 5 10 15

Gly Asn Gln Leu Tyr Glu Tyr Ala Phe Cys Glu Trp Leu Arg Gln Gln
20 25 30

Tyr Ser Asp Tyr Glu Val Leu Ala Asp Leu Ser Tyr Tyr Lys Ile Arg
35 40 45

Ser Ala His Gly Glu Leu Gly Ile Trp Asn Ile Phe Pro Asn Ile Asn
50 55 60

Ile Glu Val Ala Ser Asn Trp Asp Ile Ile Lys Tyr Ser Asp Gln Ile
65 70 75 80

Pro Ile Met Tyr Gly Gly Lys Gly Ala Asp Arg Leu Asn Ser Val Arg
85 90 95

Thr Asn Val Asn Asp Arg Phe Phe Ser Lys Arg Lys His Ser Tyr Tyr
100 105 110

37847-517 Sequence Listing_ST25 (003)

Thr Glu Ile Ser Asn Thr Asp Val Ser Glu Val Ile Asn Ala Leu Asn
115 120 125

Asn Gly Ile Arg Tyr Phe Asp Gly Tyr Trp Gln Asn Ile Asp Tyr Phe
130 135 140

Lys Gly Asn Ile Glu Asp Leu Arg Asn Lys Leu Lys Phe Ser Glu Lys
145 150 155 160

Cys Asp Lys Tyr Ile Thr Asp Glu Met Leu Arg Asp Asn Ala Val Ser
165 170 175

Leu His Val Arg Arg Gly Asp Tyr Val Gly Ser Glu Tyr Glu Lys Glu
180 185 190

Val Gly Leu Ser Tyr Tyr Lys Lys Ala Val Glu Tyr Val Leu Asp Arg
195 200 205

Val Asp Gln Ala Lys Phe Phe Ile Phe Ser Asp Asp Lys Tyr Tyr Ala
210 215 220

Glu Thr Ala Phe Glu Trp Ile Asp Asn Lys Thr Val Val Ala Gly Tyr
225 230 235 240

Asp Asn Glu Leu Ala His Val Asp Met Leu Leu Met Ser Arg Met Lys
245 250 255

Asn Asn Ile Ile Ala Asn Ser Thr Phe Ser Leu Trp Ala Ala Tyr Leu
260 265 270

Asn Asp Ser Met Asn Pro Leu Ile Val Tyr Pro Asp Val Glu Ser Leu
275 280 285

Asp Lys Lys Thr Phe Ser Asp Trp Asn Gly Ile Lys
290 295 300

<210> 231
<211> 292
<212> PRT
<213> Prevotella nanceiensis

37847-517 Sequence Listing_ST25 (003)

<400> 231

Met Asp Ser Gln Phe Leu Lys His Ile Lys Leu Ser Gly Gly Phe Gly
1 5 10 15

Asn Gln Leu Phe Gln Tyr Phe Phe Gly Glu Tyr Leu Lys Glu Lys Tyr
20 25 30

Asn Cys Ser Ile Ser Phe Phe Ser Glu Pro Ala Leu Asp Ile Asn Gln
35 40 45

Leu Gln Ile His Arg Phe Phe Pro Ala Leu Arg Ile Ser His Asn Thr
50 55 60

Glu Leu Arg Pro Tyr His Tyr Ser Phe Thr Gln Gln Leu Ala Tyr Arg
65 70 75 80

Cys Met Arg Lys Leu Leu Leu Leu Phe Pro Phe Leu Asn Arg Lys Val
85 90 95

Lys Ile Glu Asn Gly Ser Asn Tyr Gln Asn Gln Ser Phe Asn Asp Thr
100 105 110

Tyr Cys Phe Asp Gly Tyr Trp Gln Ser Tyr Arg Tyr Leu Ser Ala Phe
115 120 125

Thr Pro Ser Leu Gln Phe Glu Asp Gln Leu Ile Asn Asp Ile Ser Ala
130 135 140

Asp Tyr Ile Asn Ala Ile Glu Gln Ser Glu Ala Val Phe Leu His Ile
145 150 155 160

Arg Arg Gly Asp Tyr Leu Asn Lys Glu Asn Gln Lys Val Phe Ala Glu
165 170 175

Cys Pro Leu Asn Tyr Phe Glu Asn Ala Ala Asn Arg Ile Lys Glu Asp
180 185 190

Ile Lys Asn Val His Phe Phe Val Phe Ser Asn Asp Ile Gln Trp Val

37847-517 Sequence Listing_ST25 (003)

195

200

205

Lys Ser His Leu Lys Leu Asn Asp Asn Glu Val Thr Phe Ile Gln Asn
210 215 220

Glu Gly Asn Ser Cys Asp Leu Lys Asp Phe Tyr Leu Met Thr Arg Cys
225 230 235 240

Lys His Ala Ile Ile Ser Asn Ser Thr Phe Ser Trp Trp Ala Ala Tyr
245 250 255

Leu Ile Asn Asn Ser Asp Lys Lys Val Ile Ala Pro Lys His Trp Tyr
260 265 270

Asn Asp Ile Ser Met Asn Asn Ala Thr Lys Asp Leu Ile Pro Pro Thr
275 280 285

Trp Ile Arg Leu
290

<210> 232

<211> 285

<212> PRT

<213> Bacteroides ovatus

<400> 232

Met Ile Ile Thr Arg Leu His Gly Arg Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Ala Gly Arg Ala Leu Ala Asp Arg Ala Gly Val Pro Leu Ala
20 25 30

Leu Asp Ser Arg Gly Ala Ile Leu Arg Gly Glu Gly Val Leu Thr Arg
35 40 45

Val Phe Asp Leu Glu Leu Ala Asp Pro Val His Leu Pro Pro Leu Lys
50 55 60

Gln Thr Asn Pro Leu Arg Tyr Ala Ile Trp Arg Gly Ile Gly Gln Lys
65 70 75 80

37847-517 Sequence Listing_ST25 (003)

Val Gly Ala Lys Pro Tyr Phe Arg Arg Glu Arg Gly Leu Gly Tyr Asn
85 90 95

Pro Ala Phe Glu Asp Trp Gly Asp Asn Ser Tyr Leu His Gly Tyr Trp
100 105 110

Gln Ser Gln Lys Tyr Phe Gln Asn Ser Ala Glu Arg Ile Arg Ser Asp
115 120 125

Phe Thr Phe Pro Ala Phe Ser Asn Gln Gln Asn Ala Glu Met Ala Ala
130 135 140

Arg Ile Ala Glu Ser Thr Ala Ile Ser Leu His Val Arg Arg Gly Asp
145 150 155 160

Tyr Leu Thr Phe Ala Ala His Val Leu Cys Asp Gln Ala Tyr Tyr Asp
165 170 175

Ala Ala Leu Ala Lys Val Leu Asp Gly Leu Gln Gly Asp Pro Ile Val
180 185 190

Tyr Val Phe Ser Asp Asp Pro Gln Trp Ala Lys Asp Asn Leu Ser Leu
195 200 205

Pro Cys Glu Lys Val Val Val Asp Phe Asn Gly Pro Glu Thr Asp Phe
210 215 220

Glu Asp Met Arg Leu Met Ser Leu Cys Gln His Asn Ile Ile Gly Asn
225 230 235 240

Ser Ser Phe Ser Trp Trp Ala Ala Trp Leu Asn Gln Thr Pro Gly Arg
245 250 255

Arg Val Ala Gly Pro Ala Lys Trp Phe Gly Asp Pro Lys Leu Ser Asn
260 265 270

Pro Asp Ile Phe Pro His Asp Trp Leu Arg Ile Ser Val
275 280 285

37847-517 Sequence Listing_ST25 (003)

28 May 2018
2015259088

<210> 233
 <211> 283
 <212> PRT
 <213> Winogradskyella psychrotolerans

<400> 233

Met Gly Asn Gln Leu Tyr Glu Tyr Ala Thr Ala Lys Ala Met Ala Val
 1 5 10 15

Ala Leu Asn Lys Lys Leu Val Ile Asp Pro Arg Pro Ile Leu Lys Glu
 20 25 30

Ala Pro Gln Arg His Tyr Asp Leu Gly Leu Phe Asn Ile Gln Asp Glu
 35 40 45

Asp Phe Gly Ser Pro Phe Val Gln Trp Leu Val Arg Trp Val Ala Ser
 50 55 60

Val Arg Leu Gly Lys Phe Phe Lys Thr Ile Met Pro Phe Ala Trp Ser
 65 70 75 80

Tyr Gln Met Ile Arg Asp Lys Glu Glu Gly Phe Asp Glu Ser Leu Leu
 85 90 95

Gln Gln Lys Ser Arg Asn Ile Val Ile Glu Gly Tyr Trp Gln Ser Phe
 100 105 110

Lys Tyr Phe Glu Ser Ile Arg Pro Thr Leu Leu Lys Glu Leu Ser Phe
 115 120 125

Lys Asp Lys Pro Asn Ala Ile Asn Gln Lys Tyr Leu Asp Glu Ile Glu
 130 135 140

Ser Val Asn Ala Val Ala Val His Ile Arg Arg Gly Asp Tyr Val Ala
 145 150 155 160

Asn Pro Val Ala Asn Ala Val His Gly Leu Cys Asp Met Asp Tyr Tyr
 165 170 175

37847-517 Sequence Listing_ST25 (003)

Lys Lys Ala Ile Ala Ile Ile Lys Asp Lys Val Glu Asn Pro Tyr Phe
180 185 190

Phe Ile Phe Thr Asp Asp Pro Asp Trp Ala Glu Asp Asn Phe Lys Ile
195 200 205

Ser Glu His Gln Lys Ile Ile Lys His Asn Ile Gly Lys Gln Asp His
210 215 220

Glu Asp Phe Arg Leu Leu Thr Asn Cys Lys Tyr Phe Ile Ile Ala Asn
225 230 235 240

Ser Ser Phe Ser Trp Trp Gly Ala Trp Leu Ser Asp Tyr Lys Asn Lys
245 250 255

Ile Val Ile Ser Pro Asn Lys Trp Phe Asn Val Asp Ala Val Pro Ile
260 265 270

Thr Glu Arg Ile Pro Glu Ser Trp Ile Arg Val
275 280

<210> 234
<211> 302
<212> PRT
<213> Lachnospiraceae bacterium NK4A179

<400> 234

Met Ile Thr Val Arg Ile Asp Gly Gly Phe Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Phe Phe Leu His Leu Lys Lys Thr Ile Thr Asp Asn Lys Ile
20 25 30

Ser Val Asp Leu Asn Cys Tyr Asn Pro His Gly Ser Gly Asp Ile Phe
35 40 45

Thr Arg Phe Lys Leu Ala Pro Glu Gln Ala Ala Pro Ser Glu Ile Lys
50 55 60

37847-517 Sequence Listing_ST25 (003)

Arg	Phe	His	Arg	Asn	Ser	Ile	Tyr	His	Leu	Leu	Arg	Pro	Leu	Asp	Ser	65	70	75	80
Ala	Gly	Ile	Thr	Thr	Asn	Pro	Tyr	Tyr	Arg	Glu	Glu	Asp	Ile	Asp	Asp	85	90	95	
Leu	Asn	Ser	Val	Leu	Asn	Lys	Lys	Arg	Val	Tyr	Leu	Arg	Gly	Tyr	Trp	100	105	110	
Gln	Asp	Lys	Arg	Tyr	Pro	Phe	Ser	Val	Lys	Asp	Gln	Leu	Ile	Asp	Cys	115	120	125	
Phe	Asp	Leu	Gly	Lys	Met	Asp	Met	Thr	Gly	Ala	Ser	Ala	Glu	Asn	Asn	130	135	140	
Val	Ile	Leu	Glu	Gln	Ile	Ala	Ser	Glu	Glu	Ser	Arg	Ser	Val	Gly	Val	145	150	155	160
His	Leu	Arg	Gly	Gly	Asp	Tyr	Ile	Gly	Asp	Pro	Val	Tyr	Ser	Gly	Ile	165	170	175	
Cys	Thr	Pro	Glu	Tyr	Tyr	Glu	Ala	Ala	Phe	Lys	His	Val	Ser	Glu	Lys	180	185	190	
Ile	Lys	Asp	Pro	Val	Phe	His	Ile	Phe	Thr	Asn	Asp	Ile	Ser	Met	Ile	195	200	205	
Glu	Lys	Cys	Gly	Leu	Ser	Gly	Lys	Tyr	Asp	Leu	Lys	Ile	Thr	Asp	Ile	210	215	220	
Asn	Asp	Glu	Ala	His	Gly	Trp	Ala	Asp	Leu	Lys	Leu	Met	Ser	Ala	Cys	225	230	235	240
Arg	His	His	Ile	Ile	Ser	Asn	Ser	Ser	Phe	Ser	Trp	Trp	Ala	Ala	Phe	245	250	255	
Leu	Gly	Glu	Ala	Thr	Thr	Glu	Ala	Ser	Ala	Asp	Val	Ile	Asn	Val	Ile	260	265	270	

37847-517 Sequence Listing_ST25 (003)

Pro Glu Tyr Met Arg Gln Gly Val Ser Ala Glu Thr Leu Arg Cys Pro
275 280 285

Cys Trp Thr Thr Val Thr Ser Asp Gly Arg Val Tyr Pro Ser
290 295 300

<210> 235

<211> 263

<212> PRT

<213> Prevotella sp. oral taxon 317

<400> 235

Met Lys Ile Val Cys Ile Lys Gly Gly Leu Gly Asn Gln Leu Phe Glu
1 5 10 15

Tyr Cys Arg Tyr Arg Ser Leu His Arg His Asp Asn Arg Gly Val Tyr
20 25 30

Leu His Tyr Asp Arg Arg Arg Thr Lys Gln His Gly Gly Val Trp Leu
35 40 45

Asp Lys Ala Phe His Ile Thr Leu Pro Asn Glu Pro Leu Arg Val Lys
50 55 60

Leu Leu Val Met Val Leu Lys Thr Leu Arg Arg Leu His Leu Phe Lys
65 70 75 80

Arg Leu Tyr Arg Glu Glu Asp Pro Arg Ala Val Leu Ile Asp Asp Tyr
85 90 95

Ser Gln His Lys Gln Tyr Ile Thr Asn Ala Ala Glu Ile Leu Asn Phe
100 105 110

Arg Pro Phe Glu Gln Leu Asp Tyr Ala Glu Glu Ile Gln Thr Thr Pro
115 120 125

Phe Ala Val Ser Val His Val Arg Arg Gly Asp Tyr Leu Leu Leu Ala
130 135 140

Asn Lys Ser Asn Phe Gly Val Cys Ser Val His Tyr Tyr Leu Ser Ala

28 May 2018
2015259088

37847-517 Sequence Listing_ST25 (003)

145 150 155 160

Ala Val Ala Val Arg Glu Arg His Pro Glu Ser Arg Phe Phe Val Phe
165 170 175

Ser Asp Asp Met Glu Trp Ala Lys Glu Asn Leu Asn Leu Pro Asn Cys
180 185 190

Val Phe Val Glu His Ala Gln Ala Gln Pro Asp His Ala Asp Leu Tyr
195 200 205

Leu Met Ser Leu Cys Lys Gly His Ile Ile Ala Asn Ser Thr Phe Ser
210 215 220

Phe Trp Gly Ala Tyr Leu Ser Lys Gly Ser Ser Ala Ile Ala Ile Tyr
225 230 235 240

Pro Lys Gln Trp Phe Ala Glu Pro Thr Trp Asn Val Pro Asp Ile Phe
245 250 255

Pro Ala His Trp Met Ala Leu
260

<210> 236
<211> 309
<212> PRT
<213> Butyrivibrio sp. XPD2006

<400> 236

Met Leu Ile Ile Gln Ile Ala Gly Gly Leu Gly Asn Gln Met Gln Gln
1 5 10 15

Tyr Ala Val Tyr Thr Lys Leu Arg Gly Met Gly Lys Asp Val Arg Leu
20 25 30

Asp Leu Ser Trp Phe Asp Pro Ser Val Gln Lys Asn Met Leu Ala Pro
35 40 45

Arg Glu Phe Glu Leu Ser Met Phe Glu Gly Val Asp Tyr Thr Glu Cys
50 55 60

37847-517 Sequence Listing_ST25 (003)

Thr Ala Glu Glu Arg Asp Ser Phe Leu Lys Gln Gly Met Ile Ala Asn
65 70 75 80

Val Thr Gly Lys Met Leu Lys Lys Leu Gly Leu Arg Asp Glu Ala Asn
85 90 95

Pro Lys Val Phe Ser Glu Lys Glu Met Tyr His Pro Glu Ile Phe Glu
100 105 110

Leu Glu Asp Arg Tyr Ile Lys Gly Tyr Phe Ala Cys Gln Lys Tyr Tyr
115 120 125

Asp Asp Ile Met Gly Glu Leu Trp Glu Lys Tyr Thr Phe Pro Ala His
130 135 140

Ser Asp Pro Asp Leu His Thr Arg Asn Met Ala Leu Val Glu Arg Met
145 150 155 160

Glu Lys Glu Thr Ser Val Ser Val His Ile Arg Arg Gly Asp Tyr Leu
165 170 175

Asp Pro Ser Asn Val Glu Ile Leu Gly Asn Ile Ala Thr Glu Glu Tyr
180 185 190

Tyr Gln Gly Ala Met Asp Tyr Phe Ser Val Lys Asp Pro Asp Thr His
195 200 205

Phe Tyr Ile Phe Thr Ser Asp His Glu Tyr Ala Arg Glu Lys Phe Ser
210 215 220

Asp Glu Ser Lys Tyr Thr Ile Val Asp Trp Asn Ser Gly Arg Asn Ser
225 230 235 240

Val Gln Asp Leu Met Leu Met Ser His Cys Lys Gly Asn Ile Cys Ala
245 250 255

Asn Ser Thr Phe Ser Phe Trp Gly Ala Arg Leu Asn Arg Arg Pro Asp
260 265 270

37847-517 Sequence Listing_ST25 (003)

Lys Thr Val Ile Arg Thr Tyr Lys Met Arg Asn Asn Gln Pro Val Asn
275 280 285

Pro Asp Ile Met His Asp Tyr Trp Lys Gly Trp Ile Leu Met Asp Glu
290 295 300

Lys Gly Ser Ile Ile
305

<210> 237

<211> 317

<212> PRT

<213> Butyrivibrio fibrisolvens

<400> 237

Met Ile Ile Ile Lys Leu Gln Gly Gly Leu Gly Asn Gln Leu Phe Leu
1 5 10 15

Tyr Gly Leu Tyr Lys Asn Leu Lys His Leu Lys Arg Asp Val Lys Met
20 25 30

Asp Ile Glu Ser Gly Phe Glu Glu Asp Lys Leu Arg Val Pro Cys Leu
35 40 45

Lys Ser Met Gly Leu Asp Tyr Glu Val Ala Thr Arg Asp Glu Ile Val
50 55 60

Ala Ile Arg Asp Ser Tyr Met Asp Ile Phe Ser Arg Ile Arg Arg Lys
65 70 75 80

Ile Thr Gly Arg Lys Thr Phe Asp Tyr Tyr Glu Pro Glu Asp Gly Asn
85 90 95

Phe Asp Pro Arg Val Leu Glu Gln Thr His Ala Tyr Leu Asp Gly Tyr
100 105 110

Phe Gln Ser Glu Lys Tyr Phe Gly Asp Ser Asp Asp Arg Lys Lys Leu
115 120 125

37847-517 Sequence Listing_ST25 (003)

Lys Asp Glu Leu Leu Lys Glu Lys Ile Arg Val Leu Asp Ser Ser Asp
130 135 140

Thr Leu Lys Asp Leu Tyr Asn Met Met Ser Ser Gly Ser Ser Val Ser
145 150 155 160

Leu His Ile Arg Arg Gly Asp Tyr Leu Thr Pro Gly Ile Met Glu Thr
165 170 175

Tyr Gly Gly Ile Cys Thr Asp Glu Tyr Tyr Asp Ile Ala Met Asn Arg
180 185 190

Ile Lys Asn Glu Tyr Pro Asp Ser Lys Phe Phe Ile Phe Ser Asn Asp
195 200 205

Ile Asp Trp Cys Lys Glu Lys Tyr Gly Ser Arg Asp Asp Val Ile Phe
210 215 220

Val Asp Ser Cys Asp Glu His Glu Gly Leu Thr Asn Val Ser Gly Asp
225 230 235 240

Gln Asp Asp Ile Gln Val Gln Gly Asp Ile Lys Glu His Gly Asn Asn
245 250 255

Ser Leu Arg Asp Ala Ala Glu Leu Tyr Leu Met Ser Ala Cys Lys His
260 265 270

His Ile Leu Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala Trp Leu Ser
275 280 285

Asp His Glu Gly Met Thr Ile Ala Pro Ser Lys Trp Leu Asn Asn Lys
290 295 300

Asn Met Thr Asp Ile Tyr Thr Lys Asp Met Leu Leu Ile
305 310 315

<210> 238

<211> 297

<212> PRT

37847-517 Sequence Listing_ST25 (003)

<213> *Cylindrospermopsis raciborskii*

<400> 238

Met Lys Lys Thr Val Val Leu Leu Lys Gly Gly Leu Gly Asn Gln Met
1 5 10 15

Phe Gln Tyr Ala Phe Ala Arg Ser Ile Ser Leu Lys Asn Ser Ser Lys
20 25 30

Leu Val Ile Asp Asn Trp Ser Gly Phe Thr Phe Asp Tyr Lys Tyr His
35 40 45

Arg Gln Tyr Glu Leu Gly Thr Phe Ser Ile Val Gly Arg Pro Ala Asn
50 55 60

Leu Thr Glu Lys Phe Pro Phe Trp Phe Tyr Glu Leu Lys Ser Lys Phe
65 70 75 80

Phe Pro Arg Leu Pro Lys Val Phe Gln Gln Gln Phe Tyr Gly Leu Leu
85 90 95

Ile Asn Glu Val Gly Gly Glu Tyr Ile Pro Glu Ile Glu Glu Thr Lys
100 105 110

Ile Ser Gln Asn Cys Trp Leu Asn Gly Tyr Trp Gln Ser Pro Leu Tyr
115 120 125

Phe Gln Lys His Ser Asp Ser Ile Thr Arg Glu Leu Met Pro Pro Glu
130 135 140

Pro Met Glu Lys His Phe Leu Glu Leu Gly Lys Leu Leu Arg Glu Thr
145 150 155 160

Glu Ser Val Ala Leu Gly Ile Arg Leu Tyr Glu Glu Ser Lys Asn Pro
165 170 175

Gly Ser His Ser Ser Ser Gly Glu Leu Lys Ser His Phe Glu Ile Asn
180 185 190

37847-517 Sequence Listing_ST25 (003)

Gln Ala Ile Leu Lys Leu Arg Glu Leu Cys Asn Gly Ala Lys Phe Phe
195 200 205

Val Phe Cys Thr His Arg Ser Pro Leu Leu Gln Glu Leu Ala Leu Pro
210 215 220

Glu Asn Thr Ile Phe Val Thr His Asp Asp Gly Tyr Val Gly Ser Met
225 230 235 240

Glu Arg Met Trp Leu Leu Thr Gln Cys Lys His His Ile Phe Thr Asn
245 250 255

Ser Thr Phe Tyr Trp Trp Gly Ala Trp Leu Ser Gln Lys Phe Tyr Ile
260 265 270

Gln Gly Ser Gln Ile Val Phe Ala Ala Asp Asn Phe Ile Asn Ser Asp
275 280 285

Ala Ile Pro Lys His Trp Lys Pro Phe
290 295

<210> 239
<211> 273
<212> PRT
<213> Prevotella multiformis

<400> 239

Met Lys Ile Val Asn Phe Gln Gly Gly Leu Gly Asn Gln Met Phe Ile
1 5 10 15

Tyr Ala Phe Ser Arg Tyr Leu Ser Arg Leu Tyr Pro Gln Glu Lys Ile
20 25 30

Tyr Gly Ser Tyr Trp Ser Arg Ser Leu Tyr Val His Ser Ala Phe Gln
35 40 45

Leu Asp Arg Ile Phe Ser Leu Gln Leu Pro Pro His Asn Leu Phe Thr
50 55 60

Asp Cys Ile Ser Lys Leu Ala Arg Phe Phe Glu Arg Leu Arg Leu Val

2015259088 28 May 2018

2015259088 28 May 2018

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

<210> 240
 <211> 290
 <212> PRT
 <213> Bacteroides sp. CAG:462

 <400> 240

 Met Ile Ile Val Asn Met Ala Cys Gly Leu Ala Asn Arg Met Phe Gln
 1 5 10 15

 Tyr Ala Phe Tyr Leu Ser Leu Lys Glu Arg Gly Tyr Asn Val Lys Val
 20 25 30

 Asp Phe Tyr Lys Ser Ala Thr Leu Pro His Glu Asn Val Pro Trp Asn
 35 40 45

 Asp Ile Phe Pro Tyr Ala Glu Ile Asp Gln Val Ser Asn Phe Arg Val
 50 55 60

 Leu Ile Leu Gly Gly Gly Ala Asn Leu Leu Ser Lys Leu Arg Arg Lys
 65 70 75 80

 Tyr Leu Pro Ser Leu Thr Asn Val Ile Thr Met Ser Thr Ala Phe Asp
 85 90 95

 Thr Asp Leu Gln Ile Asp Asp Asp Arg Lys Asp Lys Tyr Ile Ile Gly
 100 105 110

 Val Phe Gln Ser Ala Ala Met Val Glu Gly Val Cys Lys Lys Val Lys
 115 120 125

 Gln Cys Phe Ser Phe Leu Pro Phe Thr Asp Leu Arg His Leu Gln Leu
 130 135 140

 Glu Lys Glu Met Gln Glu Cys Glu Ser Val Ala Ile His Val Arg Lys
 145 150 155 160

 Gly Asn Asp Tyr Gln Gln Arg Ile Trp Tyr Gln Asn Thr Cys Thr Met
 165 170 175

37847-517 Sequence Listing_ST25 (003)

Asp Tyr Tyr Arg Lys Ala Ile Ala Glu Ile Lys Gly Lys Val Lys Asp
180 185 190

Pro Arg Phe Tyr Val Phe Thr Asp Asn Ala Asp Trp Val Arg Arg Asn
195 200 205

Phe Thr Asp Phe Asp Tyr Lys Met Val Glu Gly Asn Pro Val Tyr Gly
210 215 220

Trp Gly Ser His Phe Asp Met Gln Leu Met Ser Arg Cys Lys Tyr Asn
225 230 235 240

Ile Ile Ser Asn Ser Thr Tyr Ser Trp Trp Gly Ala Tyr Leu Asn Ala
245 250 255

Asn Arg Asn Lys Ile Val Ile Cys Pro Asn Ile Trp Phe Asn Pro Glu
260 265 270

Ser Cys Asn Glu Tyr Thr Ser Cys Lys Leu Leu Cys Lys Gly Trp Ile
275 280 285

Ala Leu
290

<210> 241
<211> 557
<212> PRT
<213> Desulfovibrio africanus

<400> 241

Met Arg Ile Gly Ile Leu Tyr Ile Cys Thr Gly Lys Tyr Thr Val Phe
1 5 10 15

Trp Asn His Phe Phe Thr Ser Cys Glu Gln His Phe Leu Arg Glu His
20 25 30

Glu Lys His Tyr Tyr Ile Phe Thr Asp Gly Glu Ile Ala His Leu Asn
35 40 45

37847-517 Sequence Listing_ST25 (003)

