

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

(19) World Intellectual Property  
Organization  
International Bureau



(43) International Publication Date  
6 April 2017 (06.04.2017)

(10) International Publication Number  
**WO 2017/059424 A1**

(51) International Patent Classification:

C12N 15/09 (2006.01) C12N 15/67 (2006.01)  
C12N 15/11 (2006.01) C12Q 1/68 (2006.01)  
C12N 15/63 (2006.01)

(21) International Application Number:

PCT/US2016/055162

(22) International Filing Date:

3 October 2016 (03.10.2016)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/236,477 2 October 2015 (02.10.2015) US  
62/263,404 4 December 2015 (04.12.2015) US

(71) Applicant: UNIVERSITY OF UTAH RESEARCH  
FOUNDATION [US/US]; 615 Arapeen Drive, Suite 310,  
Salt Lake City, Utah 84108 (US).

(72) Inventors: HUGHES, Kelly T.; 1534 Bryan Avenue, Salt  
Lake City, Utah 84105 (US). CHEVANCE, Fabienne;  
1534 Bryan Avenue, Salt Lake City, Utah 84105 (US).

(74) Agents: CALVO, Paul A. et al.; Sterne, Kessler, Goldstein  
& Fox P.L.L.C., 1100 New York Avenue, N.W., Washing-  
ton, District of Columbia 20005 (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, DJ, 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,  
KW, 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))



WO 2017/059424 A1

(54) Title: COMPOSITIONS FOR ADJUSTABLE RIBOSOME TRANSLATION SPEED AND METHODS OF USE

(57) Abstract: The present invention relates to the effects of codon context and synonymous codon changes on mRNA translation.

# COMPOSITIONS FOR ADJUSTABLE RIBOSOME TRANSLATION SPEED AND METHODS OF USE

## BACKGROUND OF THE INVENTION

In 1961, Crick *et al.* published their landmark *Nature* paper, "General nature of the genetic  
5 code for proteins" (Crick, F. H., *et al. Nature* 192:1227-32 (1961)). Using a simple, yet clever  
genetic approach of combining +1 and -1 frameshift mutations, Crick *et al.* provided evidence for  
the triplet code that is part of the central dogma of molecular biology: DNA is transcribed into  
mRNA by RNA polymerase, which is translated into protein by ribosomes. The genetic code for  
the amino acid sequences of proteins is composed of 64 triplet codons that specify the twenty  
10 amino acids and sites of translation termination (stop codons). The interaction between a specific  
codon in the mRNA being translated and the three bases that define the anticodon of the cognate  
aminoacyl-tRNA (bases 36, 35 and 34 within the anticodon loop) determines the decoding process  
(Grosjean, H., *et al., FEBS Lett.* 584:252-264 (2010)).

It is now known that the decoding process is more complicated. The ability of the ribosome  
15 to translate a given codon is highly influenced by its neighboring codons (referred to as the codon  
context effect on translation) and perhaps mRNA secondary structure (Goodman, D. B., *et al.,  
Science* 342:475-478 (2013)). Translation of a given codon is also influenced by the location of the  
codon within the mRNA open reading frame. Further, while synonymous mutations (nucleotide  
changes within the codon that do not change the encoded amino acid) were long thought to be  
20 evolutionarily silent (Nei, M. *Mol. Biol. Evol.* 22:2318-2342 (2005)), it is now known that  
synonymous codons affect translation speed and accuracy, co-translational folding, protein  
secretion and overall expression levels (Hunt, R. C., *et al., Trends Genet.* 30:308-321 (2014)).  
Recent evidence supports an association between mutations resulting in synonymous codons and  
human diseases such as cancer (Sauna, Z. E., and C. Kimchi-Sarfaty, *Nat Rev Genet* 12:683-91  
25 (2011); Supek, F., *et al., Cell* 156:1324-35 (2014)).

## BRIEF SUMMARY OF THE INVENTION

The present invention provides synonymous changes in codons pairs that result in differing  
ribosome translation speed and accuracy, protein folding and expression. In one embodiment, the  
invention provides a 15-base pair sequence for manipulation of expression of proteins in

- 2 -

genetically modified organisms. The present invention also provides a method to screen for polynucleotide mutations that result in deleterious effects associated