Cys	Asn	Arg	Val	His	Arg	Ile	Glu	Gln	Gln	His	Leu	Gly	Trp	Pro	Asp	50	55	60	
Ser	Thr	Leu	Lys	Arg	Phe	His	Met	Phe	Glu	Arg	Ile	Ala	Asp	Thr	Leu	65	70	75	80
Arg	Gln	Asn	Ser	Asp	Phe	Ile	Val	Phe	Phe	Asn	Ala	Asn	Met	Val	Phe	85	90	95	
Leu	Arg	Asp	Val	Gly	Lys	Glu	Phe	Leu	Pro	Thr	Arg	Glu	Gln	Ala	Leu	100	105	110	
Val	Phe	His	Arg	His	Pro	Gly	Leu	Phe	Arg	Arg	Pro	Ala	Trp	Leu	Leu	115	120	125	
Pro	Tyr	Glu	Arg	Arg	Pro	Glu	Ser	Thr	Ala	Tyr	Ile	Pro	Tyr	Gly	Ser	130	135	140	
Gly	Ser	Ile	Tyr	Val	Cys	Gly	Gly	Val	Asn	Gly	Gly	Tyr	Thr	Gln	Pro	145	150	155	160
Tyr	Leu	Asp	Phe	Val	Ala	Met	Leu	Arg	Arg	Asn	Ile	Asp	Ile	Asp	Val	165	170	175	
Glu	Arg	Gly	Ile	Ile	Ala	Arg	Trp	His	Asp	Glu	Ser	His	Ile	Asn	Arg	180	185	190	
Phe	Val	Ile	Gly	Arg	His	Tyr	Lys	Ile	Gly	His	Pro	Gly	Tyr	Val	Tyr	195	200	205	
Pro	Asp	Arg	Arg	Asn	Leu	Pro	Phe	Pro	Arg	Ile	Ile	Arg	Val	Ile	Asp	210	215	220	
Lys	Ala	Ser	Val	Gly	Gly	His	Thr	Phe	Leu	Arg	Gly	Gln	Thr	Pro	Glu	225	230	235	240
Pro	Ala	Pro	Glu	Glu	Gln	Ser	Lys	Thr	Val	Ala	Lys	Lys	Leu	Arg	Ser	245	250	255	

37847-517 Sequence Listing_ST25 (003)

Gln Leu Lys Arg Pro Cys Met Pro Arg Ala Ala Gln Asp Glu Pro Ile
260 265 270

Ile Leu Ala Arg Met Met Gly Gly Leu Gly Asn Gln Met Phe Ile Tyr
275 280 285

Ala Ala Ala Arg Val Leu Ala Glu Arg Gln Gly Ala Gln Leu His Leu
290 295 300

Asp Thr Gly Lys Leu Ser Gly Asp Ser Ile Arg Gln Tyr Asp Leu Pro
305 310 315 320

Ala Phe Ser Ile Asp Ala Pro Leu Trp His Ile Pro Cys Gly Cys Asp
325 330 335

Arg Ile Val Gln Ala Trp Phe Ala Leu Arg His Val Ala Ala Gly Cys
340 345 350

Gly Met Pro Lys Pro Thr Met Gln Val Leu Arg Ser Gly Phe His Leu
355 360 365

Asp Gln Arg Phe Phe Ser Ile Arg His Ser Ala Tyr Leu Ile Gly Tyr
370 375 380

Trp Gln Ser Pro His Tyr Trp Arg Gly His Glu Asp Arg Val Arg Ser
385 390 395 400

Ser Phe Asp Leu Thr Arg Phe Glu Arg Pro His Leu Arg Glu Ala Leu
405 410 415

Ala Ala Val Ser Gln Pro Asn Thr Ile Ser Val His Leu Arg Arg Gly
420 425 430

Asp Phe Arg Ala Pro Lys Asn Ser Asp Lys His Leu Leu Ile Asp Gly
435 440 445

Ser Tyr Tyr Glu Arg Ala Arg Lys Leu Leu Leu Glu Met Thr Pro Gln
450 455 460

37847-517 Sequence Listing_ST25 (003)

Ser His Phe Tyr Ile Phe Ser Asp Glu Pro Glu Glu Ala Gln Arg Leu
465 470 475 480

Phe Ala His Trp Glu Asn Thr Ser Phe Gln Pro Arg Arg Ser Gln Glu
485 490 495

Glu Asp Leu Leu Leu Met Ser Arg Cys Ser Ala Ser Ile Ile Ala Asn
500 505 510

Ser Ser Phe Ser Trp Trp Gly Ala Trp Leu Gly Arg Pro Lys Gln His
515 520 525

Val Ile Ala Pro Arg Met Trp Phe Thr Arg Asp Val Leu Met His Thr
530 535 540

Tyr Thr Leu Asp Leu Phe Pro Glu Lys Trp Ile Leu Leu
545 550 555

<210> 242

<211> 299

<212> PRT

<213> Roseburia sp. CAG:100

<400> 242

Met Ile Leu Ile His Val Met Gly Gly Leu Gly Asn Gln Leu Tyr Gln
1 5 10 15

Tyr Ala Leu Tyr Glu Lys Met Lys Ser Leu Gly Lys Lys Val Lys Leu
20 25 30

Asp Thr Tyr Ala Tyr Asn Asp Ala Ala Gly Glu Asp Lys Glu Trp Arg
35 40 45

Ser Leu Glu Leu Asp Arg Phe Pro Ala Ile Glu Tyr Asp Lys Ala Thr
50 55 60

Ser Glu Asp Arg Thr Lys Leu Leu Asp Asn Ser Gly Leu Leu Thr Ala
65 70 75 80

37847-517 Sequence Listing_ST25 (003)

Lys	Ile	Arg	Arg	Lys	Leu	Leu	Gly	Arg	Lys	Asp	Lys	Thr	Ile	Arg	Glu	85	90	95	
Ser	Lys	Glu	Tyr	Met	Pro	Glu	Ile	Phe	His	Met	Asp	Asp	Val	Tyr	Leu	100	105	110	
Tyr	Gly	Phe	Trp	Asn	Cys	Glu	Arg	Tyr	Tyr	Glu	Asp	Ile	Ile	Pro	Leu	115	120	125	
Leu	Gln	Asp	Lys	Leu	Gln	Phe	Pro	Ile	Ser	Asn	Asn	Pro	Arg	Asn	Gln	130	135	140	
Gln	Cys	Ile	Glu	Gln	Met	Gln	Lys	Glu	Asn	Ala	Val	Ser	Ile	His	Ile	145	150	155	160
Arg	Arg	Thr	Asp	Tyr	Leu	Thr	Val	Ala	Asp	Gly	Ala	Arg	Tyr	Met	Gly	165	170	175	
Ile	Cys	Thr	Glu	Asp	Tyr	Tyr	Lys	Gly	Ala	Met	Ala	Tyr	Ile	Glu	Glu	180	185	190	
Arg	Val	Ser	Asn	Pro	Val	Tyr	Tyr	Ile	Phe	Ser	Asp	Asp	Val	Glu	Tyr	195	200	205	
Ala	Lys	Gln	His	Tyr	His	Gln	Asp	Asn	Met	His	Val	Val	Asp	Trp	Asn	210	215	220	
Ser	Lys	Ala	Asp	Ser	Ile	Tyr	Asp	Met	Gln	Leu	Met	Ser	Lys	Cys	Lys	225	230	235	240
His	Asn	Ile	Cys	Ala	Asn	Ser	Thr	Phe	Ser	Met	Trp	Ala	Ala	Arg	Leu	245	250	255	
Asn	Gln	Asn	Lys	Glu	Lys	Ile	Met	Ile	Arg	Pro	Leu	His	His	Asp	Asn	260	265	270	
Tyr	Glu	Thr	Thr	Thr	Ala	Thr	Gln	Val	Lys	Gln	Asn	Trp	Lys	Asn	Trp	275	280	285	

37847-517 Sequence Listing_ST25 (003)

Ile Leu Leu Asp Gln Asn Gly Gln Val Cys Glu
290 295

<210> 243
<211> 326
<212> PRT
<213> Desulfovibrio africanus

<400> 243

Met Thr Met Asn Ile Ile Arg Met Ser Gly Gly Leu Gly Ser Gln Met
1 5 10 15

Phe Gln Tyr Ala Leu Tyr Leu Lys Leu Lys Ser Met Gly Lys Glu Val
20 25 30

Lys Phe Asp Asp Ile Asn Glu Tyr Arg Gly Glu Lys Ala Arg Pro Ile
35 40 45

Met Leu Ala Val Phe Gly Ile Glu Tyr Pro Arg Ala Thr Trp Asp Glu
50 55 60

Ile Thr Ser Phe Thr Asp Gly Ser Met Asp Leu Leu Lys Arg Leu Arg
65 70 75 80

Arg Lys Ile Phe Gly Arg Lys Ala Ile Glu Tyr Glu Glu Gln Gly Phe
85 90 95

Tyr Asp Pro Asn Val Leu Asn Phe Asp Ser Met Tyr Leu Arg Gly Asn
100 105 110

Phe Gln Ser Glu Lys Tyr Phe Gln Asp Ile Lys Glu Glu Val Arg Lys
115 120 125

Leu Tyr Arg Phe Ser Thr Leu Glu Asp Met Arg Leu Pro Glu Arg Leu
130 135 140

Tyr Lys Ala Thr Lys Ala Cys Leu Asp Gly Ile Glu Ser Ser Glu Ser
145 150 155 160

Val Gly Leu His Met Tyr Arg Ser Asp Ser Arg Val Asp Gly Glu Leu

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

165 170 175

Tyr Asp Gly Ile Cys Thr Gly Asn Tyr Tyr Lys Gly Ala Val Arg Phe
180 185 190

Ile Gln Asp Lys Val Pro Asp Ala Lys Phe Tyr Ile Phe Ser Asn Glu
195 200 205

Pro Lys Trp Val Arg Gly Trp Val Val Asp Leu Ile Gln Ser Gln Ile
210 215 220

Gln Glu Gly Met Ser Pro Ser Gln Val Lys Glu Met Glu Lys Arg Phe
225 230 235 240

Val Met Val Glu Ala Asn Thr Glu Tyr Thr Gly Tyr Leu Asp Met Met
245 250 255

Leu Met Ser Lys Cys Lys His Asn Ile Ile Ser Asn Ser Ser Phe Ser
260 265 270

Trp Trp Ser Ala Trp Met Asn Asp His Pro Glu Lys Val Val Val Ala
275 280 285

Pro Asp Arg Trp Ser Ser Asp Lys Glu Gly Asn Glu Ile Tyr Thr Thr
290 295 300

Gly Met Thr Leu Val Asn Glu Lys Gly Arg Val Asn Tyr Thr Ile His
305 310 315 320

Glu Asn Ser Thr Val Lys
325

<210> 244
<211> 311
<212> PRT
<213> Prevotella nigrescens

<400> 244

Met Ile Leu Ser Tyr Ile Thr Gly Arg Leu Gly Asn Gln Leu Phe Glu
1 5 10 15

37847-517 Sequence Listing_ST25 (003)

Tyr Ala Tyr Ala Arg Ser Leu Leu Leu Lys Arg Gly Lys Asn Glu Glu
20 25 30

Leu Ile Leu Asn Phe Ser Leu Val Arg Ala Ala Gly Lys Glu Ile Glu
35 40 45

Gly Phe Asp Asp Asn Leu Arg Tyr Phe Asn Val Tyr Ser Tyr Thr Glu
50 55 60

Leu Asp Lys Asp Ile Val Leu Ser Lys Gly Asp Leu Leu Gln Leu Phe
65 70 75 80

Ile Tyr Ile Leu Phe Lys Leu Asp Gln Lys Leu Phe Arg Ile Ile Lys
85 90 95

Lys Glu Lys Trp Phe Ser Phe Phe Arg Arg Phe Gly Ile Ile Phe Gln
100 105 110

Asp Tyr Leu Asp Asn Ile Ser Asn Leu Ile Ile Pro Arg Thr Lys Asn
115 120 125

Val Phe Cys Tyr Gly Lys Tyr Glu Asn Pro Lys Tyr Phe Asp Asp Ile
130 135 140

Arg Ser Ile Leu Leu Lys Glu Phe Thr Pro Arg Ile Pro Pro Leu Lys
145 150 155 160

Asn Asn Asp Gln Leu Tyr Ser Val Ile Glu Ser Thr Asn Ser Val Cys
165 170 175

Ile Ser Ile Arg Arg Gly Asp Phe Leu Cys Asp Lys Phe Lys Asp Arg
180 185 190

Phe Leu Val Cys Asp Lys Glu Tyr Phe Leu Glu Ala Met Glu Glu Ala
195 200 205

Lys Lys Arg Ile Ser Asn Ser Thr Phe Ile Phe Phe Ser Asp Asp Ile
210 215 220

37847-517 Sequence Listing_ST25 (003)

Glu Trp Val Arg Glu Asn Ile His Ser Asp Val Pro Cys Tyr Tyr Glu
225 230 235 240

Ser Gly Lys Asp Pro Val Trp Glu Lys Leu Arg Leu Met Tyr Ser Cys
245 250 255

Lys His Phe Ile Ile Ser Asn Ser Thr Phe Ser Trp Trp Ala Gln Tyr
260 265 270

Leu Ser Arg Asn Glu Glu Lys Val Val Ile Ala Pro Asp Arg Trp Ser
275 280 285

Asn Val Pro Gly Glu Lys Ser Phe Leu Leu Ser Asn Ser Phe Ile Lys
290 295 300

Ile Pro Ile Gly Ile Leu Pro
305 310

<210> 245

<211> 279

<212> PRT

<213> Bacteroides sp. CAG:875

<400> 245

Met Ile Tyr Val Glu Ile Asn Gly Arg Leu Gly Asn Asn Met Phe Glu
1 5 10 15

Ile Ala Ala Ala Lys Ser Leu Thr Asp Glu Val Thr Leu Trp Cys Lys
20 25 30

Gly Asp Trp Gln Leu Asn Cys Ile Lys Met Tyr Ser Asp Thr Leu Phe
35 40 45

Lys Asn Tyr Pro Ile Val Lys Ser Leu Pro Asn Asn Ile Arg Ile Tyr
50 55 60

Glu Glu Pro Glu Phe Thr Phe His Pro Ile Pro Tyr Lys Glu Asn Gln
65 70 75 80

37847-517 Sequence Listing_ST25 (003)

Asp Leu Leu Ile Lys Gly Tyr Phe Gln Ser Tyr Lys Tyr Leu Asp Arg
 85 90 95

Glu Lys Val Leu Lys Leu Tyr Pro Cys Pro Met Pro Val Lys Leu Asp
 100 105 110

Ile Glu Lys Arg Phe Gly Asp Ile Leu Ser Gln Tyr Thr Val Val Ser
 115 120 125

Ile Asn Val Arg Arg Gly Asp Tyr Leu Asn Leu Pro His Arg His Pro
 130 135 140

Phe Val Gly Lys Lys Phe Leu Glu Arg Ala Met Leu Trp Phe Gly Asp
 145 150 155 160

Lys Val His Tyr Ile Ile Ser Ser Asp Asp Ile Glu Trp Cys Lys Ala
 165 170 175

His Phe Lys Gln Phe Asp Asn Val His Tyr Leu Thr Asn Ser Tyr Pro
 180 185 190

Leu Leu Asp Leu Tyr Ile Gln Thr Ala Cys His His Asn Ile Ile Ser
 195 200 205

Asn Ser Ser Phe Ser Trp Trp Gly Ala Tyr Leu Asn Asn His Pro Gln
 210 215 220

Lys Ile Val Ile Ala Pro His Arg Trp Phe Gly Met Ser Thr Asn Ile
 225 230 235 240

Asn Thr Gln Asp Leu Leu Pro Pro Glu Trp Met Ile Glu Gln Cys Val
 245 250 255

Tyr Glu Pro Lys Val Phe Leu Lys Ala Leu Pro Leu His Ala Lys Tyr
 260 265 270

Leu Leu Lys Arg Val Leu Lys
 275

37847-517 Sequence Listing_ST25 (003)

<210> 246
 <211> 292
 <212> PRT
 <213> Prevotella sp. oral taxon 299

<400> 246

Met Asp Ser Gln Leu Leu Lys His Ile Lys Leu Ser Gly Gly Phe Gly
 1 5 10 15

Asn Gln Leu Phe Gln Tyr Phe Phe Gly Glu Tyr Leu Lys Glu Lys Tyr
 20 25 30

Asn Cys Ser Ile Ser Phe Phe Ser Glu Pro Ala Leu Asp Ile Asn Gln
 35 40 45

Leu Gln Ile His Arg Phe Phe Pro Thr Leu Arg Ile Ser His Asn Thr
 50 55 60

Glu Leu Arg Arg Phe His Tyr Ala Phe Thr Gln Gln Leu Ala Tyr Arg
 65 70 75 80

Cys Met Arg Lys Leu Leu Leu Leu Phe Pro Phe Leu Asn Arg Lys Val
 85 90 95

Lys Ile Glu Asn Gly Ser Asn Tyr Gln Asn Gln Ser Phe Asn Asp Thr
 100 105 110

Tyr Cys Phe Asp Gly Tyr Trp Gln Ser Tyr Arg Tyr Leu Ser Ala Phe
 115 120 125

Thr Pro Ser Leu Gln Phe Glu Asp Gln Leu Ile Asn Asp Ile Ser Ala
 130 135 140

Asp Tyr Ile Asn Ala Ile Glu Gln Ser Glu Ala Val Phe Leu His Ile
 145 150 155 160

Arg Arg Gly Asp Tyr Leu Asn Lys Glu Asn Gln Lys Val Phe Ala Glu
 165 170 175

37847-517 Sequence Listing_ST25 (003)

Cys Pro Leu Asn Tyr Phe Glu Asn Ala Val Asn Lys Ile Lys Glu Gly
180 185 190

Asn Lys Thr Tyr His Phe Phe Val Phe Ser Asn Asp Ile Glu Trp Val
195 200 205

Lys Cys His Leu Lys Leu Asn Asn Asn Glu Val Thr Phe Ile Gln Asn
210 215 220

Glu Gly Ser Ser Cys Asp Leu Lys Asp Phe Tyr Leu Met Thr Arg Cys
225 230 235 240

Lys His Ala Ile Ile Ser Asn Ser Thr Phe Ser Trp Trp Ala Ala Tyr
245 250 255

Leu Ile Asn Asn Asn Asp Lys Lys Val Ile Ala Pro Lys Arg Trp Tyr
260 265 270

Asn Asp Leu Ser Met Asn Asn Ala Thr Lys Asp Leu Ile Pro Pro Thr
275 280 285

Trp Ile Arg Leu
290

<210> 247
<211> 265
<212> PRT
<213> Paraprevotella xylaniphila

<400> 247

Met Lys Ile Val Cys Leu Lys Gly Gly Leu Gly Asn Gln Met Phe Glu
1 5 10 15

Tyr Cys Arg Phe Arg Asp Leu Met Asp Ser Gly Asn Gly Lys Val Tyr
20 25 30

Leu Phe Tyr Asp Arg Arg Arg Leu Lys Gln His Asp Gly Leu Arg Leu
35 40 45

Ser Asp Cys Phe Glu Leu Glu Leu Pro Ser Cys Pro Trp Gly Ile Arg

28 May 2018
2015259088

37847-517 Sequence Listing_ST25 (003)

50

55

60

Leu Val Val Trp Gly Leu Lys Ile Cys Arg Ala Ile Gly Val Leu Lys
65 70 75 80

Arg Leu Tyr Asp Asp Glu Lys Pro Asp Ala Val Leu Ile Asp Asp Tyr
85 90 95

Ser Gln His Arg Arg Phe Ile Pro Asn Ala Arg Arg Tyr Phe Ser Phe
100 105 110

Arg Gln Phe Leu Ala Glu Leu Gln Ser Gly Phe Val Gln Met Ile Arg
115 120 125

Ala Val Asp Tyr Pro Val Ser Val His Val Arg Arg Gly Asp Tyr Leu
130 135 140

His Pro Ser Asn Ser Ser Phe Val Leu Cys Gly Val Asp Tyr Phe Arg
145 150 155 160

Gln Ala Ile Ala Tyr Val Arg Lys Lys Arg Pro Asp Ala Arg Phe Phe
165 170 175

Phe Phe Ser Asp Asp Met Glu Trp Val Arg Glu Asn Leu Trp Met Glu
180 185 190

Asp Ala Val Tyr Val Glu His Thr Glu Leu Met Pro Asp Tyr Met Asp
195 200 205

Leu Tyr Leu Met Thr Leu Cys Arg Gly His Ile Ile Ser Asn Ser Thr
210 215 220

Phe Ser Phe Trp Gly Ala Tyr Leu Ala Val Asp Gly Asn Gly Met Lys
225 230 235 240

Ile Tyr Pro Arg Arg Trp Phe Arg Asp Pro Thr Trp Ile Thr Pro Pro
245 250 255

Ile Phe Ser Glu Glu Trp Val Gly Leu

37847-517 Sequence Listing_ST25 (003)

260

265

<210> 248
<211> 281
<212> PRT
<213> Dethiosulfovibrio peptidovorans

<400> 248

Met Phe Gln Tyr Ala Phe Gly Arg Ala Leu Ala Leu Asp Leu Gly Leu
1 5 10 15

Asp Leu Lys Leu Asp Ile Ser Asn Phe Gly Ser Asp Ser Arg Pro Phe
20 25 30

Ser Leu Gly Ile Tyr Ser Leu Thr Lys Asn Ile Pro Phe Gly Cys Tyr
35 40 45

Leu Ser Thr Ser Thr Arg Leu Lys Val Lys Met Thr Lys Lys Leu Arg
50 55 60

Arg Trp Gly Val Trp Gly Met Asp Lys Asn Met Pro Gly Val Leu Val
65 70 75 80

Glu Pro Phe Pro Pro Val Leu Val Ser Leu Asp Glu Val Leu Ser Glu
85 90 95

Lys Leu Ser His Leu Phe Val Asp Gly Tyr Trp Gln Ser Glu Lys Tyr
100 105 110

Phe Ser Arg Tyr Ser Asp Val Ile Arg Ser Asp Phe Arg Val Ile Glu
115 120 125

Glu Ser Ser Ala Phe Leu Ala Trp Lys Lys Arg Met Leu Ser Glu Pro
130 135 140

Gly Gly Ser Ile Ser Val His Val Arg Arg Gly Asp Tyr Val Thr Asp
145 150 155 160

Ser Ser Ala Asn Arg Val His Gly Val Leu Pro Ile Glu Tyr Tyr Leu
165 170 175

37847-517 Sequence Listing_ST25 (003)

Arg Ala Lys Glu Ile Leu Asn Thr Ile Ser Asp Gly Leu Val Phe Tyr
180 185 190

Val Phe Thr Asp Asp Pro Val Trp Ala Arg Asn Asn Leu Cys Leu Gly
195 200 205

Asp Lys Thr Ile Tyr Val Ser Gly Glu Asp Leu Lys Asp Tyr Glu Glu
210 215 220

Leu Ala Leu Met Ser Cys Cys Asp His His Val Val Ala Asn Ser Ser
225 230 235 240

Phe Ser Trp Trp Gly Ala Trp Leu Gly Gln Asp Thr Ser Thr Val Thr
245 250 255

Ile Ala Pro Gly Arg Trp Phe Arg Lys Met Asp Ser Ser Phe Val Ile
260 265 270

Pro Asp Asn Trp Ile Lys Ile Trp Thr
275 280

<210> 249

<211> 299

<212> PRT

<213> Lachnospiraceae bacterium 10-1

<400> 249

Met Ile Ile Ile Gln Val Met Gly Gly Leu Gly Asn Gln Leu Gln Gln
1 5 10 15

Tyr Ala Leu Tyr Arg Lys Phe Val Arg Met Gly Lys Glu Ala Arg Leu
20 25 30

Asp Ile Ser Trp Phe Leu Asp Lys Glu Lys Arg Gly Glu Val Leu Ala
35 40 45

Glu Arg Glu Leu Glu Leu Asp Tyr Phe Asp Arg Leu Ile Tyr Glu Thr
50 55 60

37847-517 Sequence Listing_ST25 (003)

Cys Thr Pro Glu Glu Lys Glu Gln Leu Ile Gly Ser Glu Gly Val Ala
65 70 75 80

Gly Lys Leu Lys Arg Lys Phe Leu Pro Gly Arg Ile Arg Trp Phe His
85 90 95

Glu Ser Lys Ile Tyr His Pro Glu Leu Leu Gln Met Glu Asn Met Tyr
100 105 110

Leu Ser Gly Tyr Phe Ala Cys Glu Lys Tyr Tyr Ala Asp Ile Leu Tyr
115 120 125

Asp Leu Arg Glu Lys Ile Gln Phe Pro Val Asn Asp His Pro Lys Asn
130 135 140

Ile Lys Met Ala Gln Glu Met Gln Glu Arg Glu Ser Val Ser Val His
145 150 155 160

Leu Arg Arg Gly Asp Tyr Leu Asp Glu Lys Asn Thr Ala Met Phe Gly
165 170 175

Asn Ile Cys Thr Asp Ala Tyr Tyr Cys Lys Ala Ile Glu Tyr Met Lys
180 185 190

Thr Leu Cys Ser Lys Pro His Phe Tyr Ile Phe Ser Asp Asp Ile Pro
195 200 205

Tyr Val Arg Gln Arg Phe Thr Gly Glu Glu Tyr Thr Val Val Asp Ile
210 215 220

Asn His Gly Arg Asp Ser Phe Phe Asp Met Trp Leu Met Ser Arg Cys
225 230 235 240

Arg His Asn Ile Cys Ala Asn Ser Thr Phe Ser Phe Trp Gly Ala Arg
245 250 255

Leu Asn Ser Asn Asp Asn Lys Ile Met Ile Arg Pro Thr Ile His Lys
260 265 270

37847-517 Sequence Listing_ST25 (003)

Asn Ser Gln Val Phe Val Lys Glu Glu Met Glu Gln Leu Trp Pro Gly
 275 280 285

Trp Lys Phe Ile Ser Pro Asp Gly Gly Ile Lys
 290 295

<210> 250
 <211> 292
 <212> PRT
 <213> Treponema maltophilum
 <400> 250

Met Phe Cys Ala Ala Phe Val Glu Ala Leu Lys His Ala Gly Gln Lys
 1 5 10 15

Val Phe Val Asp Thr Ser Leu Tyr Asn Lys Gly Thr Val Arg Ser Gly
 20 25 30

Ile Asp Phe Cys His Asn Gly Leu Glu Thr Glu His Leu Phe Gly Ile
 35 40 45

Lys Phe Asp Glu Ala Asp Lys Ala Asp Val His Arg Leu Ser Thr Ser
 50 55 60

Ala Glu Gly Leu Leu Asn Arg Ile Arg Arg Lys Tyr Phe Thr Lys Lys
 65 70 75 80

Thr His Tyr Ile Asp Thr Val Phe Arg Tyr Thr Pro Glu Val Leu Ser
 85 90 95

Asp Lys Ser Asp Arg Tyr Leu Glu Gly Phe Trp Gln Thr Glu Lys Tyr
 100 105 110

Phe Leu Pro Ile Glu Ser Asp Ile Arg Thr Leu Phe Arg Phe Arg Gln
 115 120 125

Pro Leu Ser Glu Lys Ser Ala Ala Val Gln Ser Ala Leu Gln Ala Gln
 130 135 140

37847-517 Sequence Listing_ST25 (003)

Glu Pro Ala Ser Leu Ser Ala Ser Ile His Val Arg Arg Gly Asp Phe
145 150 155 160

Leu His Thr Lys Thr Leu Asn Val Cys Thr Glu Thr Tyr Tyr Asn Asn
165 170 175

Ala Ile Glu Tyr Ala Ala Lys Lys Tyr Ala Val Ser Ala Phe Tyr Val
180 185 190

Phe Ser Asp Asp Ile Gln Trp Cys Arg Glu His Leu Asn Phe Phe Gly
195 200 205

Ala Arg Ser Val Phe Ile Asp Trp Asn Ile Gly Ala Asp Ser Trp Gln
210 215 220

Asp Met Val Leu Met Ser Met Cys Arg Cys Asn Ile Ile Ala Asn Ser
225 230 235 240

Ser Phe Ser Trp Trp Ala Ala Trp Leu Asn Ala Ala Ser Asp Lys Ile
245 250 255

Val Leu Ala Pro Ala Ile Trp Asn Arg Arg Gln Leu Glu Tyr Ala Asp
260 265 270

Arg Tyr Tyr Gly Tyr Asp Tyr Ser Asp Val Ile Pro Glu Thr Trp Ile
275 280 285

Arg Ile Pro Ile
290

<210> 251
<211> 266
<212> PRT
<213> Bacteroides massiliensis

<400> 251

Met Lys Leu Val Ser Phe Thr Ala Gly Leu Gly Asn Gln Leu Phe Gln
1 5 10 15

Tyr Cys Phe Tyr Arg Tyr Leu Leu Asn Lys Phe Pro Asn Glu Lys Ile

37847-517 Sequence Listing_ST25 (003)

20

25

30

Tyr Gly Tyr Tyr Asn Lys Lys Trp Leu Lys Lys His Gly Gly Ile Ile
35 40 45

Ile Glu His Phe Phe Asp Val Lys Leu Pro Arg Ser Thr Arg Trp Ile
50 55 60

Asn Leu Tyr Gly Gln Tyr Leu Arg Ile Ile Tyr Lys Cys Phe Ser Cys
65 70 75 80

Gly Val Ser Lys Asp Asp Asp Phe Glu Met Asn Arg Thr Met Phe Val
85 90 95

Gly Tyr Trp Gln Asp Gln Cys Phe Phe Ser Gly Ile Asn Ile Ser Tyr
100 105 110

Lys Lys Asn Leu Val Ile Ser Glu Lys Asn Thr Trp Leu Leu Gly Glu
115 120 125

Ile Leu Lys Cys Asn Ser Val Ala Ile His Phe Arg Arg Gly Asp Tyr
130 135 140

Met Leu Pro Gln Phe Lys Lys Ile Phe Gly Glu Val Cys Thr Val Lys
145 150 155 160

Tyr Tyr Leu Lys Ser Ile Arg Lys Val Glu Glu Lys Ile Ser Glu Pro
165 170 175

Val Phe Phe Val Phe Ser Asp Asp Ile Asp Trp Val Lys Gln Asn Phe
180 185 190

Thr Phe Asn Lys Val Tyr Phe Val Asp Trp Asn Lys Gly Gln Asn Ser
195 200 205

Phe Trp Asp Met Tyr Leu Met Ser Gln Cys Ser Ala Asn Ile Ile Ala
210 215 220

Asn Ser Thr Phe Ser Phe Trp Gly Ala Tyr Leu Asn Lys Asn Asn Pro

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

225 230 235 240

Phe Val Ile Tyr Pro Gln Lys Trp Val Arg Thr Asn Leu Lys Gln Pro
245 250 255

Asn Ile Phe Pro Lys Thr Trp Met Ala Leu
260 265

<210> 252
<211> 303
<212> PRT
<213> Enterococcus faecium

<400> 252

Met Ile Val Leu Thr Leu Gly Gly Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Gly Tyr Ala Arg Tyr Ile Gln Lys Ile His Arg Glu Lys Phe Ile
20 25 30

Tyr Ile Asn Asp Ser Glu Val Ile Lys Glu Ala Asp Arg Phe Asn Ser
35 40 45

Leu Gly Asn Leu Asn Thr Val Asn Ile Lys Val Leu Pro Arg Ile Ile
50 55 60

Ser Lys Pro Leu Asn Glu Thr Glu Arg Leu Val Arg Lys Ile Met Val
65 70 75 80

Arg Leu Phe Gly Val Ala Gly Phe Asn Glu Ser Ala Ile Phe Gln Ser
85 90 95

Leu Asn Lys Phe Gly Ile Tyr Tyr His Pro Ser Val Tyr Lys Phe Tyr
100 105 110

Glu Ser Leu Lys Thr Gly Phe Pro Ile Lys Ile Ile Glu Gly Gly Phe
115 120 125

Gln Ser Trp Lys Tyr Leu Glu Thr Cys Pro Glu Ile Lys Gln Glu Leu
130 135 140

37847-517 Sequence Listing_ST25 (003)

Arg Val Lys Tyr Glu Pro Met Gly Glu Asn Leu Arg Leu Leu Asn Leu
145 150 155 160

Ile Ser Gln Ser Glu Ser Val Cys Val His Ile Arg Arg Gly Asp Tyr
165 170 175

Leu Ser Pro Lys Tyr Lys His Leu Asn Val Cys Asp Tyr Gln Tyr Tyr
180 185 190

Phe Glu Ser Met Asn Tyr Ile Ile Ser Lys Leu Asn Asn Pro Thr Phe
195 200 205

Phe Ile Phe Ser Asn Thr Ser Asp Asp Leu Asp Trp Ile Lys Glu Asn
210 215 220

Tyr Ser Leu Pro Gly Lys Ile Val Tyr Val Lys Asn Asp Asn Pro Asp
225 230 235 240

Tyr Glu Glu Leu Arg Leu Met Tyr Ser Cys Lys His Phe Ile Ile Ser
245 250 255

Asn Ser Thr Phe Ser Trp Trp Ala Gln Tyr Leu Ser Asn Asn Ser Gly
260 265 270

Ile Val Ile Ala Pro Glu Ile Trp Asn Arg Leu Asn His Asp Gly Ile
275 280 285

Ala Asp Leu Tyr Met Pro Asn Trp Ile Thr Met Lys Val Asn Arg
290 295 300

<210> 253

<211> 290

<212> PRT

<213> Bacteroides ovatus

<400> 253

Met Asp Val Val Val Ile Phe Asn Gly Leu Gly Asn Gln Met Ser Gln
1 5 10 15

37847-517 Sequence Listing_ST25 (003)

Tyr Ala Tyr Tyr Leu Ala Lys Lys Lys Val Asn Pro Asn Thr Lys Val
 20 25 30
 Ile Phe Asp Ile Met Ser Lys His Asn His Tyr Gly Tyr Asp Leu Glu
 35 40 45
 Arg Ala Phe Gly Ile Glu Val Asn Lys Thr Leu Leu Ile Lys Val Leu
 50 55 60
 Gln Ile Ile Tyr Val Leu Ser Arg Lys Phe Arg Leu Phe Lys Ser Val
 65 70 75 80
 Gly Val Arg Thr Ile Tyr Glu Pro Leu Asn Tyr Asp Tyr Thr Pro Leu
 85 90 95
 Leu Met Gln Lys Gly Pro Trp Gly Ile Asn Tyr Tyr Val Gly Gly Trp
 100 105 110
 His Ser Glu Lys Asn Phe Met Asn Val Pro Asp Glu Val Lys Lys Ala
 115 120 125
 Phe Met Phe Arg Glu Gln Pro Asn Glu Asp Arg Phe Asn Glu Trp Leu
 130 135 140
 Gln Val Ile Arg Gly Asp Asn Ser Ser Val Ser Val His Ile Arg Arg
 145 150 155 160
 Gly Asp Tyr Met Asn Ile Glu Pro Thr Gly Tyr Tyr Gln Leu Asn Gly
 165 170 175
 Val Ala Thr Leu Asp Tyr Tyr His Glu Ala Ile Asp Tyr Ile Arg Gln
 180 185 190
 Tyr Val Asp Thr Pro His Phe Tyr Val Phe Ser Asn Asp Leu Asp Trp
 195 200 205
 Cys Lys Glu Gln Phe Gly Val Glu Asn Phe Phe Tyr Ile Glu Cys Asn
 210 215 220

37847-517 Sequence Listing_ST25 (003)

Gln Gly Val Asn Ser Trp Arg Asp Met Tyr Leu Met Ser Glu Cys His
225 230 235 240

Tyr His Ile Asn Ala Asn Ser Thr Phe Ser Trp Trp Gly Ala Trp Leu
245 250 255

Cys Lys Phe Glu Asp Ser Ile Thr Val Cys Pro Glu Arg Phe Ile Arg
260 265 270

Asn Val Val Thr Lys Asp Phe Tyr Pro Glu Arg Trp His Lys Ile Lys
275 280 285

Ser Cys
290

<210> 254

<211> 271

<212> PRT

<213> Synechococcus phage S-SM2

<400> 254

Met Ile Gly Phe Asn Ala Leu Gly Arg Met Gly Arg Leu Ala Asn Gln
1 5 10 15

Met Phe Gln Tyr Ala Ser Leu Lys Gly Ile Ala Arg Asn Thr Gly Val
20 25 30

Asp Phe Cys Val Pro Tyr His Glu Glu Ala Val Asn Asp Gly Ile Gly
35 40 45

Asn Met Leu Arg Thr Glu Ile Phe Asp Ser Phe Asp Leu Gln Val Asn
50 55 60

Val Gly Leu Leu Asn Lys Gly His Ala Pro Val Val Gln Glu Arg Phe
65 70 75 80

Phe His Phe Asp Glu Glu Leu Phe Arg Met Cys Pro Asp His Val Asp
85 90 95

37847-517 Sequence Listing_ST25 (003)

Ile Arg Gly Tyr Phe Gln Thr Glu Lys Tyr Phe Lys His Ile Glu Asp
 100 105 110

Glu Ile Arg Glu Asp Phe Thr Phe Lys Asp Glu Ile Leu Asn Pro Cys
 115 120 125

Lys Glu Met Ile Ala Gly Val Asp Asn Pro Leu Ala Leu His Val Arg
 130 135 140

Arg Thr Asp Tyr Val Thr Asn Ser Ala Asn His Pro Pro Cys Thr Leu
 145 150 155 160

Glu Tyr Tyr Glu Ala Ala Leu Lys His Phe Asp Asp Asp Arg Asn Val
 165 170 175

Ile Val Phe Ser Asp Asp Pro Ala Trp Cys Lys Glu Gln Glu Leu Phe
 180 185 190

Ser Asp Asp Arg Phe Met Ile Ser Glu Asn Glu Asp Asn Arg Ile Asp
 195 200 205

Leu Cys Leu Met Ser Leu Cys Asp Asp Phe Ile Ile Ala Asn Ser Thr
 210 215 220

Tyr Ser Trp Trp Gly Ala Trp Leu Ser Ala Asn Lys Asp Lys Lys Val
 225 230 235 240

Ile Ala Pro Val Gln Trp Phe Gly Thr Gly Tyr Thr Lys Asp His Asp
 245 250 255

Thr Ser Asp Leu Ile Pro Asp Gly Trp Thr Arg Ile Ala Thr Ala
 260 265 270

<210> 255

<211> 294

<212> PRT

<213> Geobacter metallireducens

<400> 255

Met Asp Ile His Val Leu Ser Tyr Gly Leu Gly Asn Gln Leu Ser Gln

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

1	5	10	15												
Tyr	Ala	Phe	Phe	Ile	Asn	Arg	Arg	Gln	Leu	Met	Gln	Arg	Ala	Tyr	Ala
		20						25					30		
Phe	Tyr	Ala	Phe	Lys	Gln	His	Asn	Gly	Tyr	Glu	Leu	Asp	Arg	Ile	Phe
		35					40					45			
Gly	Leu	Lys	Glu	Gly	Leu	Pro	Trp	Tyr	Leu	Gln	Phe	Val	Arg	Val	Val
	50					55					60				
Phe	Arg	Leu	Gly	Ile	Ser	Arg	Arg	Phe	Tyr	Ser	Lys	Arg	Thr	Ala	Asp
65					70					75					80
Phe	Val	Leu	Ser	Leu	Phe	Arg	Ile	Lys	Val	Ile	Asp	Glu	Ala	Tyr	Asn
			85						90					95	
Tyr	Glu	Phe	Asp	Pro	Ser	Leu	Leu	Lys	Pro	Trp	Phe	Gly	Ile	Arg	Ile
			100					105					110		
Leu	Tyr	Gly	Gly	Trp	His	Asp	Ser	Arg	Tyr	Phe	His	Pro	Ser	Glu	Ala
		115					120					125			
Ala	Val	Arg	Thr	Ala	Phe	Ser	Phe	Pro	Pro	Leu	Asp	Asp	Val	Asn	Asp
		130				135					140				
Ala	Ile	Leu	Gln	Gln	Ile	Asp	Ala	Val	Tyr	Gly	Val	Ser	Ile	His	Val
145					150					155					160
Arg	Arg	Gly	Asp	Tyr	Leu	Lys	Gly	Ile	Asn	Ser	Asn	Leu	Phe	Gly	Gly
			165					170						175	
Ile	Ala	Thr	Leu	Glu	Tyr	Tyr	Arg	Asn	Ala	Ile	Gly	Trp	Ala	Ile	Thr
			180					185					190		
Tyr	Cys	Lys	His	Arg	Ser	Leu	Glu	Ile	Lys	Phe	Tyr	Val	Phe	Ser	Asp
		195					200					205			
Asp	Ile	Asp	Trp	Cys	Lys	Gln	Asn	Leu	Gly	Leu	Arg	Asp	Ala	Val	Tyr

37847-517 Sequence Listing_ST25 (003)

210

215

220

Val Ser Gly Asn Ser Lys Thr Asp Ser Trp Lys Asp Ile Leu Leu Met
 225 230 235 240

Ser His Cys Arg Ala Asn Ile Ile Ala Asn Ser Thr Phe Ser Trp Trp
 245 250 255

Ala Ala Trp Leu Asn Gln Gln Pro Asn Lys Val Val Ile Cys Pro Thr
 260 265 270

Lys Phe Ile Asn Thr Asp Ser Pro Asn Gln Thr Ile Tyr Pro Ala Ala
 275 280 285

Trp His Gln Ile Glu Gly
 290

<210> 256

<211> 299

<212> PRT

<213> Lachnospiraceae bacterium NK4A136

<400> 256

Met Ile Ile Val Arg Phe His Gly Gly Leu Gly Asn Gln Met Phe Glu
 1 5 10 15

Tyr Ala Phe Tyr Arg Tyr Met Thr Asn Lys Tyr Gly Ala Asp Asn Val
 20 25 30

Ile Gly Asp Met Thr Trp Phe Asp Arg Asn Tyr Ser Glu His Gln Gly
 35 40 45

Tyr Glu Leu Lys Lys Val Phe Asp Ile Asp Ile Pro Ala Ile Asp Tyr
 50 55 60

Lys Thr Leu Ala Lys Ile His Glu Tyr Tyr Pro Arg Tyr His Arg Phe
 65 70 75 80

Ala Gly Leu Arg Tyr Leu Ser Arg Met Tyr Ala Lys Tyr Lys Asn Lys
 85 90 95

37847-517 Sequence Listing_ST25 (003)

His Leu Lys Pro Thr Gly Glu Tyr Ile Met Asp Phe Gly Pro Ser Gln
100 105 110

Tyr Ile His Asn Asp Ala Phe Asp Lys Leu Asp Thr Asn Lys Asp Tyr
115 120 125

Tyr Ile Glu Gly Val Phe Cys Ser Asp Ala Tyr Ile Lys Tyr Tyr Glu
130 135 140

Asn Gln Ile Lys Lys Asp Leu Thr Phe Lys Pro Asn Tyr Ser Gln His
145 150 155 160

Thr Lys Asp Met Leu Pro Lys Ile Glu Glu Thr Asn Ser Val Ala Ile
165 170 175

His Val Arg Arg Gly Asp Tyr Val Gly Asn Val Phe Asp Ile Val Thr
180 185 190

Pro Asp Tyr Tyr Arg Gln Ala Val Asn Tyr Ile Arg Glu Arg Val Glu
195 200 205

Asn Pro Val Phe Phe Val Phe Ser Asp Asp Met Asp Tyr Ile Lys Ala
210 215 220

Asn Phe Asp Phe Leu Gly Asp Phe Val Pro Val His Asn Cys Gly Lys
225 230 235 240

Asp Ser Phe Gln Asp Met Tyr Leu Ile Ser Arg Cys Arg His Met Ile
245 250 255

Ile Ala Asn Ser Ser Phe Ser Tyr Phe Gly Ala Leu Leu Gly Glu Lys
260 265 270

Asp Ser Thr Ile Val Ile Ala Pro Lys Lys Tyr Lys Ala Asp Glu Asp
275 280 285

Leu Ala Leu Ala Arg Glu Asn Trp Val Leu Leu
290 295

37847-517 Sequence Listing_ST25 (003)

28 May 2018
2015259088

<210> 257
 <211> 290
 <212> PRT
 <213> Bacteroides coprophilus

<400> 257

Met Gly Phe Ile Val Asn Met Ala Cys Gly Leu Ala Asn Arg Met Phe
 1 5 10 15

Gln Tyr Ser Tyr Tyr Leu Phe Leu Lys Lys Gln Gly Tyr Lys Val Thr
 20 25 30

Val Asp Phe Tyr Arg Ser Ala Lys Leu Ala His Glu Lys Val Ala Trp
 35 40 45

Asn Ser Ile Phe Pro Tyr Ala Glu Ile Lys Gln Ala Ser Arg Leu Lys
 50 55 60

Val Phe Leu Trp Gly Gly Gly Ser Asp Leu Cys Ser Lys Val Arg Arg
 65 70 75 80

Arg Tyr Phe Pro Ser Ser Thr Asn Val Arg Thr Thr Thr Gly Ala Phe
 85 90 95

Asp Ala Ser Leu Pro Ala Asn Thr Ala Arg Asn Glu Tyr Ile Ile Gly
 100 105 110

Val Phe Leu Asn Ala Ser Ile Val Glu Ala Val Asp Asp Glu Ile Lys
 115 120 125

Lys Cys Phe Thr Phe Leu Pro Phe Thr Asp Glu Met Asn Leu Arg Leu
 130 135 140

Lys Lys Glu Ile Glu Glu Cys Glu Ser Val Ala Ile His Val Arg Lys
 145 150 155 160

Gly Lys Asp Tyr Gln Ser Arg Ile Trp Tyr Gln Asn Thr Cys Ser Met
 165 170 175

37847-517 Sequence Listing_ST25 (003)

Glu Tyr Tyr Arg Lys Ala Ile Leu Gln Met Lys Glu Lys Leu Gln His
180 185 190

Ser Lys Phe Tyr Val Phe Thr Asp Asn Val Asp Trp Val Lys Glu Asn
195 200 205

Phe Gln Glu Ile Asp Tyr Thr Leu Val Glu Gly Asn Pro Ala Asp Gly
210 215 220

Tyr Gly Ser His Phe Asp Met Gln Leu Met Ser Leu Cys Lys His Asn
225 230 235 240

Ile Ile Ser Asn Ser Thr Tyr Ser Trp Trp Ser Ala Phe Leu Asn Arg
245 250 255

Asn Pro Glu Lys Val Val Ile Ala Pro Glu Ile Trp Phe Asn Pro Asp
260 265 270

Ser Cys Asp Glu Phe Arg Ser Asp Arg Ala Leu Cys Lys Gly Trp Ile
275 280 285

Val Leu
290

<210> 258

<211> 265

<212> PRT

<213> Paraprevotella clara YIT 11840

<400> 258

Met Lys Ile Val Cys Leu Lys Gly Gly Leu Gly Asn Gln Met Phe Glu
1 5 10 15

Tyr Cys Arg Phe Arg Asp Leu Met Glu Ser Gly His Asp Glu Val Tyr
20 25 30

Leu Phe Tyr Asp His Arg Arg Leu Lys Gln His Asn Gly Leu Arg Leu
35 40 45

37847-517 Sequence Listing_ST25 (003)

Ser	Asp	Cys	Phe	Glu	Leu	Glu	Leu	Pro	Ser	Cys	Pro	Trp	Gly	Ile	Lys	50	55	60
Leu	Val	Val	Trp	Gly	Leu	Lys	Ile	Cys	Arg	Ala	Val	Gly	Val	Leu	Lys	65	70	75
Arg	Leu	Tyr	Asp	Asp	Glu	Lys	Pro	Glu	Ala	Val	Leu	Ile	Asp	Asp	Tyr	85	90	95
Ser	Gln	His	Arg	Arg	Phe	Ile	Pro	Asn	Ala	Arg	Arg	Tyr	Phe	Phe	Phe	100	105	110
Arg	Gln	Phe	Leu	Ala	Glu	Leu	Gln	Ser	Gly	Phe	Val	Gln	Met	Ile	Arg	115	120	125
Ala	Val	Asp	Tyr	Pro	Val	Ser	Val	His	Val	Arg	Arg	Gly	Asp	Tyr	Leu	130	135	140
His	Pro	Ser	Asn	Ser	Ser	Phe	Gly	Leu	Cys	Gly	Val	Asp	Tyr	Phe	Gln	145	150	155
Gln	Ala	Ile	Ala	Tyr	Val	Arg	Lys	Lys	Arg	Pro	Asp	Ala	Arg	Phe	Phe	165	170	175
Phe	Phe	Ser	Asp	Asp	Met	Glu	Trp	Val	Arg	Glu	Asn	Leu	Trp	Met	Glu	180	185	190
Asp	Ala	Val	Tyr	Val	Glu	His	Thr	Glu	Leu	Leu	Pro	Asp	Tyr	Val	Asp	195	200	205
Leu	Tyr	Leu	Met	Thr	Leu	Cys	Arg	Gly	His	Ile	Ile	Ser	Asn	Ser	Thr	210	215	220
Phe	Ser	Phe	Trp	Gly	Ala	Tyr	Leu	Ala	Val	Asp	Gly	Asn	Gly	Met	Lys	225	230	235
Ile	Tyr	Pro	Arg	Arg	Trp	Phe	Arg	Asp	Pro	Thr	Trp	Thr	Ser	Pro	Pro	245	250	255

28 May 2018
2015259088

37847-517 Sequence Listing_ST25 (003)

Ile Phe Ser Glu Glu Trp Val Gly Leu
260 265

<210> 259

<211> 309

<212> PRT

<213> Butyrivibrio sp. NC2007

<400> 259

Met Leu Ile Ile Gln Ile Ala Gly Gly Leu Gly Asn Gln Met Gln Gln
1 5 10 15

Tyr Ala Val Tyr Thr Lys Leu Arg Glu Met Gly Lys Asp Val Lys Leu
20 25 30

Asp Leu Ser Trp Phe Asp Pro Gln Val Gln Lys Asn Met Leu Ala Pro
35 40 45

Arg Glu Phe Glu Leu Pro Ile Phe Gly Gly Thr Asp Tyr Glu Glu Cys
50 55 60

Ser Ala Tyr Glu Arg Asp Ala Leu Leu Lys Gln Gly Ala Phe Ala Ala
65 70 75 80

Ile Ala Gly Lys Val Leu Lys Lys Leu Gly Leu Arg Asp Glu Ala Asn
85 90 95

Pro Lys Val Phe Ser Glu Lys Glu Met Tyr His Pro Glu Val Phe Glu
100 105 110

Leu Glu Asp Lys Tyr Ile Lys Gly Tyr Phe Ala Cys Gln Lys Tyr Tyr
115 120 125

Gly Asp Ile Met Asp Lys Leu Gln Glu Lys Phe Ile Phe Pro Glu His
130 135 140

Ser Asp Pro Asp Leu His Ala Arg Asn Met Ala Leu Val Glu Arg Met
145 150 155 160

Glu Arg Glu Pro Ser Val Ser Val His Ile Arg Arg Gly Asp Tyr Leu

37847-517 Sequence Listing_ST25 (003)

165 170 175

Asp Pro Ser Asn Val Glu Ile Leu Gly Asn Ile Ala Thr Glu Gln Tyr
180 185 190

Tyr Gln Gly Ala Met Asp Tyr Phe Thr Val Lys Glu Pro Asp Thr His
195 200 205

Phe Tyr Ile Phe Thr Ser Asp His Glu Tyr Ala Arg Glu Lys Phe Ser
210 215 220

Asp Glu Ser Lys Tyr Thr Ile Val Asp Trp Asn Asn Gly Lys Asn Ser
225 230 235 240

Val Gln Asp Leu Met Leu Met Ser His Cys Lys Gly Asn Ile Cys Ala
245 250 255

Asn Ser Thr Phe Ser Phe Trp Gly Ala Arg Leu Asn Lys Arg Pro Asp
260 265 270

Lys Thr Val Ile Arg Thr Tyr Lys Met Arg Asn Asn Gln Pro Val Asn
275 280 285

Pro Gln Ile Met His Asp Tyr Trp Lys Gly Trp Ile Leu Met Asp Glu
290 295 300

Lys Gly Ser Ile Ile
305

<210> 260
<211> 273
<212> PRT
<213> Paraprevotella xylaniphila

<400> 260

Met Lys Ile Leu Val Phe Thr Gly Gly Leu Gly Asn Gln Met Phe Ala
1 5 10 15

Tyr Ala Phe Tyr Leu Tyr Leu Lys Arg Leu Phe Pro Gln Glu Arg Phe
20 25 30

37847-517 Sequence Listing_ST25 (003)

Tyr Gly Leu Tyr Gly Lys Lys Leu Ser Glu His Tyr Gly Leu Glu Ile
 35 40 45
 Asp Lys Trp Phe Lys Val Ser Leu Pro Arg Gln Pro Trp Trp Val Leu
 50 55 60
 Pro Val Thr Gly Leu Phe Tyr Leu Tyr Lys Gln Cys Val Pro Asn Ser
 65 70 75 80
 Lys Trp Leu Asp Leu Asn Gln Glu Ile Cys Lys Asn Pro Arg Ala Ile
 85 90 95
 Val Phe Phe Pro Phe Lys Phe Thr Lys Lys Tyr Ile Pro Asp Asp Asn
 100 105 110
 Ile Trp Leu Glu Trp Lys Val Asp Glu Ser Gly Leu Ser Glu Lys Asn
 115 120 125
 Arg Leu Leu Leu Ser Glu Ile Arg Ser Ser Asp Cys Cys Phe Val His
 130 135 140
 Val Arg Arg Gly Asp Tyr Leu Ser Pro Thr Phe Lys Ser Leu Phe Glu
 145 150 155 160
 Gly Cys Cys Thr Leu Ser Tyr Tyr Gln Arg Ala Leu Lys Ser Met Lys
 165 170 175
 Glu Ile Ser Pro Phe Val Lys Phe Val Cys Phe Ser Asp Asp Ile Gln
 180 185 190
 Trp Val Lys Gln Asn Leu Glu Leu Gly Asn Arg Ala Val Phe Val Asp
 195 200 205
 Trp Asn Ser Gly Thr Asp Ser Pro Leu Asp Met Tyr Leu Met Ser Gln
 210 215 220
 Cys Arg Tyr Gly Ile Met Ala Asn Ser Thr Phe Ser Tyr Trp Gly Ala
 225 230 235 240

37847-517 Sequence Listing_ST25 (003)

Arg Leu Gly Arg Lys Lys Lys Arg Ile Tyr Tyr Pro Gln Lys Trp Trp
245 250 255

Asn His Gly Thr Gly Leu Pro Asp Ile Phe Pro Asn Thr Trp Val Lys
260 265 270

Ile

<210> 261

<211> 310

<212> PRT

<213> *Blautia hydrogenotrophica* DSM 10507

<400> 261

Met Glu Ile His Val Tyr Leu Thr Gly Arg Leu Gly Asn Gln Leu Phe
1 5 10 15

Gln Tyr Ala Phe Ala Arg His Leu Gln Lys Glu Tyr Gly Gly Lys Ile
20 25 30

Ile Cys Asn Ile Tyr Glu Leu Glu His Arg Ser Glu Lys Ala Ala Trp
35 40 45

Val Pro Gly Lys Phe Asn Tyr Glu Met Ser Asn Tyr Lys Leu Asn Asp
50 55 60

Ser Ile Leu Ile Glu Asp Ile Lys Leu Pro Trp Phe Ala Asp Phe Ser
65 70 75 80

Asn Pro Ile Ile Arg Ile Val Lys Lys Val Ile Pro Arg Ile Tyr Phe
85 90 95

Asn Leu Met Ala Ser Lys Gly Tyr Leu Leu Trp Gln Lys Asn Ser Tyr
100 105 110

Ile Asn Ile Pro Ala Ile Arg Asn Asn Glu Ile Ile Val Asn Gly Trp
115 120 125

37847-517 Sequence Listing_ST25 (003)

Trp Gln Asp Val Arg Phe Phe His Asp Val Glu Ala Glu Leu Ser Asn
130 135 140

Glu Ile Val Pro Thr Thr Lys Pro Ile Ser Glu Asn Glu Tyr Leu Tyr
145 150 155 160

Asn Ile Ala Glu Arg Glu Asn Ser Val Cys Val Ser Ile Arg Gly Gly
165 170 175

Asn Tyr Leu Val Pro Lys Val Lys Lys Lys Leu Phe Val Cys Asp Lys
180 185 190

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

Ala Ile Phe Ile Val Phe Ser Asp Asp Leu Glu Trp Val Lys Ser Tyr
210 215 220

Ile Lys Leu Glu Glu Lys Phe Pro Glu Cys Lys Phe Tyr Tyr Glu Ser
225 230 235 240

Gly Lys Asp Thr Val Glu Glu Lys Leu Arg Met Met Thr Lys Cys Lys
245 250 255

His Phe Ile Ile Ser Asn Ser Ser Phe Ser Trp Trp Ala Gln Tyr Leu
260 265 270

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

Asn Gly Asp Lys Asn Gly Leu Tyr Ile Asp Asp Trp Ile Leu Ile Pro
290 295 300

Thr Gln Thr Lys Asp Met
305 310

<210> 262

<211> 291

<212> PRT

37847-517 Sequence Listing_ST25 (003)

<213> Geobacter lovleyi

<400> 262

Met Ile Thr Val Leu Leu Asn Gly Gly Leu Gly Asn Gln Leu Phe Gln
1 5 10 15

Tyr Ala Ala Gly Arg Ala Leu Ala Glu Lys His Asp Val Glu Leu Leu
20 25 30

Leu Asp Leu Ser Arg Leu Gln His Pro Lys Pro Gly Asp Thr Pro Arg
35 40 45

Cys Phe Glu Leu Ala Pro Phe Asn Ile Lys Ala Ser Leu Leu Ala Glu
50 55 60

Glu Gly Arg Gln Pro Leu Gly Ser Tyr Gln Ala Cys Met His Arg Leu
65 70 75 80

Leu Leu Lys Ala Ser Ile Pro Leu Trp Gly Ser Ile Ile Leu Lys Glu
85 90 95

Gln Gly Cys Gly Phe Asp Pro Leu Ile Phe Arg Ala Pro Ser Ser Cys
100 105 110

Ile Leu Asp Gly Phe Trp Gln Ser Glu Cys Tyr Phe Lys Gln Ile Thr
115 120 125

Ser Leu Leu Gln Gln Glu Leu Ser Leu Lys Ala Pro Ser Pro Ala Leu
130 135 140

Arg Lys Ala Ser Ser Val Leu Ser Asp Ala Thr Val Ala Val His Val
145 150 155 160

Arg Arg Gly Asp Tyr Val Thr Asn Pro Ala Ala Ala Ser Phe His Gly
165 170 175

Ile Cys Ser Gln Asp Tyr Tyr Gln Ala Ala Val Ala Asn Ile Leu Thr
180 185 190

37847-517 Sequence Listing_ST25 (003)

Ser Tyr Pro Asp Ser Gln Phe Leu Val Phe Ser Asp Asp Pro Ala Trp
195 200 205

Cys Gln Glu His Leu Asp Leu Gly Gln Pro Phe Arg Leu Ala Ala Asp
210 215 220

Phe Gly Leu Asn Gly Ser Ala Glu Glu Leu Val Leu Ile Ser Arg Cys
225 230 235 240

Ala His Gln Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala Trp
245 250 255

Leu Asn Pro Ser Pro His Lys Leu Val Val Ala Pro Cys Arg Trp Phe
260 265 270

Thr Asp Pro Ala Ile Thr Thr Asn Asp Leu Leu Pro Glu Thr Trp Val
275 280 285

Arg Leu Pro
290

<210> 263
<211> 290
<212> PRT
<213> Lachnospiraceae bacterium NK4A136

<400> 263

Met Val Ile Ser His Leu Ser Gly Gly Phe Gly Asn Gln Leu Tyr Ser
1 5 10 15

Tyr Ala Phe Ala Tyr Ala Val Ala Lys Ala Arg Lys Glu Glu Leu Trp
20 25 30

Ile Asp Thr Ala Ile Gln Asp Ala Pro Trp Phe Phe Arg Asn Pro Asp
35 40 45

Ile Leu Asn Leu Asn Ile Lys Tyr Asp Lys Arg Val Ser Tyr Lys Ile
50 55 60

Gly Glu Lys Lys Ile Asp Lys Ile Phe Asn Arg Ile Asn Phe Arg Asn

2015259088 28 May 2018

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

275
Leu Lys
290

<210> 264
<211> 269
<212> PRT
<213> Bacteroides vulgatus

<400> 264

Met Phe His Tyr Lys Phe Leu Leu Phe Gly Gly Gly Leu Gly Asn Gln
1 5 10 15

Ile Phe Glu Tyr Tyr Phe Tyr Leu Trp Leu Arg Lys Lys Tyr Pro Asn
20 25 30

Ile Val Phe Leu Gly Cys Tyr Arg Lys Ala Ser Phe Lys Ala His Asn
35 40 45

Gly Leu Glu Ile Ser Asp Val Phe Asp Val Asp Leu Pro Asn Asp Gly
50 55 60

Gly Leu Ser Gly Arg Phe Ile Ser Tyr Val Leu Ser Val Leu Ser Arg
65 70 75 80

Ile Ile Pro Ser Leu Ser Met Lys Ala Asn Thr Glu Tyr Ser Ser Lys
85 90 95

Tyr Leu Leu Ile Asn Ala Tyr Gln Pro Asn Leu Leu Phe Tyr Leu Asn
100 105 110

Glu Glu Lys Ile Lys Phe Arg Pro Phe Lys Leu Asp Glu Val Asn Arg
115 120 125

Arg Leu Leu Asn Ser Ile Lys Met Glu Ser Ser Val Ser Ile His Val
130 135 140

Arg Arg Gly Asp Tyr Leu Phe Gly Gln Tyr Arg Asp Ile Tyr Ser Asn
145 150 155 160

37847-517 Sequence Listing_ST25 (003)

Ile Cys Thr Leu Ala Tyr Tyr Gln Lys Ala Val Asp Lys Cys Lys Gly
 165 170 175

Ile Leu Glu Ser Pro Arg Phe Phe Val Phe Ser Asp Asp Ile Glu Trp
 180 185 190

Ala Arg Asp Val Phe Val Gly Arg Glu Tyr Glu Phe Val Ser Asn Asn
 195 200 205

Ile Gly Lys Asn Ser Phe Ile Asp Met Phe Leu Met Ser Asn Cys Lys
 210 215 220

Ile Gln Ile Ile Ala Asn Ser Thr Phe Ser Tyr Trp Ala Ala Tyr Leu
 225 230 235 240

Ser Asn Ser Leu Val Lys Ile Tyr Pro Ala Lys Trp Ile Asn Gly Ile
 245 250 255

Glu Arg Pro Asn Ile Phe Pro Asp Asn Trp Ile Gly Leu
 260 265

<210> 265
 <211> 287
 <212> PRT
 <213> Planctomyces brasiliensis

<400> 265

Met Ile Ile Ala Arg Ile Glu Asn Gly Leu Gly Asn Gln Leu Phe Lys
 1 5 10 15

Tyr Ala Ala Gly Arg Ala Leu Ser Leu Lys His Arg Thr Ser Leu Tyr
 20 25 30

Thr Ile Pro Gly Ser Val Arg Lys Pro His Glu Thr Phe Ile Leu Ser
 35 40 45

Lys Tyr Phe Asn Val Gln Ala Lys Ser Val Ser Pro Phe Leu Leu Gln
 50 55 60

37847-517 Sequence Listing_ST25 (003)

Thr Gly Phe Arg Leu Arg Leu Leu Lys Gly Tyr Glu Asn His Ser Phe
65 70 75 80

Gly Phe Asp Pro Arg Phe Glu Thr Thr Arg Asn Asn Thr Val Val Ser
85 90 95

Gly Asn Phe Gln Ser Ala Arg Tyr Phe Leu Pro Phe Phe Asp Gln Ile
100 105 110

Asn Arg Glu Leu Thr Leu Lys Pro Glu Val Val Asp Gly Leu Glu Ser
115 120 125

Val Tyr Pro His Val Leu Glu Ser Leu Arg Thr Pro Asn Ser Val Cys
130 135 140

Val His Ile Arg Leu Gly Asp Tyr Val Ser Ser Gly Tyr Asp Ile Cys
145 150 155 160

Gly Pro Glu Tyr Tyr Ala Lys Ala Ile Ser Arg Leu Gln Gln Leu His
165 170 175

Gly Glu Leu Arg Ala Phe Val Phe Ser Asp Thr Pro Gln Ala Ala Ser
180 185 190

Arg Phe Leu Pro Ala Asp Ile Asp Ala Gln Ile Met Ser Glu Phe Pro
195 200 205

Glu Val Arg Asp Ala Ala Arg Ser Leu Thr Val Glu Arg Ser Thr Ile
210 215 220

Arg Asp Tyr Phe Leu Met Gln Gln Cys Arg His Phe Val Ile Pro Asn
225 230 235 240

Ser Ser Phe Ser Tyr Trp Ala Ala Leu Leu Ser Ser Ser Asp Gly Asp
245 250 255

Val Ile Tyr Pro Asn Arg Trp Tyr Ile Asp Ile Asp Thr Ser Pro Arg
260 265 270

37847-517 Sequence Listing_ST25 (003)

Asp Leu Gly Leu Ala Pro Ala Glu Trp Thr Pro Ile Pro Leu Thr
 275 280 285

<210> 266

<211> 309

<212> PRT

<213> Butyrivibrio sp. AE2015

<400> 266

Met Ile Ile Leu Gln Ile Ala Gly Gly Leu Gly Asn Gln Met Gln Gln
 1 5 10 15

Tyr Ala Leu Tyr Arg Lys Leu Leu Lys Cys Gly Lys Thr Val Lys Leu
 20 25 30

Asp Leu Ser Trp Phe Gly Pro Glu Ile Gln Lys Asn Met Leu Ala Pro
 35 40 45

Arg Glu Phe Glu Leu Val Leu Phe Lys Asp Leu Pro Phe Glu Ile Cys
 50 55 60

Thr Lys Glu Glu Lys Asp Ala Leu Ile Lys Gln Asn Leu Phe Gln Lys
 65 70 75 80

Ile Ala Gly Lys Val Ser Gln Lys Leu Gly Lys Ser Ala Ser Ser Asn
 85 90 95

Ala Lys Val Phe Val Glu Thr Lys Met Tyr His Glu Glu Ile Phe Asp
 100 105 110

Leu Asp Asp Val Tyr Ile Thr Gly Tyr Phe Ala Cys Gln Tyr Tyr Tyr
 115 120 125

Asp Asp Val Met Ala Glu Leu Gln Asp Leu Phe Val Phe Pro Ser His
 130 135 140

Ser Ile Pro Glu Leu Asp Gln Arg Asn Ala Val Leu Ala Ser Lys Met
 145 150 155 160

37847-517 Sequence Listing_ST25 (003)

Glu Lys Glu Asn Ser Val Ser Val His Ile Arg Arg Gly Asp Tyr Leu
165 170 175

Ser Pro Glu Asn Val Gly Ile Leu Gly Asn Ile Ala Ser Asp Lys Tyr
180 185 190

Tyr Glu Ser Ala Met Asn Tyr Phe Leu Glu Lys Asp Glu Asn Thr His
195 200 205

Phe Tyr Ile Phe Thr Asn Asp His Glu Tyr Ala Arg Glu His Tyr Ser
210 215 220

Asp Glu Ser Arg Tyr Thr Ile Ile Asp Trp Asn Thr Gly Lys Asn Ser
225 230 235 240

Leu Gln Asp Leu Met Leu Met Ser His Cys Lys Gly Asn Ile Cys Ala
245 250 255

Asn Ser Thr Phe Ser Phe Trp Gly Ala Arg Leu Asn Lys Arg Pro Asp
260 265 270

Arg Glu Leu Val Arg Thr Leu Lys Met Arg Asn Asn Gln Glu Ala Gln
275 280 285

Pro Glu Ile Met His Glu Tyr Trp Lys Asn Trp Ile Leu Ile Asp Glu
290 295 300

Asn Gly Val Ile Val
305

<210> 267
<211> 291
<212> PRT
<213> Roseovarius nubinhibens

<400> 267

Met Thr Asp Thr Pro Pro Pro Ser Gln Val Ile Thr Ser Arg Leu Phe
1 5 10 15

Gly Gly Ala Gly Asn Gln Leu Phe Gln Tyr Ala Ala Gly Arg Ala Leu

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

20

25

30

Ala Asp Arg Leu Gly Cys Asp Leu Met Ile Asp Ala Arg Tyr Val Ala
35 40 45

Gly Ser Arg Asp Arg Gly Asp Cys Phe Thr His Phe Ala Lys Ala Arg
50 55 60

Leu Arg Arg Asp Val Ala Leu Pro Pro Ala Lys Ser Asp Gly Pro Leu
65 70 75 80

Arg Tyr Ala Leu Trp Arg Lys Phe Gly Arg Ser Pro Arg Phe His Arg
85 90 95

Glu Arg Gly Leu Gly Val Asp Pro Glu Phe Phe Asn Leu Pro Arg Gly
100 105 110

Thr Tyr Leu His Gly Tyr Trp Gln Ser Glu Gln Tyr Phe Gly Pro Asp
115 120 125

Thr Asp Ala Leu Arg Arg Asp Leu Thr Leu Thr Thr Ala Leu Asp Ala
130 135 140

Pro Asn Ala Ala Met Ala Ala Gln Ile Asp Ala Ala Pro Cys Pro Val
145 150 155 160

Ser Phe His Val Arg Arg Gly Asp Tyr Ile Ala Ala Gly Ala Tyr Ala
165 170 175

Ala Cys Thr Pro Asp Tyr Tyr Arg Ala Ala Ala Asp His Leu Ala Thr
180 185 190

Thr Leu Gly Lys Pro Leu Thr Cys Phe Ile Phe Ser Asn Asp Pro Ala
195 200 205

Trp Ala Arg Asp Asn Leu Asp Leu Gly Gln Asp Gln Val Ile Val Asp
210 215 220

Leu Asn Asp Glu Ala Thr Gly His Phe Asp Met Ala Leu Met Ala Arg

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

225 230 235 240

Cys Ala His His Val Ile Ala Asn Ser Thr Phe Ser Trp Trp Gly Ala
245 250 255

Trp Leu Asn Pro Asp Pro Asp Lys Leu Val Val Ala Pro Arg Asn Trp
260 265 270

Phe Ala Thr Gln Ala Leu His Asn Pro Asp Leu Ile Pro Glu Gln Trp
275 280 285

His Arg Leu
290

<210> 268
<211> 300
<212> PRT
<213> Eubacterium sp. CAG:581

<400> 268

Met Ile Glu Val Asn Ile Val Gly Gln Leu Gly Asn Gln Met Phe Glu
1 5 10 15

Tyr Ala Cys Ala Arg Gln Leu Gln Lys Lys Tyr Gly Gly Glu Ile Val
20 25 30

Leu Asn Thr Tyr Glu Met Arg Lys Glu Thr Pro Asn Phe Lys Leu Ser
35 40 45

Ile Leu Asp Tyr Lys Leu Ser Glu Asn Val Lys Ile Ile Ser Asp Lys
50 55 60

Pro Leu Ser Ser Ala Asn Ala Asn Asn Tyr Leu Val Lys Ile Met Arg
65 70 75 80

Gln Tyr Phe Pro Asn Trp Tyr Phe Asn Phe Met Ala Lys Arg Gly Thr
85 90 95

Phe Val Trp Lys Ser Ala Arg Lys Tyr Lys Glu Leu Pro Glu Leu Asn
100 105 110

37847-517 Sequence Listing_ST25 (003)

Glu Gln Leu Ser Lys His Ile Val Leu Asn Gly Tyr Trp Gln Cys Asp
 115 120 125

Lys Tyr Phe Asn Asp Val Val Asp Thr Ile Arg Glu Asp Phe Thr Pro
 130 135 140

Lys Tyr Pro Leu Lys Ala Glu Asn Glu Gln Leu Leu Glu Lys Ile Lys
 145 150 155 160

Ser Thr Glu Ser Val Cys Val Thr Ile Arg Arg Gly Asp Phe Met Asn
 165 170 175

Glu Lys Asn Lys Asp Thr Phe Tyr Ile Cys Asp Asp Asp Tyr Phe Asn
 180 185 190

Lys Ala Leu Ser Lys Ile Lys Glu Leu Cys Pro Asp Cys Thr Phe Phe
 195 200 205

Gly Phe Ser Asp Asp Val Glu Trp Ile Lys Lys Asn Val Asn Phe Pro
 210 215 220

Gly Glu Val Tyr Phe Glu Ser Gly Asn Asp Pro Val Trp Glu Lys Leu
 225 230 235 240

Arg Leu Met Ser Ala Cys Lys His Phe Val Leu Ser Asn Ser Ser Phe
 245 250 255

Ser Trp Trp Ala Gln Tyr Leu Ser Asp Asn Asn Asn Lys Ile Val Val
 260 265 270

Ala Pro Asp Ile Trp Tyr Lys Thr Gly Asp Pro Lys Lys Thr Ala Leu
 275 280 285

Tyr Gln Asp Gly Trp Asn Leu Ile His Ile Gly Asp
 290 295 300

<210> 269

<211> 303

37847-517 Sequence Listing_ST25 (003)

<212> PRT

<213> Providencia alcalifaciens

<400> 269

Met Lys Ile Asn Gly Lys Glu Ser Ser Met Lys Ile Lys Gln Lys Lys
 1 5 10 15

Ile Ile Ser His Leu Ile Gly Gly Leu Gly Asn Gln Leu Phe Gln Tyr
 20 25 30

Ala Thr Ser Tyr Ala Leu Ala Lys Glu Asn Asn Ala Lys Ile Val Ile
 35 40 45

Asp Asp Arg Leu Phe Lys Lys Tyr Lys Leu His Gly Gly Tyr Arg Leu
 50 55 60

Asp Lys Leu Asn Ile Ile Gly Glu Lys Ile Ser Ser Ile Asp Lys Leu
 65 70 75 80

Leu Phe Pro Leu Ile Leu Cys Lys Leu Ser Gln Lys Glu Asn Phe Ile
 85 90 95

Phe Lys Ser Thr Lys Lys Phe Ile Leu Glu Lys Lys Thr Ser Ser Phe
 100 105 110

Lys Tyr Leu Thr Phe Ser Asp Lys Glu His Thr Lys Met Leu Ile Gly
 115 120 125

Tyr Trp Gln Asn Ala Ile Tyr Phe Gln Lys Tyr Phe Ser Glu Leu Lys
 130 135 140

Glu Met Phe Val Pro Leu Asp Ile Ser Gln Glu Gln Leu Asp Leu Ser
 145 150 155 160

Ile Gln Ile His Ala Gln Gln Ser Val Ala Leu His Val Arg Arg Gly
 165 170 175

Asp Tyr Ile Ser Asn Lys Asn Ala Leu Ala Met His Gly Ile Cys Ser
 180 185 190

37847-517 Sequence Listing_ST25 (003)

Ile Asp Tyr Tyr Lys Asn Ser Ile Gln His Ile Asn Ala Lys Leu Glu
195 200 205

Lys Pro Phe Phe Tyr Ile Phe Ser Asn Asp Lys Leu Trp Cys Glu Glu
210 215 220

Asn Leu Thr Pro Leu Phe Asp Gly Asn Phe His Ile Val Glu Asn Asn
225 230 235 240

Ser Gln Glu Ile Asp Leu Trp Leu Ile Ser Gln Cys Gln His His Ile
245 250 255

Ile Ala Asn Ser Thr Phe Ser Trp Trp Gly Ala Trp Leu Ala Asn Ser
260 265 270

Asp Ser Gln Ile Val Ile Thr Pro Asp Pro Trp Phe Asn Lys Glu Ile
275 280 285

Asp Ile Pro Ser Pro Val Leu Ser His Trp Leu Lys Leu Lys Lys
290 295 300

<210> 270
<211> 301
<212> PRT
<213> Salmonella enterica

<400> 270

Met Phe Ser Cys Leu Ser Gly Gly Leu Gly Asn Gln Met Phe Gln Tyr
1 5 10 15

Ser Ala Ala Tyr Ile Leu Lys Lys Asn Ile Cys His Ala Gln Leu Ile
20 25 30

Ile Asp Asp Ser Tyr Phe Tyr Cys Gln Pro Gln Lys Asp Thr Pro Arg
35 40 45

Asn Phe Glu Ile Asn Gln Phe Asn Ile Val Phe Asp Arg Val Thr Thr
50 55 60

37847-517 Sequence Listing_ST25 (003)

Asp	Glu	Glu	Lys	Arg	Ala	Ile	Ser	Lys	Leu	Arg	Lys	Phe	Lys	Lys	Ile	65	70	75	80
Pro	Leu	Pro	Leu	Phe	Lys	Ser	Asn	Val	Ile	Thr	Glu	Phe	Leu	Phe	Gly	85	90	95	
Lys	Ser	Leu	Leu	Thr	Asp	Glu	Asp	Phe	Tyr	Lys	Val	Leu	Lys	Lys	Asn	100	105	110	
Gln	Phe	Thr	Val	Lys	Met	Asn	Ala	Cys	Leu	Phe	Ser	Leu	Tyr	Gln	Asp	115	120	125	
Ser	Ser	Leu	Ile	Asn	Lys	Tyr	Arg	Asp	Leu	Ile	Leu	Pro	Leu	Phe	Thr	130	135	140	
Ile	Asn	Asp	Glu	Leu	Leu	Gln	Val	Cys	Gln	Gln	Leu	Asp	Ser	Tyr	Gly	145	150	155	160
Phe	Ile	Cys	Glu	His	Thr	Asn	Thr	Thr	Ser	Leu	His	Ile	Arg	Arg	Gly	165	170	175	
Asp	Tyr	Val	Thr	Asn	Pro	His	Ala	Ala	Lys	Phe	His	Gly	Thr	Leu	Ser	180	185	190	
Met	Asn	Tyr	Tyr	Ser	Gln	Ala	Met	Asn	Tyr	Val	Asp	His	Lys	Leu	Gly	195	200	205	
Lys	Gln	Leu	Phe	Ile	Ile	Phe	Ser	Asp	Asp	Val	Gln	Trp	Ala	Ala	Glu	210	215	220	
Lys	Phe	Gly	Gly	Arg	Ser	Asp	Cys	Tyr	Ile	Val	Asn	Asn	Val	Asn	Cys	225	230	235	240
Gln	Phe	Ser	Ala	Ile	Asp	Met	Tyr	Leu	Met	Ser	Leu	Cys	Asn	Asn	Asn	245	250	255	
Ile	Ile	Ala	Asn	Ser	Thr	Tyr	Ser	Trp	Trp	Gly	Ala	Trp	Leu	Asn	Lys	260	265	270	

37847-517 Sequence Listing_ST25 (003)

Ser Glu Glu Lys Leu Val Ile Ala Pro Arg Lys Trp Phe Ala Glu Asp
275 280 285

Lys Glu Ser Leu Leu Ala Val Asn Asp Trp Ile Ser Ile
290 295 300

<210> 271

<211> 289

<212> PRT

<213> *Sulfurospirillum deleyianum*

<400> 271

Met Ile Ile Ile Lys Ile Met Gly Gly Leu Ala Ser Gln Leu His Lys
1 5 10 15

Tyr Ser Val Gly Arg Ala Leu Ser Leu Lys Tyr Asn Thr Glu Leu Lys
20 25 30

Leu Asp Ile Phe Trp Phe Asp Asn Ile Ser Gly Ser Asp Thr Ile Arg
35 40 45

Glu Tyr His Leu Asp Lys Tyr Asn Val Val Ala Lys Ile Ala Thr Glu
50 55 60

Gln Glu Ile Lys Gln Phe Lys Pro Asn Lys Tyr Leu Leu Lys Ile Asn
65 70 75 80

Asn Leu Phe Gln Lys Phe Thr Asn Trp Lys Ile Asn Tyr Arg Asn Tyr
85 90 95

Cys Asn Glu Ser Phe Ile Ser Leu Glu Asn Phe Asn Leu Leu Pro Asp
100 105 110

Asn Ile Tyr Val Glu Gly Glu Trp Ser Gly Asp Arg Tyr Phe Ser His
115 120 125

Ile Lys Glu Ile Leu Gln Lys Glu Leu Thr Leu Lys Ser Glu Tyr Met
130 135 140

Asp Ser Thr Asn His Phe Leu Ala Lys Gln Ser Ser Asp Phe Ala His

37847-517 Sequence Listing_ST25 (003)

145 150 155 160

Asp Asp Asn Ala Ser Lys Leu His Cys Thr Cys Ser Leu Glu Tyr Tyr
165 170 175

Lys Lys Ala Leu Gln Tyr Ile Ser Lys Asn Leu Leu Lys Met Lys Leu
180 185 190

Leu Ile Phe Ser Asp Asp Leu Asp Trp Leu Lys Pro Asn Phe Asn Phe
195 200 205

Leu Asp Asn Val Glu Phe Glu Phe Val Glu Gly Phe Gln Asp Tyr Glu
210 215 220

Glu Phe His Leu Met Thr Leu Ser Lys His Asn Ile Ile Ala Asn Ser
225 230 235 240

Gly Phe Ser Leu Phe Phe Ala Trp Leu Asn Ile Asn His Asn Lys Ile
245 250 255

Ile Ile Ser Leu Ser Glu Trp Val Phe Glu Glu Lys Leu Asn Lys Tyr
260 265 270

Ile Ile Asp Asn Ile Lys Asp Lys Asn Ile Leu Phe Leu Glu Asn Leu
275 280 285

Glu

<210> 272

<211> 309

<212> PRT

<213> Pseudovibrio sp. FO-BEG1

<400> 272

Met Ser Val Ala Ser Gln Val Arg Ile Ser Gly Ala Ala Arg Arg Arg
1 5 10 15

Lys Leu Lys Pro Thr Leu Ile Val Arg Ile Arg Gly Gly Ile Gly Asn
20 25 30

37847-517 Sequence Listing_ST25 (003)

Gln Leu Phe Gln Tyr Ala Leu Gly Arg Lys Ile Ala Leu Glu Thr Gly
 35 40 45
 Met Lys Leu Arg Phe Asp Arg Ser Glu Tyr Asp Gln Tyr Phe Asn Arg
 50 55 60
 Ser Tyr Cys Leu Asn Leu Phe Lys Thr Gln Gly Leu Ser Ala Thr Glu
 65 70 75 80
 Ser Glu Met Ser Ala Val Leu Trp Pro Ala Gln Ser Phe Gly Gln Thr
 85 90 95
 Val Lys Leu Cys Arg Lys Phe Tyr Pro Phe Tyr Gln Arg Arg Tyr Ile
 100 105 110
 Arg Glu Asp Glu Leu Leu Gln Asp Ser Glu Thr Pro Val Leu Lys Gln
 115 120 125
 Ser Ala Tyr Leu Asp Gly Tyr Trp Gln Thr Trp Glu Ile Pro Phe Ser
 130 135 140
 Ile Met Glu Gln Leu Arg Asp Glu Ile Thr Leu Lys Lys Pro Met Val
 145 150 155 160
 Leu Glu Arg Leu Lys Leu Leu Gln Arg Ile Lys Ser Gly Pro Ser Ala
 165 170 175
 Ala Leu His Val Arg Tyr Gly Asp Tyr Ser Gln Ala His Asn Leu Gln
 180 185 190
 Asn Phe Gly Leu Cys Ser Ala Gly Tyr Tyr Lys Gly Ala Met Asp Phe
 195 200 205
 Leu Thr Glu Arg Val Pro Gly Leu Thr Phe Tyr Val Phe Ser Asp Ser
 210 215 220
 Pro Glu Arg Ala Arg Glu Val Val Pro Gln Gln Glu Asn Val Tyr Phe
 225 230 235 240

37847-517 Sequence Listing_ST25 (003)

Ser Asp Pro Met Gln Asp Gly Lys Asp His Glu Asp Leu Met Val Met
245 250 255

Ser Ser Cys Asp His Ile Val Thr Ala Asn Ser Thr Phe Ser Trp Trp
260 265 270

Ala Ala Phe Leu Asn Gly Asn Glu Asp Lys His Val Ile Ala Pro Leu
275 280 285

Lys Trp Phe Lys Asn Pro Asn Leu Asp Asp Ser Leu Ile Val Pro Pro
290 295 300

His Trp Gln Arg Leu
305

<210> 273
<211> 264
<212> PRT
<213> Prevotella sp. oral taxon 472

<400> 273

Met Lys Ile Val Cys Ile Lys Gly Gly Leu Gly Asn Gln Leu Phe Glu
1 5 10 15

Tyr Cys Arg Tyr His Gly Leu Leu Arg Gln His Asn Asn His Gly Val
20 25 30

Tyr Leu His Tyr Asp Arg Arg Arg Thr Lys Gln His Gly Gly Val Trp
35 40 45

Leu Asp Lys Ala Phe Leu Ile Thr Leu Pro Thr Glu Pro Trp Arg Val
50 55 60

Lys Leu Met Val Met Ala Leu Lys Met Leu Arg Lys Leu His Leu Phe
65 70 75 80

Lys Arg Leu Tyr Arg Glu Asp Asp Pro Arg Ala Val Leu Ile Asp Asp
85 90 95

37847-517 Sequence Listing_ST25 (003)

Tyr Ser Gln His Lys Gln Phe Ile Thr Asn Ala Ala Glu Ile Leu Asn
 100 105 110

Phe Arg Pro Phe Ala Gln Leu Asp Tyr Val Asp Glu Ile Thr Ser Glu
 115 120 125

Pro Phe Ala Val Ser Val His Val Arg Arg Gly Asp Tyr Leu Leu Pro
 130 135 140

Ala Asn Lys Ala Asn Phe Gly Val Cys Ser Val His Tyr Tyr Leu Ser
 145 150 155 160

Ala Ala Val Ala Val Arg Glu Arg His Pro Asp Ala Arg Phe Phe Val
 165 170 175

Phe Ser Asp Asp Ile Glu Trp Ala Lys Met Asn Leu Asn Leu Pro Asn
 180 185 190

Cys Val Phe Val Glu His Ala Gln Pro Gln Pro Asp His Ala Asp Leu
 195 200 205

Tyr Leu Met Ser Leu Cys Lys Gly His Ile Ile Ala Asn Ser Thr Phe
 210 215 220

Ser Phe Trp Gly Ala Tyr Leu Ser Met Gly Ser Ser Ala Ile Ala Ile
 225 230 235 240

Tyr Pro Lys Gln Trp Phe Ala Glu Pro Thr Trp Asn Ala Pro Asp Ile
 245 250 255

Phe Leu Gly His Trp Ile Ala Leu
 260

<210> 274
 <211> 304
 <212> PRT
 <213> Butyrivibrio fibrisolvens

<400> 274

37847-517 Sequence Listing_ST25 (003)

Met	Leu	Ile	Ile	Arg	Val	Ala	Gly	Gly	Leu	Gly	Asn	Gln	Met	Gln	Gln	1	5	10	15
Tyr	Ala	Met	Tyr	Arg	Lys	Leu	Lys	Ser	Leu	Gly	Lys	Glu	Val	Lys	Leu	20	25	30	
Asp	Leu	Ser	Trp	Phe	Asp	Val	Glu	Asn	Gln	Glu	Gly	Gln	Leu	Ala	Pro	35	40	45	
Arg	Lys	Cys	Glu	Leu	Lys	Tyr	Phe	Asp	Gly	Val	Asp	Phe	Glu	Glu	Cys	50	55	60	
Thr	Asp	Ala	Glu	Arg	Ala	Tyr	Phe	Thr	Lys	Arg	Ser	Ile	Leu	Thr	Lys	65	70	75	80
Ala	Leu	Asn	Lys	Val	Phe	Pro	Ala	Thr	Cys	Lys	Ile	Phe	Glu	Glu	Thr	85	90	95	
Glu	Met	Phe	His	Pro	Glu	Ile	Tyr	Ser	Phe	Lys	Asp	Lys	Tyr	Leu	Glu	100	105	110	
Gly	Tyr	Phe	Leu	Cys	Asn	Lys	Tyr	Tyr	Asp	Asp	Ile	Leu	Pro	Phe	Ile	115	120	125	
Gln	Asn	Glu	Ile	Val	Phe	Pro	Lys	His	Ser	Asp	Pro	Lys	Met	Gln	Lys	130	135	140	
Arg	Asn	Glu	Glu	Leu	Met	Glu	Arg	Met	Asp	Gly	Trp	His	Thr	Ala	Ser	145	150	155	160
Ile	His	Leu	Arg	Arg	Gly	Asp	Tyr	Ile	Thr	Glu	Pro	Gln	Asn	Glu	Ala	165	170	175	
Leu	Phe	Gly	Asn	Ile	Ala	Thr	Asp	Ala	Tyr	Tyr	Asp	Ala	Ala	Ile	Arg	180	185	190	
Tyr	Val	Leu	Asp	Lys	Asp	Tyr	Gln	Thr	His	Phe	Tyr	Ile	Phe	Ser	Asn	195	200	205	

37847-517 Sequence Listing_ST25 (003)

Asp Pro Glu Tyr Ala Arg Glu His Tyr Ser Asp Glu Ser Arg Tyr Thr
210 215 220

Ile Val Thr Gly Asn Asp Gly Asp Asn Ser Leu Leu Asp Met Glu Leu
225 230 235 240

Met Ser His Cys Arg Tyr Asn Ile Cys Ala Asn Ser Thr Phe Ser Phe
245 250 255

Trp Gly Ala Arg Leu Asn Lys Arg Ser Asp Lys Glu Met Ile Arg Thr
260 265 270

Phe Lys Met Arg Asn Asn Gln Glu Val Thr Ala Arg Glu Met Thr Asp
275 280 285

Tyr Trp Lys Asp Trp Ile Leu Ile Asp Glu Lys Gly Asn Arg Ile Phe
290 295 300

<210> 275
<211> 294
<212> PRT
<213> Lewinella persica

<400> 275

Met Val Ile Ser Arg Leu His Ser Gly Leu Gly Asn Gln Met Phe Gln
1 5 10 15

Tyr Ala Phe Ala Arg Arg Ile Gln Leu Gln Leu Asn Val Lys Leu Arg
20 25 30

Ile Asp Leu Ser Ile Leu Leu Asp Ser Arg Pro Pro Asp Gly Tyr Ile
35 40 45

Lys Arg Glu Tyr Asp Leu Asp Ile Phe Lys Leu Ser Pro Ala Tyr His
50 55 60

Cys Asn Pro Thr Ser Leu Arg Ile Leu Tyr Ala Pro Gly Lys Tyr Arg
65 70 75 80

Trp Ser Gln Val Val Arg Asp Leu Ala Arg Lys Gly Tyr Pro Val Tyr

37847-517 Sequence Listing_ST25 (003)

85 90 95

Met Glu Lys Ser Phe Ser Val Asp Asn Thr Leu Leu Asp Ser Pro Pro
100 105 110

Asp Asn Val Ile Tyr Gln Gly Tyr Trp Gln Ser Glu Arg Tyr Phe Ser
115 120 125

Glu Val Ala Asn Thr Ile Arg Lys Asp Phe Ala Phe Gln His Ser Ile
130 135 140

Gln Pro Gln Ser Glu Ser Leu Ala Arg Glu Ile Arg Lys Glu Asp Ser
145 150 155 160

Val Cys Leu Asn Ile Arg Arg Lys Asp Tyr Leu Ala Ser Pro Thr His
165 170 175

Asn Val Thr Asp Glu Thr Tyr Tyr Glu Asn Cys Ile Gln Gln Met Arg
180 185 190

Glu Arg Phe Ser Gly Ala Arg Phe Phe Leu Phe Ser Asp Asp Leu Val
195 200 205

Trp Cys Arg Glu Phe Phe Ala Asp Phe His Asp Val Val Ile Val Gly
210 215 220

His Asp His Ala Gly Pro Lys Phe Gly Asn Tyr Leu Gln Leu Met Ala
225 230 235 240

Gln Cys His His Tyr Ile Ile Pro Asn Ser Thr Phe Ala Trp Trp Ala
245 250 255

Ala Trp Leu Gly Glu Arg Thr Gly Ser Val Ile Met Ala Pro Glu Arg
260 265 270

Trp Phe Gly Thr Asp Glu Phe Asp Tyr Arg Asp Val Val Pro Glu Arg
275 280 285

Trp Leu Lys Val Pro Asn

37847-517 Sequence Listing_ST25 (003)

290

<210> 276
<211> 916
<212> DNA
<213> Artificial Sequence

<220>
<223> artifical construct

<400> 276
cagtcagtca gaattcaaga aggagatata catatgaaaa tcgtcaaaat cctgggcggt 60
ctgggcaatc agatgttcca gtatgctctg tacctgagcc tgcaagaaag ttttccaaaa 120
gaacgtgtgg ccctggacct gtcctccttc cacggctatc acctgcataa tggctttgag 180
ctggagaaca ttttctccgt taccgctcag aaagcatccg ccgcagatat catgcgtatt 240
gcttattact acccgaacta tctgctgtgg cgcattggca aacgttttct gccgcgtcgt 300
aaaggatatgt gcctggaatc tagctccctg cgtttcgatg aaagcgttct gcgtcaggaa 360
ggtaaccgtt attttgacgg ttactggcaa gacgaacgct acttcgcagc ctatcgtgaa 420
aaagtgtga aggctttcac ctttcctgca ttcaaacgcg cagaaaacct gagcctgctg 480
gaaaaactgg acgaaaacag cattgctctg catgttcgtc gcggtgatta cgtaggtaat 540
aacctgtacc aaggcatctg tgacctggac tactaccgta ccgctatcga gaaaatgtgt 600
gcacacgtta ctccgtctct gttttgtatc ttttccaacg acatcacgtg gtgccagcag 660
cacctgcaac cgtacctgaa ggcccctgtg gtgtacgtta cttggaacac cgggtgttgaa 720
tcctaccgcg atatgcagct gatgtcctgc tgcgcacata acatcatcgc gaatagctcc 780
ttctcttggt ggggtgcttg gctgaatcag aaccgtgaaa aagttgttat cgccccgaaa 840
aaatggctga acatggaaga atgtcacttc acgctgccgg caagctggat caaaatttag 900
ctcgagtgac tgactg 916

<210> 277
<211> 886
<212> DNA
<213> Artificial Sequence

<220>
<223> artifical construct

37847-517 Sequence Listing_ST25 (003)

<400> 277
 cagtcagtca gaattcaaga aggagatata catatggtga ttatcaaaat gatgggtggt 60
 ctgggcaacc agatgttcca gtacgcactg tacaaagcat tcgagcagaa gcacatcgat 120
 gtgtatgcag acctggcatg gtacaaaaac aaatccgtga aatttgaact gtacaacttc 180
 ggcatataaaa tcaacgtagc atccgagaaa gacatcaacc gtctgagcga ttgccaggcg 240
 gactttgttt cccgcatccg ccgtaaaatc tttggtaaaa aaaagagctt cgtatctgaa 300
 aaaaatgact cctgctatga aaacgacatc ctgcgtatgg acaacgttta tctgagcggg 360
 tattggcaga ccgaaaaata cttctctaac acgcgtgaga agctgctgga ggattattcc 420
 ttcgctctgg taaactctca ggtgtccgaa tgggaagact ccattcgcaa caaaaacagc 480
 gttagcatcc atatccgtcg tgggtgattat ctacaggggcg aactgtatgg tgggtatttc 540
 acctctctgt actacgccga agcaatcgag tacattaaaa tgcgtgttcc gaacgcaaaa 600
 ttcttcgttt tctctgatga cggtgaatgg gttaaacagc aagaagactt caaaggcttc 660
 gtaatcgttg atcgcaacga gtattctagc gctctgtccg atatgtacct gatgtccctg 720
 tgcaagcata acattattgc taactcctct ttcagctggt gggcagcttg gctgaaccgt 780
 aacgaagaaa aaattgtaat cgcgccgcgc cgttggctga acggcaagtg caccacagat 840
 atctggtgta aaaaatggat tcgtatctag ctcgagtgc tgactg 886

<210> 278
 <211> 919
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> artificial construct

<400> 278

Cys Ala Gly Thr Cys Ala Gly Thr Cys Ala Gly Ala Ala Thr Thr Cys
 1 5 10 15

Ala Ala Gly Ala Ala Gly Gly Ala Gly Ala Thr Ala Thr Ala Cys Ala
 20 25 30

Thr Ala Thr Gly Gly Thr Gly Ala Thr Cys Gly Thr Ala Cys Ala Gly

37847-517 Sequence Listing_ST25 (003)

35	40	45																	
Cys	Thr	Gly	Ala	Gly	Cys	Gly	Gly	Cys	Gly	Gly	Thr	Cys	Thr	Gly	Gly				
50					55						60								
Gly	Cys	Ala	Ala	Cys	Cys	Ala	Gly	Ala	Thr	Gly	Thr	Thr	Cys	Gly	Ala				
65				70						75					80				
Ala	Thr	Ala	Cys	Gly	Cys	Gly	Cys	Thr	Gly	Thr	Ala	Cys	Cys	Thr	Gly				
			85						90					95					
Ala	Gly	Cys	Cys	Thr	Gly	Ala	Ala	Ala	Gly	Cys	Ala	Ala	Ala	Ala	Gly				
			100					105						110					
Gly	Cys	Ala	Ala	Ala	Gly	Ala	Ala	Gly	Thr	Gly	Ala	Ala	Ala	Ala	Thr				
		115					120					125							
Thr	Gly	Ala	Cys	Gly	Ala	Thr	Gly	Thr	Thr	Ala	Cys	Gly	Thr	Gly	Thr				
	130					135					140								
Thr	Ala	Cys	Gly	Ala	Gly	Gly	Gly	Cys	Cys	Cys	Thr	Gly	Gly	Cys	Ala				
145				150						155					160				
Cys	Cys	Cys	Gly	Thr	Cys	Cys	Gly	Cys	Gly	Thr	Cys	Ala	Ala	Cys	Thr				
				165					170					175					
Gly	Gly	Ala	Thr	Gly	Thr	Thr	Thr	Thr	Thr	Gly	Gly	Thr	Ala	Thr	Cys				
			180					185					190						
Ala	Cys	Gly	Thr	Ala	Cys	Gly	Ala	Thr	Cys	Gly	Cys	Gly	Cys	Gly	Thr				
		195					200					205							
Cys	Thr	Cys	Gly	Thr	Gly	Ala	Gly	Gly	Ala	Gly	Cys	Thr	Gly	Ala	Cys				
	210					215					220								
Thr	Gly	Ala	Gly	Ala	Thr	Gly	Ala	Cys	Gly	Gly	Ala	Cys	Gly	Cys	Gly				
225					230					235					240				
Ala	Gly	Cys	Ala	Thr	Gly	Gly	Ala	Thr	Gly	Cys	Gly	Cys	Thr	Gly	Thr				

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

245 250 255

Cys Thr Cys Gly Thr Gly Thr Thr Cys Gly Thr Cys Gly Cys Ala Ala
260 265 270

Ala Cys Thr Gly Ala Cys Cys Gly Gly Thr Cys Gly Cys Cys Gly Cys
275 280 285

Ala Cys Thr Ala Ala Ala Gly Cys Gly Thr Ala Cys Cys Gly Cys Gly
290 295 300

Ala Ala Cys Gly Cys Gly Ala Cys Ala Thr Cys Ala Ala Cys Thr Thr
305 310 315 320

Cys Gly Ala Thr Cys Cys Ala Cys Thr Gly Gly Thr Thr Ala Thr Gly
325 330 335

Gly Ala Ala Ala Ala Ala Gly Ala Cys Cys Cys Gly Gly Cys Ala Cys
340 345 350

Thr Gly Cys Thr Gly Gly Ala Ala Gly Gly Cys Thr Gly Thr Thr Thr
355 360 365

Cys Cys Ala Gly Thr Cys Thr Gly Ala Cys Ala Ala Ala Thr Ala Cys
370 375 380

Thr Thr Thr Cys Gly Thr Gly Ala Thr Thr Gly Cys Gly Ala Ala Gly
385 390 395 400

Gly Cys Cys Gly Cys Gly Thr Gly Cys Gly Cys Gly Ala Ala Gly Cys
405 410 415

Gly Thr Ala Thr Cys Gly Thr Thr Thr Cys Cys Gly Cys Gly Gly Cys
420 425 430

Ala Thr Thr Gly Ala Ala Thr Cys Cys Gly Gly Cys Gly Cys Gly Thr
435 440 445

Thr Cys Cys Cys Gly Cys Thr Gly Cys Cys Gly Gly Ala Ala Gly Ala

37847-517 Sequence Listing_ST25 (003)

450

455

460

Cys Thr Ala Thr Cys Thr Gly Cys Gly Cys Cys Thr Gly Gly Ala Ala
465 470 475 480

Ala Ala Gly Cys Ala Gly Ala Thr Cys Gly Ala Ala Gly Ala Thr Thr
485 490 495

Gly Thr Cys Ala Gly Thr Cys Cys Gly Thr Ala Thr Cys Cys Gly Thr
500 505 510

Ala Cys Ala Cys Ala Thr Cys Cys Gly Thr Cys Gly Thr Gly Gly Cys
515 520 525

Gly Ala Cys Thr Ala Cys Cys Thr Gly Gly Ala Cys Gly Ala Ala Thr
530 535 540

Cys Thr Cys Ala Thr Gly Gly Thr Gly Gly Thr Cys Thr Gly Thr Ala
545 550 555 560

Cys Ala Cys Cys Gly Gly Cys Ala Thr Thr Thr Gly Thr Ala Cys Thr
565 570 575

Gly Ala Gly Gly Cys Gly Thr Ala Cys Thr Ala Thr Ala Ala Ala Gly
580 585 590

Ala Gly Gly Cys Thr Thr Thr Thr Gly Cys Thr Cys Gly Cys Ala Thr
595 600 605

Gly Gly Ala Ala Cys Gly Thr Cys Thr Gly Gly Thr Thr Cys Cys Gly
610 615 620

Gly Gly Cys Gly Cys Ala Cys Gly Thr Thr Thr Cys Thr Thr Cys Cys
625 630 635 640

Thr Gly Thr Thr Cys Thr Cys Thr Ala Ala Cys Gly Ala Thr Cys Cys
645 650 655

Ala Gly Ala Ala Thr Gly Gly Ala Cys Thr Cys Gly Thr Gly Ala Gly

37847-517 Sequence Listing_ST25 (003)

660

665

670

Cys Ala Cys Thr Thr Thr Gly Ala Gly Ala Gly Cys Ala Ala Gly Ala
675 680 685

Ala Cys Thr Gly Cys Gly Thr Thr Cys Thr Gly Gly Thr Thr Gly Ala
690 695 700

Ala Gly Gly Thr Ala Gly Cys Ala Cys Cys Gly Ala Ala Gly Ala Cys
705 710 715 720

Ala Cys Gly Gly Gly Thr Thr Ala Cys Ala Thr Gly Gly Ala Cys Cys
725 730 735

Thr Gly Thr Ala Cys Cys Thr Gly Ala Thr Gly Ala Gly Cys Cys Gly
740 745 750

Cys Thr Gly Cys Cys Gly Cys Cys Ala Cys Ala Ala Thr Ala Thr Thr
755 760 765

Ala Thr Thr Gly Cys Cys Ala Ala Cys Thr Cys Thr Thr Cys Thr Thr
770 775 780

Thr Cys Ala Gly Cys Thr Gly Gly Thr Gly Gly Gly Gly Cys Gly Cys
785 790 795 800

Thr Thr Gly Gly Cys Thr Gly Ala Ala Thr Gly Ala Gly Ala Ala Cys
805 810 815

Cys Cys Thr Gly Ala Gly Ala Ala Ala Ala Ala Gly Thr Cys Ala
820 825 830

Thr Cys Gly Cys Ala Cys Cys Gly Gly Cys Thr Ala Ala Ala Thr Gly
835 840 845

Gly Cys Thr Gly Ala Ala Cys Gly Gly Thr Cys Gly Thr Gly Ala Gly
850 855 860

Thr Gly Cys Cys Gly Thr Gly Ala Thr Ala Thr Cys Thr Ala Thr Ala

37847-517 Sequence Listing_ST25 (003)

865 870 875 880

Cys Cys Gly Ala Ala Cys Gly Cys Ala Thr Gly Ala Thr Thr Cys Gly
885 890 895

Thr Cys Thr Gly Thr Ala Gly Cys Thr Cys Gly Ala Gly Thr Gly Ala
900 905 910

Cys Thr Gly Ala Cys Thr Gly
915

<210> 279

<211> 937

<212> PRT

<213> Artificial Sequence

<220>

<223> artifical construct

<400> 279

Cys Ala Gly Thr Cys Ala Gly Thr Cys Ala Gly Ala Ala Thr Thr Cys
1 5 10 15

Ala Ala Gly Ala Ala Gly Gly Ala Gly Ala Thr Ala Thr Ala Cys Ala
20 25 30

Thr Ala Thr Gly Ala Thr Cys Ala Thr Thr Gly Thr Thr Cys Gly Thr
35 40 45

Cys Thr Gly Ala Ala Ala Gly Gly Cys Gly Gly Thr Cys Thr Gly Gly
50 55 60

Gly Cys Ala Ala Cys Cys Ala Ala Cys Thr Gly Thr Cys Thr Cys Ala
65 70 75 80

Gly Thr Ala Thr Gly Cys Ala Cys Thr Gly Gly Gly Cys Cys Gly Thr
85 90 95

Ala Ala Gly Ala Thr Cys Gly Cys Gly Cys Ala Thr Cys Thr Gly Cys
100 105 110

37847-517 Sequence Listing_ST25 (003)

Ala Cys Ala Ala Thr Ala Cys Cys Gly Ala Ala Cys Thr Gly Ala Ala
115 120 125

Ala Cys Thr Gly Gly Ala Cys Ala Cys Cys Ala Cys Thr Thr Gly Gly
130 135 140

Thr Thr Cys Ala Cys Cys Ala Cys Thr Ala Thr Cys Thr Cys Cys Thr
145 150 155 160

Cys Cys Gly Ala Cys Ala Cys Thr Cys Cys Ala Cys Gly Thr Ala Cys
165 170 175

Cys Thr Ala Cys Cys Gly Thr Cys Thr Gly Ala Ala Cys Ala Ala Thr
180 185 190

Thr Ala Thr Ala Ala Cys Ala Thr Cys Ala Thr Cys Gly Gly Cys Ala
195 200 205

Cys Thr Ala Thr Thr Gly Cys Ala Thr Cys Cys Gly Cys Ala Ala Ala
210 215 220

Gly Gly Ala Ala Ala Thr Cys Cys Ala Gly Cys Thr Gly Ala Thr Cys
225 230 235 240

Gly Ala Ala Cys Gly Thr Gly Gly Thr Cys Gly Cys Gly Cys Gly Cys
245 250 255

Ala Ala Gly Gly Cys Cys Gly Thr Gly Gly Cys Thr Ala Cys Cys Thr
260 265 270

Gly Cys Thr Gly Thr Cys Thr Ala Ala Ala Ala Thr Thr Thr Cys Thr
275 280 285

Gly Ala Thr Cys Thr Gly Cys Thr Gly Ala Cys Thr Cys Cys Gly Ala
290 295 300

Thr Gly Thr Ala Cys Cys Gly Thr Cys Gly Thr Ala Cys Cys Thr Ala
305 310 315 320

37847-517 Sequence Listing_ST25 (003)

Cys Gly Thr Cys Cys Gly Thr Gly Ala Ala Cys Gly Thr Ala Thr Gly
325 330 335

Cys Ala Thr Ala Cys Cys Thr Thr Cys Gly Ala Thr Ala Ala Ala Gly
340 345 350

Cys Thr Ala Thr Cys Cys Thr Gly Ala Cys Cys Gly Thr Thr Cys Cys
355 360 365

Gly Gly Ala Cys Ala Ala Cys Gly Thr Gly Thr Ala Cys Cys Thr Gly
370 375 380

Gly Ala Thr Gly Gly Thr Thr Ala Cys Thr Gly Gly Cys Ala Gly Ala
385 390 395 400

Cys Cys Gly Ala Ala Ala Ala Gly Thr Ala Cys Thr Thr Cys Ala Ala
405 410 415

Ala Gly Ala Cys Ala Thr Cys Gly Ala Ala Gly Ala Ala Ala Thr Cys
420 425 430

Cys Thr Gly Cys Gly Cys Cys Gly Thr Gly Ala Gly Gly Thr Thr Ala
435 440 445

Cys Gly Cys Thr Gly Ala Ala Ala Gly Ala Thr Gly Ala Ala Cys Cys
450 455 460

Gly Gly Ala Thr Ala Gly Cys Ala Thr Cys Ala Ala Cys Cys Thr Gly
465 470 475 480

Gly Ala Ala Ala Thr Gly Gly Cys Thr Gly Ala Ala Cys Gly Thr Ala
485 490 495

Thr Thr Cys Ala Gly Gly Cys Thr Thr Gly Cys Cys Ala Cys Ala Gly
500 505 510

Cys Gly Thr Thr Thr Cys Cys Cys Thr Gly Cys Ala Cys Gly Thr Gly
515 520 525

37847-517 Sequence Listing_ST25 (003)

Cys Gly Thr Cys Gly Thr Gly Gly Cys Gly Ala Cys Thr Ala Cys Gly
530 535 540

Thr Thr Thr Cys Cys Ala Ala Cys Cys Cys Gly Ala Cys Cys Ala Cys
545 550 555 560

Thr Cys Ala Ala Cys Ala Ala Thr Thr Cys Cys Ala Cys Gly Gly Cys
565 570 575

Thr Gly Thr Thr Gly Cys Thr Cys Cys Ala Thr Thr Gly Ala Cys Thr
580 585 590

Ala Cys Thr Ala Cys Ala Ala Cys Cys Gly Cys Gly Cys Thr Ala Thr
595 600 605

Cys Thr Cys Thr Cys Thr Gly Ala Thr Thr Gly Ala Ala Gly Ala Ala
610 615 620

Ala Ala Ala Gly Thr Gly Gly Ala Thr Gly Ala Cys Cys Cys Gly Thr
625 630 635 640

Cys Thr Thr Thr Cys Thr Thr Thr Ala Thr Thr Thr Thr Thr Cys
645 650 655

Thr Gly Ala Cys Gly Ala Thr Cys Thr Gly Cys Cys Gly Thr Gly Gly
660 665 670

Gly Cys Thr Ala Ala Ala Gly Ala Ala Ala Ala Cys Cys Thr Gly Gly
675 680 685

Ala Cys Ala Thr Cys Cys Cys Thr Gly Gly Cys Gly Ala Ala Ala Ala
690 695 700

Ala Ala Cys Cys Thr Thr Cys Gly Thr Thr Gly Cys Gly Cys Ala Thr
705 710 715 720

Ala Ala Cys Gly Gly Cys Cys Cys Gly Gly Ala Ala Ala Ala Ala Gly
725 730 735

37847-517 Sequence Listing_ST25 (003)

Ala Gly Thr Ala Thr Thr Gly Cys Gly Ala Thr Cys Thr Gly Thr Gly
740 745 750

Gly Cys Thr Gly Ala Thr Gly Thr Cys Thr Cys Thr Gly Thr Gly Cys
755 760 765

Cys Ala Gly Cys Ala Cys Cys Ala Thr Ala Thr Cys Ala Thr Cys Gly
770 775 780

Cys Ala Ala Ala Cys Thr Cys Thr Thr Cys Thr Thr Thr Cys Ala Gly
785 790 795 800

Cys Thr Gly Gly Thr Gly Gly Gly Gly Thr Gly Cys Cys Thr Gly Gly
805 810 815

Cys Thr Gly Gly Gly Thr Cys Ala Ala Gly Ala Cys Gly Cys Cys Gly
820 825 830

Ala Ala Ala Ala Gly Ala Thr Gly Gly Thr Gly Ala Thr Cys Gly Cys
835 840 845

Gly Cys Cys Gly Cys Gly Thr Cys Gly Cys Thr Gly Gly Gly Cys Cys
850 855 860

Cys Thr Gly Thr Cys Cys Gly Ala Gly Ala Gly Cys Thr Thr Thr Gly
865 870 875 880

Ala Cys Ala Cys Thr Thr Cys Thr Gly Ala Cys Ala Thr Cys Ala Thr
885 890 895

Thr Cys Cys Gly Gly Ala Cys Thr Cys Thr Thr Gly Gly Ala Thr Thr
900 905 910

Ala Cys Thr Ala Thr Cys Thr Ala Gly Cys Thr Cys Gly Ala Gly Thr
915 920 925

Gly Ala Cys Thr Gly Ala Cys Thr Gly
930 935

37847-517 Sequence Listing_ST25 (003)

28 May 2018
2015259088

<210> 280
 <211> 913
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> artifical construct

<400> 280
 cagtcagtca gaattcaaga aggagatata catatggtac gcattgtgga aatcatcggc 60
 ggtctgggta accagatggt ccagtacgca ttctccctgt acctgaaaaa caaatctcac 120
 atctgggacc gtctgtatgt ggacatcgag gcgatgaaaa cctacgatcg tcactatggt 180
 ctggaactgg agaaagtttt caatctgagc ctgtgtccaa tctctaaccg tctgcaccgc 240
 aacctgcaaa aacgctcctt cgcaaaacac tttgtaaaga gcctgtacga gcactctgaa 300
 tgcgagttcg acgaaccggt gtaccgtggc ctgcgtcctt atcgctatta tcgcggctac 360
 tggcaaaacg aaggttactt cgttgatatt gaaccgatga tccgtgaggc ttttcagttc 420
 aacgttaaca tcctgagcaa aaagactaaa gcgatcgcat ccaaaatgcg ccgtgaactg 480
 tccgtatcta tccatgttcg ccgtgggtgat tacgaaaacc tgccggaagc gaaagcgatg 540
 catggcggtg tttgttctct ggactattac caaaaagcga tcgacttcat ccgccagcgt 600
 ctggataata acatctgttt ctatctgttc tccgacgata tcaattgggt agaagaaaac 660
 ctgcaactgg aaaaccgttg catcatcgac tggaaccagg gcgaagatag ctggcaggac 720
 atgtacctga tgagctgctg ccgccaccac attatcgcaa acagctcttt ctcttggtgg 780
 gcggcatggc tgaatccaaa caagaacaaa atcgtactga cccgaacaa atggttcaac 840
 catactgacg cagtgggtat cgtcccaaag tcctggatta aaattcctgt gttttagctc 900
 gagtgactga ctg 913

<210> 281
 <211> 919
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> artificial construct

<400> 281

37847-517 Sequence Listing_ST25 (003)

cagtcagtca gaattcaaga aggagatata catatgaaaa tcgttaaaat cctgggcggc	60
ctgggtaacc agatgtttca gtacgccctg ttcctgtctc tgaaagaacg cttcccgcat	120
gaacagggtga tgattgacac cagctgcttc cgcaattacc cactgcacaa cggtttcgaa	180
gtggatcgta tcttcgcccc gaaagcaccg gttgcctctt ggcgtaacat cctgaaggtt	240
gcctacccgt acccgaacta ccgtttcttg aaaatcggtg aatacatcct gcctaaacgt	300
aaaaccatgt gtgtagagcg taaaaacttc agctttgacg ccgcagtcct gacccgtaaa	360
ggcgattgct actatgatgg ctactggcag catgaggaat atttctgtga tatgaaagaa	420
acgatttggg aggctttctc cttccctgag ccggttgatg gtcgtaacaa ggagatcggg	480
gccctgctac aggcatctga tagcgcttcc ctgcacgttc gtcgcggtga ctacgtgaac	540
caccactgt ttcgtggtat ttgtgacctg gactattata aacgtgccat ccactacatg	600
gaagaacgcg tcaaccaca gctgtactgc gttttcagca acgatatggc ctggtgcgag	660
tcccacctgc gtgcactgct gccaggcaaa gaagtagttt atgttgactg gaacaagggt	720
gcggaatctt acgttgatat gcgtctgatg agcctgtgcc gtcacaacat catcgctaac	780
tcttctttca gctgggtgggg cgcatggctg aaccgtaacc cgcagaaagt ggtggtagcg	840
ccggaacgtt ggatgaacag cccgattgaa gaccagtgga gcgacaaatg gattaaactg	900
tagctcgagt gactgactg	919

<210> 282
 <211> 922
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> artificial construct

<400> 282

Cys	Ala	Gly	Thr	Cys	Ala	Gly	Thr	Cys	Ala	Gly	Ala	Ala	Thr	Thr	Cys
1				5				10						15	

Ala	Ala	Gly	Ala	Ala	Gly	Gly	Ala	Gly	Ala	Thr	Ala	Thr	Ala	Cys	Ala
			20					25					30		

Thr	Ala	Thr	Gly	Ala	Thr	Cys	Ala	Thr	Cys	Ala	Thr	Cys	Cys	Ala	Gly

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

35	40	45																	
Cys	Thr	Gly	Ala	Ala	Ala	Gly	Gly	Thr	Gly	Gly	Cys	Cys	Thr	Gly	Gly				
50						55					60								
Gly	Cys	Ala	Ala	Cys	Cys	Ala	Ala	Ala	Thr	Gly	Thr	Thr	Cys	Cys	Ala				
65					70					75					80				
Gly	Thr	Ala	Cys	Gly	Cys	Gly	Cys	Thr	Gly	Thr	Ala	Cys	Ala	Ala	Ala				
				85					90					95					
Thr	Cys	Cys	Cys	Thr	Gly	Ala	Ala	Ala	Ala	Ala	Ala	Cys	Gly	Thr	Gly				
			100					105					110						
Gly	Thr	Ala	Ala	Ala	Gly	Ala	Ala	Gly	Thr	Thr	Ala	Ala	Ala	Ala	Thr				
		115						120				125							
Thr	Gly	Ala	Thr	Gly	Ala	Cys	Ala	Ala	Ala	Ala	Cys	Thr	Gly	Gly	Cys				
	130					135					140								
Thr	Thr	Cys	Gly	Thr	Gly	Ala	Ala	Cys	Gly	Ala	Cys	Ala	Ala	Ala	Cys				
145					150					155					160				
Thr	Gly	Cys	Gly	Thr	Ala	Thr	Cys	Cys	Cys	Gly	Gly	Thr	Ala	Cys	Thr				
				165					170					175					
Gly	Thr	Cys	Cys	Cys	Gly	Thr	Thr	Gly	Gly	Gly	Gly	Thr	Gly	Thr	Thr				
			180					185					190						
Gly	Ala	Gly	Thr	Ala	Cys	Gly	Ala	Thr	Cys	Gly	Thr	Gly	Cys	Ala	Ala				
		195					200					205							
Cys	Cys	Gly	Ala	Cys	Gly	Ala	Ala	Gly	Ala	Gly	Ala	Thr	Thr	Ala	Thr				
	210					215					220								
Thr	Ala	Ala	Cys	Cys	Thr	Gly	Ala	Cys	Cys	Gly	Ala	Cys	Thr	Cys	Cys				
225					230					235					240				
Ala	Ala	Ala	Ala	Thr	Gly	Gly	Ala	Cys	Cys	Thr	Gly	Thr	Thr	Cys	Thr				

2015259088 28 May 2018

37847-517 Sequence Listing_ST25 (003)

245 250 255

Cys Thr Cys Gly Cys Ala Thr Cys Cys Gly Cys Cys Gly Thr Ala Ala
260 265 270

Ala Cys Thr Gly Ala Cys Thr Gly Gly Cys Cys Gly Cys Ala Ala Ala
275 280 285

Ala Cys Gly Thr Thr Cys Cys Gly Thr Ala Thr Cys Gly Ala Cys Gly
290 295 300

Ala Ala Gly Ala Ala Thr Cys Cys Gly Gly Thr Ala Ala Ala Thr Thr
305 310 315 320

Cys Ala Ala Cys Cys Cys Gly Gly Ala Ala Ala Thr Cys Cys Thr Gly
325 330 335

Gly Ala Ala Ala Ala Ala Gly Ala Gly Ala Ala Cys Gly Cys Thr Thr
340 345 350

Ala Thr Cys Thr Gly Gly Thr Gly Gly Gly Thr Thr Ala Thr Thr Gly
355 360 365

Gly Cys Ala Gly Thr Gly Cys Gly Ala Cys Ala Ala Gly Thr Ala Cys
370 375 380

Thr Thr Cys Gly Ala Cys Gly Ala Cys Ala Ala Ala Gly Ala Thr Gly
385 390 395 400

Thr Gly Gly Thr Thr Cys Gly Cys Gly Ala Ala Ala Thr Thr Cys Gly
405 410 415

Thr Gly Ala Ala Gly Cys Gly Thr Thr Cys Gly Ala Gly Ala Ala Ala
420 425 430

Ala Ala Ala Cys Cys Gly Cys Ala Gly Gly Ala Gly Cys Thr Gly Ala
435 440 445

Thr Gly Ala Cys Cys Gly Ala Cys Gly Cys Cys Ala Gly Cys Thr Cys

28 May 2018
2015259088

37847-517 Sequence Listing_ST25 (003)

450

455

460

Thr Thr Gly Gly Thr Cys Thr Ala Cys Thr Cys Thr Ala Cys Ala Gly
465 470 475 480

Cys Ala Gly Ala Thr Thr Gly Ala Ala Thr Gly Cys Thr Gly Cys Gly
485 490 495

Ala Gly Thr Cys Cys Gly Thr Ala Thr Cys Cys Cys Thr Gly Cys Ala
500 505 510

Cys Gly Thr Ala Cys Gly Thr Cys Gly Thr Ala Cys Thr Gly Ala Thr
515 520 525

Thr Ala Cys Gly Thr Gly Gly Ala Cys Gly Ala Gly Gly Ala Ala Cys
530 535 540

Ala Thr Ala Thr Thr Cys Ala Thr Ala Thr Cys Cys Ala Thr Ala Ala
545 550 555 560

Cys Ala Thr Cys Thr Gly Thr Ala Cys Gly Gly Ala Ala Ala Ala Ala
565 570 575

Thr Ala Cys Thr Ala Thr Ala Ala Ala Ala Cys Gly Cys Cys Ala
580 585 590

Thr Thr Gly Ala Thr Cys Gly Thr Gly Thr Gly Cys Gly Thr Ala Ala
595 600 605

Ala Cys Ala Gly Thr Ala Cys Cys Cys Gly Ala Gly Cys Gly Cys Ala
610 615 620

Gly Thr Gly Thr Thr Cys Thr Thr Cys Ala Thr Cys Thr Thr Cys Ala
625 630 635 640

Cys Cys Gly Ala Thr Gly Ala Thr Ala Ala Ala Gly Ala Ala Thr Gly
645 650 655

Gly Thr Gly Cys Cys Gly Cys Gly Ala Cys Cys Ala Cys Thr Thr Thr

37847-517 Sequence Listing_ST25 (003)

660

665

670

Ala Ala Ala Gly Gly Thr Cys Cys Gly Ala Ala Cys Thr Thr Cys Ala
675 680 685

Thr Cys Gly Thr Ala Gly Thr Cys Gly Ala Ala Cys Thr Gly Gly Ala
690 695 700

Ala Gly Ala Ala Gly Gly Cys Gly Ala Cys Gly Gly Thr Ala Cys Cys
705 710 715 720

Gly Ala Cys Ala Thr Cys Gly Cys Thr Gly Ala Ala Ala Thr Gly Ala
725 730 735

Cys Thr Cys Thr Gly Ala Thr Gly Thr Cys Cys Cys Gly Cys Thr Gly
740 745 750

Thr Ala Ala Ala Cys Ala Thr Cys Ala Cys Ala Thr Cys Ala Thr Cys
755 760 765

Gly Cys Thr Ala Ala Thr Thr Cys Thr Ala Gly Cys Thr Thr Thr Ala
770 775 780

Gly Cys Thr Gly Gly Thr Gly Gly Gly Cys Gly Gly Cys Gly Thr Gly
785 790 795 800

Gly Cys Thr Gly Ala Ala Cys Gly Ala Cys Thr Cys Cys Cys Cys Gly
805 810 815

Gly Ala Ala Ala Ala Ala Ala Thr Cys Gly Thr Gly Ala Thr Cys Gly
820 825 830

Cys Thr Cys Cys Thr Cys Ala Gly Ala Ala Ala Thr Gly Gly Ala Thr
835 840 845

Thr Ala Ala Cys Ala Ala Cys Cys Gly Cys Gly Ala Cys Ala Thr Gly
850 855 860

Gly Ala Cys Gly Ala Thr Ala Thr Thr Thr Ala Cys Ala Cys Cys Gly

37847-517 Sequence Listing_ST25 (003)

865 870 875 880

Ala Gly Cys Gly Thr Ala Thr Gly Ala Cys Thr Ala Ala Ala Ala Thr
885 890 895

Cys Gly Cys Ala Cys Thr Gly Thr Ala Gly Cys Thr Cys Gly Ala Gly
900 905 910

Thr Gly Ala Cys Thr Gly Ala Cys Thr Gly
915 920

<210> 283
<211> 898
<212> DNA
<213> Artificial Sequence

<220>
<223> artificial construct

<400> 283
cagtcagtca gaattcaaga aggagatata catatgcgcc tggttaaaat gatcggcgggt 60
ctgggtaatc agatgttcat ctacgcgttt tacctacaga tgcgtaagcg tttctccaac 120
gttcgtatcg acctgaccga tatgatgcac tacaacgtac actatggcta cgaactgcac 180
aaagttttcg gtctgcccg caccgagttc tgtatgaacc agcctctgaa aaaggttctg 240
gagttcctgt tcttccgtac cattgttgaa cgtaaacagc acggtcgtat ggagccgtat 300
acttgccagt atgtttggcc gctggtttac tttaagggtt tctatcagtc cgaacgttac 360
ttctccgaag ttaaggacga agttcgtgag tgtttcacct tcaatccggc actggcgaat 420
cgttcttccc aacagatgat ggaacagatc cagaatgatc ctcaggctgt ctctatccac 480
atccgtcgtg gcgactatct gaatccgaag cactacgaca ctatcggttg tatctgtcag 540
ctgccgtatt acaagcacgc cgtttccgaa attaaaaagt acgtttctaa ccctcacttt 600
tacgttttct ccgaagacct ggattgggtc aaagcaaacc tgccgctgga aaacgcacag 660
tacatcgatt ggaacaaagg cgcagatagc tggcaggata tgatgctgat gagctgttgc 720
aaacaccaca ttatctgtaa ctccaccttt agctgggtggg cggcgtggct gaacccatct 780
gtcgaaaaaa ccgtgatcat gccggaacag tggacgtctc gtcaagattc cgtggacttt 840

37847-517 Sequence Listing_ST25 (003)

gtggctagct gtggccgttg ggtccgtgtt aaaacggagt agctcgagtg actgactg 898

<210> 284
<211> 895
<212> DNA
<213> Artificial Sequence

<220>
<223> artificial construct

<400> 284
cagtcagtca gaattcaaga aggagatata catatgcgtc tgatcaagat gatcggcggc 60
ctgggtaacc agatgtttat ctacgcgttc tacctgaaaa tgaaacacca ttacccggat 120
acgaacatcg atctgtctga catggttcat tataaagttc acaacggtta tgagatgaac 180
cgtatctttg acctgagcca gactgaattt tgcatacaacc gtaccctgaa aaaaatcctg 240
gagttcctgt tcttcaaaaa aatctacgaa cgtcgccagg acccgtctac tctgtatcca 300
tacgaaaaac gttatTTTTG gccgctgctg tacttttaaag gtttctacca gtctgaacgc 360
ttcttcttcg atatcaaaga cgacgttcgt aaagccttct cttttaacct gaacatcgct 420
aaccggaaa gcctggaact gctgaaacag atcgaagttg acgaccaagc tgtttctatc 480
cacatccgcc gtggtgacta cctgctgccg cgtcactggg caaacacggg ttccgtgtgc 540
cagctgccgt attacaagaa cgcgatcgcg gaaatggaga accgtattac tggcccgagc 600
tactacgtgt tctctgatga tatctcttgg gttaaagaaa acatcccgct gaagaaagcg 660
gtctacgtga cgtggaacaa gggcgaagac agctggcagg atatgatgct gatgagccac 720
tgtcgtcacc acattatctg taattctacg ttctcctggt ggggtgcttg gctgaaccca 780
cgtaaagaga aaatcgatcat cgcgccgtgt cgctggttcc agcataaaga aaccccgga 840
atgtacccga aagaatggat caaagtaccg attaactagc tcgagtgact gactg 895

<210> 285
<211> 946
<212> PRT
<213> Artificial Sequence

<220>
<223> artificial construct

<400> 285

37847-517 Sequence Listing_ST25 (003)

Cys Ala Gly Thr Cys Ala Gly Thr Cys Ala Gly Ala Ala Thr Thr Cys
 1 5 10 15
 Ala Ala Gly Ala Ala Gly Gly Ala Gly Ala Thr Ala Thr Ala Cys Ala
 20 25 30
 Thr Ala Thr Gly Thr Ala Thr Thr Cys Thr Thr Gly Cys Cys Thr Gly
 35 40 45
 Thr Cys Thr Gly Gly Thr Gly Gly Cys Cys Thr Gly Gly Gly Thr Ala
 50 55 60
 Ala Cys Cys Ala Ala Ala Thr Gly Thr Thr Thr Cys Ala Ala Thr Ala
 65 70 75 80
 Cys Gly Cys Ala Gly Cys Ala Gly Cys Gly Thr Ala Thr Ala Thr Cys
 85 90 95
 Cys Thr Gly Ala Ala Gly Cys Ala Gly Thr Ala Thr Thr Thr Cys
 100 105 110
 Ala Gly Thr Cys Thr Ala Cys Cys Ala Cys Thr Cys Thr Gly Gly Thr
 115 120 125
 Cys Cys Thr Gly Gly Ala Thr Gly Ala Thr Ala Gly Cys Thr Ala Thr
 130 135 140
 Thr Ala Cys Thr Ala Thr Thr Cys Cys Cys Ala Gly Cys Cys Gly Ala
 145 150 155 160
 Ala Ala Cys Gly Thr Gly Ala Thr Ala Cys Cys Gly Thr Thr Cys Gly
 165 170 175
 Thr Ala Gly Cys Cys Thr Gly Gly Ala Ala Cys Thr Gly Ala Ala Thr
 180 185 190
 Cys Ala Gly Thr Thr Cys Ala Ala Cys Ala Thr Cys Thr Cys Thr Thr
 195 200 205

37847-517 Sequence Listing_ST25 (003)

Ala Thr Gly Ala Thr Cys Gly Thr Thr Thr Thr Ala Gly Cys Thr Thr
 210 215 220

Cys Gly Cys Gly Gly Ala Thr Gly Ala Ala Ala Ala Ala Gly Ala Gly
 225 230 235 240

Ala Ala Gly Ala Thr Cys Ala Ala Ala Cys Thr Gly Cys Thr Gly Cys
 245 250 255

Gly Cys Ala Ala Ala Thr Thr Cys Ala Ala Ala Cys Gly Thr Ala Ala
 260 265 270

Cys Cys Cys Gly Thr Thr Cys Cys Cys Thr Ala Ala Ala Cys Ala Gly
 275 280 285

Ala Thr Thr Thr Cys Cys Gly Ala Gly Ala Thr Cys Cys Thr Gly Thr
 290 295 300

Cys Thr Ala Thr Thr Gly Cys Gly Cys Thr Gly Thr Thr Cys Gly Gly
 305 310 315 320

Cys Ala Ala Ala Thr Ala Cys Gly Cys Gly Cys Thr Gly Thr Cys Cys
 325 330 335

Gly Ala Cys Cys Gly Thr Gly Cys Ala Thr Thr Thr Thr Ala Cys Ala
 340 345 350

Cys Cys Thr Thr Cys Gly Ala Ala Ala Cys Thr Ala Thr Cys Ala Ala
 355 360 365

Ala Ala Ala Cys Ala Thr Cys Gly Ala Cys Ala Ala Ala Gly Cys Gly
 370 375 380

Thr Gly Cys Cys Thr Gly Thr Thr Cys Thr Cys Cys Thr Thr Thr Thr
 385 390 395 400

Ala Cys Cys Ala Gly Gly Ala Cys Gly Cys Cys Gly Ala Thr Cys Thr
 405 410 415

37847-517 Sequence Listing_ST25 (003)

Gly Cys Thr Gly Ala Ala Thr Ala Ala Ala Thr Ala Thr Ala Ala Gly
420 425 430

Cys Ala Gly Cys Thr Gly Ala Thr Cys Cys Thr Gly Cys Cys Gly Cys
435 440 445

Thr Gly Thr Thr Cys Gly Ala Ala Cys Thr Gly Cys Gly Cys Gly Ala
450 455 460

Thr Gly Ala Cys Cys Thr Gly Cys Thr Gly Gly Ala Thr Ala Thr Cys
465 470 475 480

Thr Gly Cys Ala Ala Gly Ala Ala Cys Cys Thr Gly Gly Ala Ala Cys
485 490 495

Thr Gly Thr Ala Thr Thr Cys Cys Cys Thr Gly Ala Thr Cys Cys Ala
500 505 510

Ala Cys Gly Cys Ala Gly Cys Ala Ala Cys Ala Ala Thr Ala Cys Cys
515 520 525

Ala Cys Thr Gly Cys Ala Cys Thr Gly Cys Ala Thr Ala Thr Cys Cys
530 535 540

Gly Cys Cys Gly Thr Gly Gly Cys Gly Ala Cys Thr Ala Cys Gly Thr
545 550 555 560

Gly Ala Cys Cys Ala Ala Cys Cys Ala Gly Cys Ala Cys Gly Cys Cys
565 570 575

Gly Cys Gly Ala Ala Ala Thr Ala Cys Cys Ala Cys Gly Gly Cys Gly
580 585 590

Thr Gly Cys Thr Gly Gly Ala Cys Ala Thr Cys Ala Gly Cys Thr Ala
595 600 605

Cys Thr Ala Thr Ala Ala Cys Cys Ala Cys Gly Cys Ala Ala Thr Gly
610 615 620

37847-517 Sequence Listing_ST25 (003)

Gly Ala Ala Thr Ala Cys Gly Thr Gly Gly Ala Ala Cys Gly Thr Gly
625 630 635 640

Ala Ala Cys Gly Cys Gly Gly Cys Ala Ala Ala Cys Ala Gly Ala Ala
645 650 655

Cys Thr Thr Cys Ala Thr Thr Ala Thr Cys Thr Thr Cys Ala Gly Cys
660 665 670

Gly Ala Thr Gly Ala Thr Gly Thr Ala Cys Gly Thr Thr Gly Gly Gly
675 680 685

Cys Ala Cys Ala Gly Ala Ala Ala Gly Cys Gly Thr Thr Thr Cys Thr
690 695 700

Gly Gly Ala Gly Ala Ala Cys Gly Ala Thr Ala Ala Thr Thr Gly Cys
705 710 715 720

Thr Ala Cys Gly Thr Gly Ala Thr Thr Ala Ala Cys Ala Ala Cys Thr
725 730 735

Cys Cys Gly Ala Cys Thr Ala Cys Gly Ala Thr Thr Thr Cys Thr Cys
740 745 750

Thr Gly Cys Gly Ala Thr Cys Gly Ala Thr Ala Thr Gly Thr Ala Thr
755 760 765

Cys Thr Gly Ala Thr Gly Thr Cys Thr Cys Thr Gly Thr Gly Cys Ala
770 775 780

Ala Ala Ala Ala Cys Ala Ala Cys Ala Thr Cys Ala Thr Cys Gly Cys
785 790 795 800

Ala Ala Ala Thr Thr Cys Cys Ala Cys Cys Thr Ala Cys Thr Cys Cys
805 810 815

Thr Gly Gly Thr Gly Gly Gly Gly Thr Gly Cys Gly Thr Gly Gly Cys
820 825 830

37847-517 Sequence Listing_ST25 (003)

Thr Gly Ala Ala Cys Ala Ala Ala Thr Ala Cys Gly Ala Gly Gly Ala
835 840 845

Cys Ala Ala Ala Cys Thr Gly Gly Thr Thr Ala Thr Cys Thr Cys Thr
850 855 860

Cys Cys Gly Ala Ala Ala Cys Ala Ala Thr Gly Gly Thr Thr Thr Cys
865 870 875 880

Thr Gly Gly Gly Thr Ala Ala Cys Ala Ala Cys Gly Ala Ala Ala Cys
885 890 895

Cys Thr Cys Thr Cys Thr Gly Cys Gly Thr Ala Ala Cys Gly Cys Gly
900 905 910

Thr Cys Thr Thr Gly Gly Ala Thr Cys Ala Cys Cys Cys Thr Gly Thr
915 920 925

Ala Gly Cys Thr Cys Gly Ala Gly Thr Gly Ala Cys Thr Gly Ala Cys
930 935 940

Thr Gly
945

<210> 286
<211> 895
<212> DNA
<213> Artificial Sequence

<220>
<223> artificial construct

<400> 286
cagtcagtca gaattcaaga aggagatata catatgcgtc tgatcaagat gaccggtggc 60
ctgggtaacc agatgttcat ctacgcgttt tatctgcgta tgaaaaaacg ttatccgaaa 120
gttcgtattg atctgtctga tatggttcat tatcacgttc accacggcta tgaaatgcac 180
cgtgttttca atctgccgca caccgaattt tgcataacc agccgctgaa aaaagtgatc 240
gagttcctgt ttttcaaaaa gatttacgaa cgtaaacagg accctaattc tctgcgtgca 300

37847-517 Sequence Listing_ST25 (003)

```

ttcgagaaga agtatctgtg gccgctgctg tacttcaaag gtttctatca atctgagcgc      360
ttcttttgctg acatcaaaga cgaggttcgt aaagcattca cctttgactc ttctaaagtg      420
aacgctcgct ctgccgaact gctgcgtcgc ctggatgccg atgctaacgc ggtagcctg      480
cacattcgtc gcggtgacta tctacagccg cagcattggg ctaccactgg ttctgtctgc      540
cagctgccgt actaccagaa cgcgatcgct gaaatgaacc gtcgcgttgc tgccccgagc      600
tactacgttt tcagcgatga catcgcgtgg gtgaaggaaa acatccctct acagaacgca      660
gtgtacatcg actggaataa aggcgaagaa agctggcagg atatgatgct gatgagccac      720
tgccgccacc acattatctg taacagcacc ttctcttggg ggggcgcgtg gctggacccg      780
cacgaggaca aaattgtaat cgttccgaat cgttggttcc agcattgcga aactcctaac      840
atctatccgg caggctgggt gaaagttgcg attaattagc tcgagtgact gactg          895

```

```

<210> 287
<211> 6153
<212> DNA
<213> Artificial Sequence

```

```

<220>
<223> artificial construct

```

```

<400> 287
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca      60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg      120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc      180
accatatatg cggtgtgaaa taccgcacag atgcgtaagg agaaaatacc gcatcaggcg      240
cctcctcaac ctgtatattc gtaaaccacg cccaatggga gctgtctcag gtttgttcct      300
gattggttac ggcgcgtttc gcatcattgt tgagtttttc cgccagcccg acgcgcagtt      360
taccggtgcc tgggtgcagt acatcagcat ggggcaaatt ctttccatcc cgatgattgt      420
cgcggtgtg atcatgatgg tctgggcata tcgtcgcagc ccacagcaac acgtttcctg      480
aggaaccatg aaacagtatt tagaactgat gaaaaagtg ctcgacgaag gcacacagaa      540
aaacgaccgt accggaaccg gaacgctttc ctttttgggt catcagatgc gttttaacct      600
gcaagatgga ttcccgtgg tgacaactaa acgttgccac ctgcgttcca tcatccatga      660

```

37847-517 Sequence Listing_ST25 (003)

actgctgtgg	tttctgcagg	gcgacactaa	cattgcttat	ctacacgaaa	acaatgtcac	720
catctgggac	gaatgggccg	atgaaaacgg	cgacctcggg	ccagtgtatg	gtaaacagtg	780
gcgcgcctgg	ccaacgccag	atggtcgtca	tattgaccag	atcactacgg	tactgaacca	840
gctgaaaaac	gacccggatt	cgcgccgcat	tattgtttca	gcgtggaacg	taggcgaact	900
ggataaaatg	gcgctggcac	cgtgccatgc	attcttccag	ttctatgtgg	cagacggcaa	960
actctcttgc	cagctttatc	agcgctcctg	tgacgtcttc	ctcggcctgc	cgttcaacat	1020
tgccagctac	gcgttattgg	tgcatatgat	ggcgcagcag	tgcgatctgg	aagtgggtga	1080
ttttgtctgg	accggtggcg	acacgcatct	gtacagcaac	catatggatc	aaactcatct	1140
gcaattaagc	cgcgaaccgc	gtccgctgcc	gaagttgatt	atcaaacgta	aacccgaatc	1200
catcttcgac	taccgtttcg	aagactttga	gattgaaggc	tacgatccgc	atccgggcat	1260
taaagcgccg	gtggctatct	aattacgaaa	catcctgcca	gagccgacgc	cagtgtgcgt	1320
cggttttttt	accctccgtt	aaattcttcg	agacgccttc	ccgaaggcgc	cattcgccat	1380
tcaggctgcg	caactgttgg	gaagggcgat	cgggtgcgggc	ctcttcgcta	ttacgccagc	1440
tggcgaaagg	gggatgtgct	gcaaggcgat	taagttgggt	aacgccaggg	ttttcccagt	1500
cacgacgttg	taaaacgacg	gccagtgcc	agctttcttt	aatgaagcag	ggcatcagga	1560
cggtatcttt	gtggagaaag	cagagtaatc	ttattcagcc	tgactggtgg	gaaaccacca	1620
gtcagaatgt	gttagcgcat	gttgacaaaa	ataccattag	tcacattatc	cgtcagtcgg	1680
acgacatggg	agataacctg	tttattatgc	gttttgatct	tacgtttaat	attaccttta	1740
tgcgatgaaa	cggctcttggc	tttgatattc	atgttggtcag	agatttgaat	ggttccctga	1800
cctgccatcc	acattcgcaa	catactcgat	tcggttcggc	tcaatgataa	cgtcggcata	1860
tttaaaaacg	aggttatcgt	tgtctctttt	ttcagaatat	cgccaaggat	atcgctgaga	1920
gattccgggt	taatcgattt	agaactgatc	aataaatttt	ttctgaccaa	tagatattca	1980
tcaaaatgaa	cattggcaat	tgccataaaa	acgataaata	acgtattggg	atgttgatta	2040
atgatgagct	tgatacgctg	actgttagaa	gcacgttgga	tgaaacagtc	ctcattaata	2100
aacaccactg	aagggcgctg	tgaatcacia	gctatggcaa	ggatcatcaac	ggtttcaatg	2160
tcgttgattt	ctcttttttt	aaccctcta	ctcaacagat	accgggttaa	acctagtcgg	2220

37847-517 Sequence Listing_ST25 (003)

gtgtaactac ataaatccat aataatcggt gacatggcat accctcactc aatgcgtaac	2280
gataattccc cttacctgaa tatttcatca tgactaaacg gaacaacatg gggtcacctaa	2340
tgcgccactc tcgcgatttt tcaggcggac ttactatccc gtaaagtgtt gtataatttg	2400
cctggaattg tcttaaagta aagtaaagtgt tgcgatatgt gaggtagctt aaaacaaata	2460
tttcgctgca ggagtatcct ggaagatgtt cgtagaagct tactgctcac aagaaaaaag	2520
gcacgtcatc tgacgtgcct tttttatttg tactaccctg tacgattact gcagctcgag	2580
ctagttaatc ggtactttga tccattcttt cgggtacatg tccgggggtt ctttatgctg	2640
gaaccagcga cacggcgcga tgacgatttt ctctttacgt gggttcagcc aagcacccca	2700
ccaggagaac gtagaattac agataatgtg gtgacgacag tggctcatca gcatcatatc	2760
ctgccagctg tcttcgccct tgttccacgt cacgtagacc gctttcttca gcgggatgtt	2820
ttctttaacc caagagatat catcagagaa cacgtagtag ctcgggccag taatacgggt	2880
ctccatttcc gcgatcgcgt tcttgtaata cggcagctgg cacacggaac ccgtgtttgc	2940
ccagtgcgcg ggcagcaggt agtcaccacg gcggatgtgg atagaaacag cttggtcgtc	3000
aacttcgatc tgtttcagca gttccaggct ttccgggtta gcgatgttca ggttaaaaga	3060
gaaggcttta cgaacgtcgt ctttgatatc gaagaagaag cgttcagact ggtagaaacc	3120
tttaaagtac agcagcggcc aaaaataacg tttttcgtat ggatacagag tagacgggtc	3180
ctggcgacgt tcgtagattt ttttgaagaa caggaaactcc aggatttttt tcagggtacg	3240
gttgatgcaa aattcagtct ggctcaggtc aaagatacgg ttcatctcat aaccgttgtg	3300
aactttataa tgaaccatgt cagacagatc gatgttcgta tccgggtaat ggtgtttcat	3360
tttcaggtag aacgcgtaga taaacatctg gttaccagg cgcgcgatca tcttgatcag	3420
acgcatatgt atatctcctt cttgaattct aaaaattgat tgaatgtatg caaataaatg	3480
catacaccat aggtgtgggt taatttgatg ccctttttca gggctggaat gtgtaagagc	3540
ggggttatth atgctgttgt ttttttgta ctcgggaagg gctttacctc ttccgcataa	3600
acgcttccat cagcgtttat agttaaaaa atctttcggga actggttttg cgcttacccc	3660
aaccaacagg ggatttgctg ctttccattg agcctgtttc tctgcgcgac gttcgcggcg	3720
gcgtgtttgt gcatccatct ggattctcct gtcagttagc tttggtggtg tgtggcagtt	3780

37847-517 Sequence Listing_ST25 (003)

gtagtcctga	acgaaaaccc	cccgcgattg	gcacattggc	agctaataccg	gaatcgcact	3840
tacggccaat	gcttcgtttc	gtatcacaca	ccccaaagcc	ttctgctttg	aatgctgccc	3900
ttcttcaggg	cttaattttt	aagagcgtca	ccttcattgg	ggtcagtgcg	tcctgctgat	3960
gtgctcagta	tcaccgccag	tggtatttat	gtcaacaccg	ccagagataa	tttatcaccg	4020
cagatggtta	tctgtatggt	ttttatatga	atttatTTTT	tgcagggggg	cattgttttg	4080
taggtgagag	atcaattctg	cattaatgaa	tcggccaacg	cgcggggaga	ggcggtttgc	4140
gtattgggcg	ctcttccgct	tcctcgctca	ctgactcgct	gcgctcggtc	gttcggctgc	4200
ggcgagcggg	atcagctcac	tcaaaggcgg	taatacgggt	atccacagaa	tcaggggata	4260
acgcaggaaa	gaacatgtga	gcaaaaggcc	agcaaaaggc	caggaaccgt	aaaaaggccg	4320
cgttgctggc	gtttttccat	aggctccgcc	cccctgacga	gcatcacaaa	aatcgacgct	4380
caagtcagag	gtggcgaaac	ccgacaggac	tataaagata	ccaggcgttt	ccccctggaa	4440
gctccctcgt	gcgctctcct	gttccgacct	tgccgcttac	cggataacct	tccgcctttc	4500
tcccttcggg	aagcgtggcg	ctttctcata	gctcacgctg	taggtatctc	agttcggtgt	4560
aggctgttcg	ctccaagctg	ggctgtgtgc	acgaaccccc	cgttcagccc	gaccgctgcg	4620
ccttatccgg	taactatcgt	cttgagtcca	acccggttaag	acacgactta	tcgccactgg	4680
cagcagccac	tggtaacagg	attagcagag	cgaggatatgt	aggcggtgct	acagagttct	4740
tgaagtgggtg	gcctaactac	ggctacacta	gaagaacagt	atttggtatc	tgcgctctgc	4800
tgaagccagt	taccttcgga	aaaagagttg	gtagctcttg	atccggcaaa	caaaccaccg	4860
ctggtagcgg	tggttttttt	gtttgcaagc	agcagattac	gcgcagaaaa	aaaggatctc	4920
aagaagatcc	tttgatcttt	tctacggggg	ctgacgtctc	gtggaacgaa	aactcacgtt	4980
aagggatttt	ggcatagaga	ttatcaaaaa	ggatcttcac	ctagatcctt	ttaaattaaa	5040
aatgaagttt	taaatcaatc	taaagtatat	atgagtaaac	ttggtctgac	agttaccaat	5100
gcttaatcag	tgaggcacct	atctcagcga	tctgtctatt	tcgttcatcc	atagttgcct	5160
gactccccgt	cgtgtagata	actacgatac	gggaggggct	accatctggc	cccagtgctg	5220
caatgatacc	gcgagacca	cgctcaccgg	ctccagattt	atcagcaata	aaccagccag	5280
ccggaagggc	cgagcgcaga	agtggctcct	caactttatc	cgctccatc	cagtctatta	5340

37847-517 Sequence Listing_ST25 (003)

attgttgccg ggaagctaga gtaagtagtt cgccagttaa tagtttgcg aacgttggtt	5400
ccattgctac aggcatcgtg gtgtcacgct cgtcgtttgg tatggcttca ttcagctccg	5460
gttcccaacg atcaaggcga gttacatgat ccccatgtt gtgcaaaaaa gcggttagct	5520
ccttcggtcc tccgatcgtt gtcagaagta agttggccgc agtgttatca ctcatggtta	5580
tggcagcact gcataattct cttactgtca tgccatccgt aagatgcttt tctgtgactg	5640
gtgagtactc aaccaagtca ttctgagaat agtgtagcg gcgaccgagt tgctcttgcc	5700
cggcgtcaat acgggataat accgcgccac atagcagaac tttaaaagtg ctcatcattg	5760
gaaaacgttc ttcggggcga aaactctcaa ggatcttacc gctgttgaga tccagttcga	5820
tgtaacccac tcgtgcaccc aactgatctt cagcatcttt tactttcacc agcgtttctg	5880
ggtgagcaaa aacaggaagg caaaatgccg caaaaaaggg aataagggcg acacggaaat	5940
gttgaatact catactcttc ctttttcaat attattgaag catttatcag ggttattgtc	6000
tcattgagcgg atacatattt gaatgtattt agaaaaataa acaaataagg gttccgcgca	6060
catttccccg aaaagtgcc cctgacgtct aagaaacct tattatcatg acattaacct	6120
ataaaaaatag gcgtatcacg aggccctttc gtc	6153

<210> 288
 <211> 1583
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> artificial construct

<400> 288	
caccatcgaa tggcgcaaaa cttttcgcgg tatggcatga tagcgcccgg aagagagtca	60
agtgtaggct ggagctgctt cgaagttcct atactttcta gagaatagga acttcggaat	120
aggaacttcg gaataggaac taaggaggat attcatatgt actatttaaa aaacacaaac	180
ttttggatgt tcggtttatt ctttttcttt tactttttta tcatgggagc ctacttccc	240
tttttcccga tttggctaca tgacatcaac catatcagca aaagtgatac gggtattatt	300
tttgccgcta tttctctgtt ctcgctatta ttccaaccgc tgtttggtct gctttctgac	360
aaactcgggc tgcgcaaata cctgctgtgg attattaccg gcatgttagt gatgtttgcg	420

37847-517 Sequence Listing_ST25 (003)

ccgttcttta tttttatctt cgggccactg ttacaataca acatttttagt aggatcgatt	480
gttggtggta tttatctagg cttttgtttt aacgccggtg cgccagcagt agaggcattt	540
attgagaaag tcagccgtcg cagtaatttc gaatttggtc gcgcgcggat gtttggctgt	600
gttggctggg cgctgtgtgc ctcgattgtc ggcatcatgt tcaccatcaa taatcagttt	660
gttttctggc tgggctctgg ctgtgcactc atcctcgccg ttttactctt tttcgccaaa	720
acggatgcgc cctcttctgc cacggttgcc aatgcggtag gtgccaacca ttcggcattt	780
agccttaagc tggcactgga actgttcaga cagccaaaac tgtgggtttt gtcactgtat	840
gttattggcg tttcctgcac ctacgatgtt tttgaccaac agtttgctaa tttctttact	900
tcgttctttg ctaccggtga acagggtacg cgggtatttg gctacgtaac gacaatgggc	960
gaattactta acgcctcgat tatgttcttt gcgccactga tcattaatcg catcgggtggg	1020
aaaaacgccc tgctgctggc tggcactatt atgtctgtac gtattattgg ctcatcgttc	1080
gccacctcag cgctggaagt ggttattctg aaaacgctgc atatgtttga agtaccgttc	1140
ctgctgggtg gctgctttaa atatattacc agccagtttg aagtgcgttt ttcagcgacg	1200
atttatctgg tctgtttctg cttctttaag caactggcga tgatttttat gtctgtactg	1260
gcgggcaata tgtatgaaag catcggtttc cagggcgctt atctggtgct gggctctggtg	1320
gcgctgggct tcaccttaat ttccgtgttc acgcttagcg gccccggccc gctttccctg	1380
ctgcgtcgtc aggtgaatga agtcgcttaa gcaatcaatg tcggatgcgg cgcgagcgcc	1440
ttatccgacc aacatatcat aacggagtga tcgcattgta aattataaaa attgcctgat	1500
acgctgcgct tatcaggcct acaagttcag cgatctacat tagccgcac cggcatgaac	1560
aaagcgcagg aacaagcgtc gca	1583

<210> 289
 <211> 2070
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> artificial construct

<400> 289	
gttcggttat atcaatgtca aaaacctcac gccgtcaag ctggtgatca actccgggaa	60

37847-517 Sequence Listing_ST25 (003)

cggcgagcg ggtccggtgg tggacgccat tgaagcccgc tttaaagccc tcggcgcgcc	120
cgtggaatta atcaaagtgc acaacacgcc ggacggcaat ttccccaacg gtattcctaa	180
cccactactg ccggaatgcc gcgacgacac ccgcaatgcg gtcatacaaac acggcgcgga	240
tatgggcatt gcttttgatg gcgattttga ccgctgtttc ctgtttgacg aaaaagggca	300
gtttattgag ggctactaca ttgtcggcct gttggcagaa gcattcctcg aaaaaatcc	360
cggcggaag atcatccacg atccacgtct ctctggaac accgttgatg tggtgactgc	420
cgcaggtggc acgccggtaa tgtcgaaaac cggacacgcc tttattaaag aacgtatgcg	480
caaggaagac gccatctatg gtggcgaaat gagcgccac cattacttcc gtgatttcgc	540
ttactgcgac agcggcatga tcccgtggct gctggtcgcc gaactgggtgt gcctgaaaga	600
taaaacgctg ggcgaactgg tacgcgaccg gatggcggcg tttccggcaa gcggtgagat	660
caacagcaaa ctggcgcaac ccgttgaggc gattaaccgc gtggaacagc attttagccg	720
tgaggcgctg gcggtggatc gcaccgatgg catcagcatg acctttgccg actggcgctt	780
taacctgcgc acctccaata ccgaaccggt ggtgcgcctg aatgtggaat cgcgcggtga	840
tgtgccgctg atggaagcgc gaacgcgaac tctgctgacg ttgctgaacg agtaatgtcg	900
gatcttcctt taccctactg cgggtaaggg gctaataaca ggaacaacga tgattccggg	960
gatccgtcga cctgcagttc gaagttccta ttctctagaa agtataggaa cttcgaagca	1020
gctccagcct acagttaaca aagcggcata ttgatatgag cttacgtgaa aaaaccatca	1080
gcggcgcgaa gtggtcggcg attgccacgg tgatcatcat cggcctcggg ctggtgcaga	1140
tgaccgtgct ggcgcggtt atcgacaacc accagttcgg cctgcttacc gtgtcgctgg	1200
tgattatcgc gctggcagat acgctttctg acttcggtat cgctaactcg attattcagc	1260
gaaaagaaat cagtcacctt gaactcacca cgttgtactg gctgaacgtc gggctgggga	1320
tcgtgggtgtg cgtggcggtg tttttgttga gtgatctcat cggcgacgtg ctgaataacc	1380
cggacctggc accgttgatt aaaacattat cgctggcggt tgtggtaatc cccacgggc	1440
aacagttccg cgcgttgatg caaaaagagc tggagttcaa caaatcggc atgatacga	1500
ccagcgcggt gctggcgggc ttcacttgta cgggtggttag cgccatttc tggccgctgg	1560
cgatgaccgc gatcctcggg tatctgggtca atagtgcggg gagaacgtg ctgtttggct	1620

37847-517 Sequence Listing_ST25 (003)

acttttgccg caaaatttat cgccccggtc tgcattttctc gctggcgtcg gtggcaccga	1680
acttacgctt tggcgcctgg ctgacggcgg acagcatcat caactatctc aataccaacc	1740
tttcaacgct cgtgctggcg cgtatttctc gcgcgggcgt ggccaggggga tacaacctgg	1800
cgtacaacgt ggccgttgtg ccaccgatga agctgaaccc aatcatcacc cgcgtgttgt	1860
ttccggcatt cgccaaaatt caggacgata ccgaaaagct gcgtgttaac ttctacaagc	1920
tgctgtcggg agtgggggatt atcaactttc cggcgtctgt cgggctaata gtggtgtcga	1980
ataactttgt accgctggtc tttggtgaga agtgaacag cattattccg gtgctgcaat	2040
tgctgtgtgt ggtgggtctg ctgcgctccg	2070

<210> 290
<211> 5046
<212> DNA
<213> Artificial Sequence

<220>
<223> artificial construct

<400> 290	
gtggatggaa gaggtggaaa aagtggttat ggaggagtgg gtaattgatg gtgaaaggaa	60
agggttggtg atttatggga agggggaagg ggaagaggga tgtggtgaat aattaaggat	120
tgggataгаа ttagttaagg aaaaaggggg gattttatgt ggggtttaat ttttggtgta	180
ttgtgggggt tgaatgtggg ggaaagatgg ggatatagtg aggtagatgt taatagatgg	240
ggtgaaggag agtgggtgtg tgtgattagg tgggggaaat taaagtaaga gagagggtga	300
tgattggggg gatgggtgga ggtggagtgt gaagttggtg ttgtgtagaa agtataggaa	360
gttgagaggg gttttgaagg tgagggtggg ggaaggagtg aggggggaag ggggtggtaaa	420
ggaaggggaa gaggtagaaa gggagtgggg agaaagggtg gtgagggggg atgaatgtga	480
ggtagtgggg tatgtggaga agggaaaagg gaaggggaaa gagaaaggag gtaggttgga	540
gtggggttag atggggatag gtagagtggg gggttttatg gagaggaagg gaaggggaat	600
tgggaggtgg ggggggggtgt ggtaagggtt ggaaggggtg gaaagtaaag tggatgggtt	660
tgttgggggg aaggatgtga tgggggaggg gatgaagatg tgatgaagag agaggatgag	720
gatggtttgg gatgattgaa gaagatggat tggagggagg ttgtgggggg ggttgggtgg	780

37847-517 Sequence Listing_ST25 (003)

agaggggtatt	gggggtatgag	tgggggagaag	agagaatggg	gtgggtgtgat	gggggggtgt	840
tgggggtgtg	aggggagggg	gggggggttg	tttttgtgaa	gagggaggtg	tgggggtggg	900
tgaatgaagt	ggaggaggag	ggaggggggg	tatgggtgggt	ggggaggagg	ggggtttggtt	960
ggggaggtgt	ggtggagggtt	gtgagtgaag	ggggaaggga	gtgggtggta	ttgggggaag	1020
tgggggggga	ggatgtgggt	tgatgtgagg	ttgggtgggtg	ggagaaagta	tggatgatgg	1080
gtgatggaat	gggggggggtg	gatagggttg	atgggggtag	gtggggattg	gaggaggaag	1140
ggaaagatgg	gatggaggga	ggaggtagtg	ggatggaagg	gggtgtttgtg	gatgaggatg	1200
atgtggagga	agaggatgag	ggggtggggg	gagggggaagt	gttggggagg	gtgaaggggg	1260
gatgggggag	ggggaggatg	tgggtggtgag	ggatggggat	gggtggttgg	ggaatatgat	1320
ggtggaaaat	gggggggtttt	gtggattgat	ggagtgtggg	ggggtgggtg	tgggggaggg	1380
gtatgaggag	atagggttgg	gtaggggtga	tattggtgaa	gaggttgggg	gggaatgggg	1440
tgaggggttg	gtgggtggtt	agggtatggg	gggtggggat	tgggagggga	tggggttgta	1500
tggggttgtt	gaggagtgtg	tgtaataagg	ggatgttgaa	gttggtattg	ggaagtgtgt	1560
attgtgtaga	aagtatagga	agttggaagg	aggtggaggg	tagataaagg	ggggggttat	1620
ttttgagagg	agaggaagtg	gtaatggtag	ggaggggggg	tgaggtggaa	ttggggggat	1680
agtgaggggg	tggaggagtg	gtggggagga	atggggatat	ggaaagggtg	gatattgagg	1740
gatgtgggtt	gttgggggtg	gaggagatgg	ggatgggtgg	tttggatgag	ttggtgttga	1800
gtgtaggggg	tgatgttgaa	gtggaagtgg	gggggggagt	ggtgtggggg	ataattgaat	1860
tgggggggtg	gggaggggag	agggttttgg	gtggggaaga	ggtagggggt	atagatgttg	1920
agaatgggag	atgggagggg	tgaaaagagg	ggggagtaag	ggggtgggga	tagttttgtt	1980
gggggggtaa	tgggagggag	tttagggggt	gtggtaggtg	ggggaggtgg	gagttgaggg	2040
gaatgggggg	gggatggggt	gtatgggtgg	ggagtgaag	atgaagggtg	atggggattt	2100
gaggagtagg	atgaatgggg	taggttttgg	gggtgataaa	taaggttttg	gggtgatggt	2160
gggaggggtg	aggggtggta	atgaggaggg	gatgaggaag	tgtatgtggg	gtggagtgga	2220
agaagggtgg	ttgggggtgg	taatgggggg	gggggttgga	gggttgagg	gaggggttag	2280
ggtgaatggg	ggtgggttga	gttaggggaa	tgtggttatg	gaggggtgga	ggggtgaagt	2340

37847-517 Sequence Listing_ST25 (003)

gatgggggag gggggtgagg agttgttttt tatggggaat ggagatgtgt gaaagaaagg	2400
gtgagtgggg gttaaattgg gaagggttat tagggaggtg gatggaaaaa tggatttggg	2460
tggtggtgag atgggggatg ggggtggagg ggggggggag ggtgagagtg aggttttggg	2520
ggagagggga gtggtgggag ggggtgatgt ggggggggtt tgaggatggg gtggggttgg	2580
gttggagtag gggtagtgtg agggagagt tt ggggggggggt gtgggggtgg ggtagttgag	2640
ggagttgaat gaagtgttta ggttgtggag ggagatggag agggagttga ggggttggga	2700
gggggttagg atggaggggg aggatggagt ggaggaggtg gttatgggta tgagggaaga	2760
ggtattgggt ggtgagttgg atggtttggg gggataaagg gaagtggaaa aagtgggtgt	2820
ggtgttttgg ttgggtgagg ggtggatggg ggggtggggtg gggaaagagg agagggttga	2880
tagagaagtg gggatgggtt ggggtatggg gaaaatgagg ggggtaaggg gaggaggggt	2940
tggggttttg atgatattta atgagggagt gatggaggga gtgggagagg aaggggggggt	3000
gtaaaggggg atagtgagga aagggttggg agtatttagg gaaaggggga agagtgttag	3060
ggatgggggtg ggggtatttg gaaaggatga ggggggggggt gtgtggaggt agggaaaggg	3120
attttttgat ggaggatttg gggagagggg ggaagggtg gtgttgatgg aggggggggt	3180
agatggggga aataatatgg gtgggggttg tgtgggttg ggggggttga tagtggaggg	3240
ggggggaagg atggagagat ttgatggagg gatagagggg gtggtgatta ggggggttgg	3300
gtgattgatt ggggagggag gagatgatga gagtggggtg attaggatgg ggggtggagga	3360
ttggggtttag gggttgggtg atggggggta gggagggggg atgatgggtg agaggattga	3420
ttgggaggat ggggtgggtt tgaatattgg gttgatggag gagatagagg gggtaggggt	3480
gggagaggggt gtaggagagg ggatggttgg gataatggga agaggggagg gggttaaagt	3540
tgttgtggtt gatgaggagg atatggtgga ggatggtgtg gtgatggatg aggtgaggat	3600
ggagaggatg atggtggtga gggttaaggg gtggaatgag gaagggttg gggttgagga	3660
ggaggagagg attttgaatg gggaggtggg ggaaagggag atgggagggt tgtggttgaa	3720
tgagggtggg gtgggggggt tggagttgaa ggaggggagg atagagattg gggatttggg	3780
gggtggagag tttggggttt tggaggttga gaggtagtgt gaggggatgg ggataaggag	3840
gagggtgatg gataatttga ggggggaaag ggggggtggg ggtggggagg tgggtttgag	3900

37847-517 Sequence Listing_ST25 (003)

ggtgggataa agaaagtgtt aggggtaggt agtgaggga gtggggggag atgtgaagtt	3960
gagggtggag tagagggggg gtgaaatgat gattaaagg agtggaaga tggaaatggg	4020
tgatttgtgt agtgggttta tggaggaagg agaggtgagg gaaaatgggg gtgatggggg	4080
agatatggtg atgttgaga taagtggggt gagtggaggg gaggaggatg agggggaggg	4140
ggttttgtgg ggggggtaaa aatggggtga ggtgaaattg agaggggaaa ggagtgtggt	4200
gggggtaagg gaggaggggg gggttggagg agagatgaaa gggggagtta aggggatgaa	4260
aaataattgg ggtgtggggt tgggttaggg aggtttgatg aagattaaat gtgagggagt	4320
aagaaggggt gggattgtgg gtgggaagaa aggggggatt gagggtaatg ggataggtga	4380
ggttggtgta gatgggggga tggtaagggt ggatgtggga gtttgagggg aggaggagag	4440
tatgggggtg aggaagatgg gagggaggga gttttggggg aggggttgtg gtgggggaaa	4500
ggagggaaaag ggggattggg gattgagggt ggggaagtgt tgggaagggg gatgggtggg	4560
ggggtgttgg gtattagggg aggtggggaa agggggatgt ggtggaaggg gattaagttg	4620
ggtaagggga gggttttggg agtgaggagg ttgtaaaagg agggggagtg aatgggtaat	4680
gatggtgata gtaggtttgg tgaggttgtg agtggaatat agtgagggtg gggaaaatgg	4740
agtaataaaa agaggggtgg gagggtaatt gggggttggg agggtttttt tgtgtgggta	4800
agttagatgg gggatggggg ttggggttat taagggtgtg tgtaagggga tgggtggggt	4860
gatataagtg gtgggggttg gtaggttgaa ggattgaagt gggatataaa ttataaagag	4920
gaagagaaga gtgaataaat gtgaattgat ggagaagatt ggtggagggg gtgatatgtg	4980
taaagggtggg ggtgggggtg ggtagatgg tattattggt tgggtaagtg aatgtgtgaa	5040
agaagg	5046

<210> 291
 <211> 6155
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> artificial construct

<400> 291	
tctagaattc taaaaattga ttgaatgtat gcaaataaat gcatacacca taggtgtggt	60

37847-517 Sequence Listing_ST25 (003)

ttaatttgat gccctttttc agggctggaa tgtgtaagag cggggttatt tatgctgttg	120
tttttttggt actcgggaag ggctttacct cttccgcata aacgcttcca tcagcgttta	180
tagttaaaaa aatctttcgg aactggtttt gcgcttacct caaccaacag gggatttgct	240
gctttccatt gagcctgttt ctctgcgcga cgttcgcggc ggctgttttg tgcattccatc	300
tggattctcc tgtcagttag ctttgggtgg gtgtggcagt ttagtcctg aacgaaaacc	360
ccccgcgatt ggcacattgg cagctaattc ggaatcgcac ttacggccaa tgcttcgttt	420
cgtatcacac accccaaagc cttctgcttt gaatgctgcc cttcttcagg gcttaatttt	480
taagagcgtc accttcatgg tggtcagtgc gtcctgctga tgtgctcagt atcaccgcca	540
gtgggtattta tgtcaacacc gccagagata atttatcacc gcagatgggt atctgtatgt	600
tttttatatg aatttatatt ttgcaggggg gcattgtttg gtaggtgaga gatcaattct	660
gcattaatga atcggccaac gcgcggggag aggcggtttg cgtattgggc gctcttcgc	720
ttcctcgctc actgactcgc tgcgctcggc cggttcggctg cggcgagcgg tatcagctca	780
ctcaaaggcg gtaatacggc tatccacaga atcaggggat aacgcaggaa agaacatgtg	840
agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca	900
taggctccgc cccctgacg agcatcaca aaatcgacgc tcaagtcaga ggtggcgaaa	960
cccgacagga ctataaagat accaggcgtt tccccctgga agctccctcg tgcgctctcc	1020
tgttccgacc ctgccgctta ccggatacct gtccgccttt ctcccttcgg gaagcgtggc	1080
gctttctcat agctcacgct gtaggtatct cagttcgggt taggtcggtc gctccaagct	1140
gggctgtgtg cacgaacccc ccgttcagcc cgaccgtgc gccttatccg gtaactatcg	1200
tcttgagtcc aaccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag	1260
gattagcaga gcgaggtatg taggcgggtg tacagagttc ttgaagtggg ggcctaacta	1320
cggctacact agaaggacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg	1380
aaaaagagtt gtagctctt gatccggcaa acaaaccacc gctggtagcg gtggtttttt	1440
tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgatctt	1500
ttctacgggg tctgacgctc agtggaacga aaactcacgt taagggattt tggatcatgag	1560
attatcaaaa aggatcttca cctagatcct tttaaattaa aaatgaagtt ttaaatcaat	1620

37847-517 Sequence Listing_ST25 (003)

ctaaagtata	tatgagtaaa	cttgggtctga	cagttacca	tgcttaata	gtgaggcacc	1680
tatctcagcg	atctgtctat	ttcgttcatc	catagttgcc	tgactccccg	tcgtgtagat	1740
aactacgata	cgggagggct	taccatctgg	ccccagtgt	gcaatgatac	cgcgagaccc	1800
acgtcaccg	gctccagatt	tatcagcaat	aaaccagcca	gccggaagg	ccgagcgag	1860
aagtggtcct	gcaactttat	ccgcctccat	ccagtctatt	aattgttgcc	gggaagctag	1920
agtaagtagt	tcgccagtta	atagtttgcg	caacgttggt	gccattgcta	caggcatcgt	1980
gggtgcacgc	tcgtcgtttg	gtatggcttc	attcagctcc	ggttcccaac	gatcaaggcg	2040
agttacatga	tccccatgt	tgtgcaaaaa	agcggttagc	tccttcggtc	ctccgatcgt	2100
tgtcagaagt	aagttggccg	cagtgttatc	actcatgggt	atggcagcac	tgcataattc	2160
tcttactgtc	atgccatccg	taagatgctt	ttctgtgact	ggtgagtact	caaccaagtc	2220
attctgagaa	tagtgtatgc	ggcgaccgag	ttgctcttgc	ccggcgtcaa	tacgggataa	2280
taccgcgcca	catagcagaa	ctttaaaagt	gctcatcatt	ggaaaacgtt	cttcggggcg	2340
aaaactctca	aggatcttac	cgctgttgag	atccagttcg	atgtaacca	ctcgtgcacc	2400
caactgatct	tcagcatctt	ttactttcac	cagcgtttct	gggtgagcaa	aaacaggaag	2460
gcaaaatgcc	gcaaaaaagg	gaataagggc	gacacggaaa	tgttgaatac	tcatactctt	2520
cctttttcaa	tattattgaa	gcatttatca	gggttattgt	ctcatgagcg	gatacatatt	2580
tgaatgtatt	tagaaaaata	aacaaatagg	ggttccgcgc	acatttcccc	gaaaagtgcc	2640
acctgacgtc	taagaaacca	ttattatcat	gacattaacc	tataaaaata	ggcgtatcac	2700
gaggcccttt	cgtctcgcgc	gtttcgggtga	tgacggtgaa	aacctctgac	acatgcagct	2760
cccggagacg	gtcacagctt	gtctgtaagc	ggatgccggg	agcagacaag	cccgtcaggg	2820
cgcgtcagcg	gggtgtggcg	gggtgcgggg	ctggcttaac	tatgcggcat	cagagcagat	2880
tgtactgaga	gtgcaccata	tatgcgggtgt	gaaataccgc	acagatgcgt	aaggagaaaa	2940
taccgcatca	ggcgcctcct	caacctgtat	attcgtaaac	cacgccaat	gggagctgtc	3000
tcaggtttgt	tcctgattgg	ttacggcgcg	tttcgcatca	ttgttgagtt	tttccgccag	3060
cccgacgcgc	agtttaccgg	tgccctgggtg	cagtacatca	gcatggggca	aattctttcc	3120
atcccgatga	ttgtcgcggg	tgtgatcatg	atggctctggg	catatcgtcg	cagcccacag	3180

37847-517 Sequence Listing_ST25 (003)

caacacgttt cctgaggaac catgaaacag tatttagaac tgatgcaaaa agtgctcgac	3240
gaaggcacac agaaaaacga ccgtaccgga accggaacgc tttccatttt tggtcatcag	3300
atgcgtttta acctgcaaga tggattcccg ctggtgacaa ctaaacgttg ccacctgcgt	3360
tccatcatcc atgaactgct gtggtttctg cagggcgaca ctaacattgc ttatctacac	3420
gaaaacaatg tcaccatctg ggacgaatgg gccgatgaaa acggcgacct cgggccagt	3480
tatggtaaac agtggcgcg cctggccaacg ccagatgggtc gtcattattga ccagatcact	3540
acggtactga accagctgaa aaacgacctg gattcgcgcc gcattattgt ttcagcgtgg	3600
aacgtaggcg aactggataa aatggcgctg gcaccgtgcc atgcattctt ccagttctat	3660
gtggcagacg gcaaactctc ttgccagctt tatcagcgct cctgtgacgt cttcctcggc	3720
ctgccgttca acattgccag ctacgcgtta ttggtgcata tgatggcgca gcagtgcgat	3780
ctggaagtgg gtgattttgt ctggaccggt ggcgacacgc atctgtacag caaccatatg	3840
gatcaaactc atctgcaatt aagccgcgaa ccgcgtccgc tgccgaagtt gattatcaaa	3900
cgtaaaccgg aatccatctt cgactaccgt ttcgaagact ttgagattga aggctacgat	3960
ccgcatccgg gcattaaagc gccggtggct atctaattac gaaacatcct gccagagccg	4020
acgccagtgt gcgtcggttt ttttaccctc cgttaaattc ttcgagacgc cttcccgaag	4080
gcgccattcg ccattcaggc tgcgcaactg ttgggaaggc cgatcggtgc gggcctcttc	4140
gctattacgc cagctggcga aagggggatg tgctgcaagg cgattaagtt gggtaacgcc	4200
agggttttcc cagtcacgac gttgtaaaac gacggccagt gccaagcttt ctttaatgaa	4260
gcagggcac aggacggtat ctttgtggag aaagcagagt aatcttattc agcctgactg	4320
gtgggaaacc accagtcaga atgtgttagc gcatgttgac aaaaatacca ttagtcacat	4380
tatccgtcag tcggacgaca tggtagataa cctgtttatt atgcgttttg atcttacgtt	4440
taatattacc tttatgcgat gaaacggtct tggctttgat attcatttgg tcagagattt	4500
gaatggttcc ctgacctgcc atccacattc gcaacatact cgattcgggt cggctcaatg	4560
ataacgtcgg catatttaaa aacgagggtta tcgttgtctc ttttttcaga atatcgccaa	4620
ggatatcgtc gagagattcc ggtttaatcg atttagaact gatcaataaa ttttttctga	4680
ccaatagata ttcataaaaa tgaacattgg caattgccat aaaaacgata aataacgtat	4740

37847-517 Sequence Listing_ST25 (003)

tgggatgttg	attaatgatg	agcttgatac	gctgactggt	agaagcatcg	tggatgaaac	4800
agtcctcatt	aataaacacc	actgaagggc	gctgtgaatc	acaagctatg	gcaaggtcat	4860
caacggtttc	aatgtcgttg	atttctcttt	ttttaacccc	tctactcaac	agatacccg	4920
ttaaaccctag	tcgggtgtaa	ctacataaat	ccataataat	cgttgacatg	gcataccctc	4980
actcaatgcg	taacgataat	tccccttacc	tgaatatttc	atcatgacta	aacggaacaa	5040
catgggtcac	ctaatacgcc	actctcgca	tttttcaggc	ggacttacta	tcccgtaaag	5100
tgttgataa	tttgcctgga	attgtcttaa	agtaaagtaa	atgttgatg	atgtgagtga	5160
gcttaaaaca	aatatttcgc	tgcaggagta	tcctggaaga	tgttcgtaga	agcttactgc	5220
tcacaagaaa	aaaggcacgt	catctgacgt	gcctttttta	tttgactac	cctgtacgat	5280
tactgcagct	cgagttagga	taccggcact	ttgatccaac	cagtcgggta	gatatccggt	5340
gcttcggagt	gctggaacca	acggctcggc	acaataacag	tcttatccat	attagggttc	5400
agccaggcac	cccaccaaga	aaacgtgctg	ttacaaatga	tgtgatgttt	gcaatgagac	5460
atcagcatca	tatcctgcc	ggagtcttca	tcagtgttcc	agtcaatata	aaccgcattc	5520
tgcatgggca	gattttcttt	aaccacgcg	atatcgtcgg	agaagatata	gtaagatggg	5580
ctagcaacac	gacgggacat	ttccgcgata	gcattctggt	aatacggcag	ctggcacacg	5640
gaaccggtag	tagcccagt	tttcggctgc	agatagtcac	cacgacgaat	gtgcagggaa	5700
accgcgtttt	catctttgtc	caggatttcc	agcatgttca	ggctgcggga	atttgctttg	5760
ttcttatcaa	aggtgaagga	ttcacgcact	tcgtctttga	tatcagcgaa	gaaacgctcg	5820
ctctgataga	aacctttaaa	gtacagcagc	ggccagaaat	acttcttctc	gaacgcacgc	5880
agagagtctg	gcgctgctt	gcgttcgtag	atTTTTTTTaa	aaaacaggaa	ttcgataact	5940
TTTTTcagcg	gttggttgat	gcagaattcg	gtgtgcggca	ggttgaacac	gcggtgcatt	6000
tcgtaaccgt	aatggacttt	gtaatgcac	atgtcgctca	ggtcgatacg	gaccttcggg	6060
taatactttt	tcatacgag	atagaaagca	tagataaaca	tctggttgcc	cagaccgcca	6120
gtcactttga	tcagacgcat	tatatctcct	tcttg			6155

<210> 292
 <211> 278
 <212> PRT

37847-517 Sequence Listing_ST25 (003)

<213> Artificial Sequence

<220>

<223> artificial construct

<220>

<221> MISC_FEATURE

<222> (1)..(13)

<223> wherein X is any amino acid

<400> 292

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Met Phe Gln
1 5 10 15

Tyr Ala Leu Tyr Lys Ala Phe Glu Gln Lys His Ile Asp Val Tyr Ala
20 25 30

Asp Leu Ala Trp Tyr Lys Asn Lys Ser Val Lys Phe Glu Leu Tyr Asn
35 40 45

Phe Gly Ile Lys Ile Asn Val Ala Ser Glu Lys Asp Ile Asn Arg Leu
50 55 60

Ser Asp Cys Gln Ala Asp Phe Val Ser Arg Ile Arg Arg Lys Ile Phe
65 70 75 80

Gly Lys Lys Lys Ser Phe Val Ser Glu Lys Asn Asp Ser Cys Tyr Glu
85 90 95

Asn Asp Ile Leu Arg Met Asp Asn Val Tyr Leu Ser Gly Tyr Trp Gln
100 105 110

Thr Glu Lys Tyr Phe Ser Asn Thr Arg Glu Lys Leu Leu Glu Asp Tyr
115 120 125

Ser Phe Ala Leu Val Asn Ser Gln Val Ser Glu Trp Glu Asp Ser Ile
130 135 140

Arg Asn Lys Asn Ser Val Ser Ile His Ile Arg Arg Gly Asp Tyr Leu
145 150 155 160

37847-517 Sequence Listing_ST25 (003)

Gln Gly Glu Leu Tyr Gly Gly Ile Cys Thr Ser Leu Tyr Tyr Ala Glu
 165 170 175

Ala Ile Glu Tyr Ile Lys Met Arg Val Pro Asn Ala Lys Phe Phe Val
 180 185 190

Phe Ser Asp Asp Val Glu Trp Val Lys Gln Gln Glu Asp Phe Lys Gly
 195 200 205

Phe Val Ile Val Asp Arg Asn Glu Tyr Ser Ser Ala Leu Ser Asp Met
 210 215 220

Tyr Leu Met Ser Leu Cys Lys His Asn Ile Ile Ala Asn Ser Ser Phe
 225 230 235 240

Ser Trp Trp Ala Ala Trp Leu Asn Arg Asn Glu Glu Lys Ile Val Ile
 245 250 255

Ala Pro Arg Arg Trp Leu Asn Gly Lys Cys Thr Pro Asp Ile Trp Cys
 260 265 270

Lys Lys Trp Ile Arg Ile
 275

<210> 293

<211> 300

<212> PRT

<213> Helicobacter pylori

<400> 293

Met Ala Phe Lys Val Val Gln Ile Cys Gly Gly Leu Gly Asn Gln Met
 1 5 10 15

Phe Gln Tyr Ala Phe Ala Lys Ser Leu Gln Lys His Ser Asn Thr Pro
 20 25 30

Val Leu Leu Asp Ile Thr Ser Phe Asp Trp Ser Asp Arg Lys Met Gln
 35 40 45

37847-517 Sequence Listing_ST25 (003)

Leu	Glu	Leu	Phe	Pro	Ile	Asp	Leu	Pro	Tyr	Ala	Ser	Ala	Lys	Glu	Ile	50	55	60	
Ala	Ile	Ala	Lys	Met	Gln	His	Leu	Pro	Lys	Leu	Val	Arg	Asp	Ala	Leu	65	70	75	80
Lys	Cys	Met	Gly	Phe	Asp	Arg	Val	Ser	Gln	Glu	Ile	Val	Phe	Glu	Tyr	85	90	95	
Glu	Pro	Lys	Leu	Leu	Lys	Pro	Ser	Arg	Leu	Thr	Tyr	Phe	Phe	Gly	Tyr	100	105	110	
Phe	Gln	Asp	Pro	Arg	Tyr	Phe	Asp	Ala	Ile	Ser	Pro	Leu	Ile	Lys	Gln	115	120	125	
Thr	Phe	Thr	Leu	Pro	Pro	Pro	Pro	Glu	Asn	Asn	Lys	Asn	Asn	Asn	Lys	130	135	140	
Lys	Glu	Glu	Glu	Tyr	Gln	Cys	Lys	Leu	Ser	Leu	Ile	Leu	Ala	Ala	Lys	145	150	155	160
Asn	Ser	Val	Phe	Val	His	Ile	Arg	Arg	Gly	Asp	Tyr	Val	Gly	Ile	Gly	165	170	175	
Cys	Gln	Leu	Gly	Ile	Asp	Tyr	Gln	Lys	Lys	Ala	Leu	Glu	Tyr	Met	Ala	180	185	190	
Lys	Arg	Val	Pro	Asn	Met	Glu	Leu	Phe	Val	Phe	Cys	Glu	Asp	Leu	Glu	195	200	205	
Phe	Thr	Gln	Asn	Leu	Asp	Leu	Gly	Tyr	Pro	Phe	Met	Asp	Met	Thr	Thr	210	215	220	
Arg	Asp	Lys	Glu	Glu	Glu	Ala	Tyr	Trp	Asp	Met	Leu	Leu	Met	Gln	Ser	225	230	235	240
Cys	Gln	His	Gly	Ile	Ile	Ala	Asn	Ser	Thr	Tyr	Ser	Trp	Trp	Ala	Ala	245	250	255	

37847-517 Sequence Listing_ST25 (003)

Tyr Leu Ile Glu Asn Pro Glu Lys Ile Ile Ile Gly Pro Lys His Trp
260 265 270

Leu Phe Gly His Glu Asn Ile Leu Cys Lys Glu Trp Val Lys Ile Glu
275 280 285

Ser His Phe Glu Val Lys Ser Gln Lys Tyr Asn Ala
290 295 300

<210> 294

<211> 297

<212> PRT

<213> Escherichia coli

<400> 294

Met Ser Ile Ile Arg Leu Gln Gly Gly Leu Gly Asn Gln Leu Phe Gln
1 5 10 15

Phe Ser Phe Gly Tyr Ala Leu Ser Lys Ile Asn Gly Thr Pro Leu Tyr
20 25 30

Phe Asp Ile Ser His Tyr Ala Glu Asn Asp Asp His Gly Gly Tyr Arg
35 40 45

Leu Asn Asn Leu Gln Ile Pro Glu Glu Tyr Leu Gln Tyr Tyr Thr Pro
50 55 60

Lys Ile Asn Asn Ile Tyr Lys Phe Leu Val Arg Gly Ser Arg Leu Tyr
65 70 75 80

Pro Glu Ile Phe Leu Phe Leu Gly Phe Cys Asn Glu Phe His Ala Tyr
85 90 95

Gly Tyr Asp Phe Glu Tyr Ile Ala Gln Lys Trp Lys Ser Lys Lys Tyr
100 105 110

Ile Gly Tyr Trp Gln Ser Glu His Phe Phe His Lys His Ile Leu Asp
115 120 125

Leu Lys Glu Phe Phe Ile Pro Lys Asn Val Ser Glu Gln Ala Asn Leu

37847-517 Sequence Listing_ST25 (003)

130

135

140

Leu Ala Ala Lys Ile Leu Glu Ser Gln Ser Ser Leu Ser Ile His Ile
145 150 155 160

Arg Arg Gly Asp Tyr Ile Lys Asn Lys Thr Ala Thr Leu Thr His Gly
165 170 175

Val Cys Ser Leu Glu Tyr Tyr Lys Lys Ala Leu Asn Lys Ile Arg Asp
180 185 190

Leu Ala Met Ile Arg Asp Val Phe Ile Phe Ser Asp Asp Ile Phe Trp
195 200 205

Cys Lys Glu Asn Ile Glu Thr Leu Leu Ser Lys Lys Tyr Asn Ile Tyr
210 215 220

Tyr Ser Glu Asp Leu Ser Gln Glu Glu Asp Leu Trp Leu Met Ser Leu
225 230 235 240

Ala Asn His His Ile Ile Ala Asn Ser Ser Phe Ser Trp Trp Gly Ala
245 250 255

Tyr Leu Gly Thr Ser Ala Ser Gln Ile Val Ile Tyr Pro Thr Pro Trp
260 265 270

Tyr Asp Ile Thr Pro Lys Asn Thr Tyr Ile Pro Ile Val Asn His Trp
275 280 285

Ile Asn Val Asp Lys His Ser Ser Cys
290 295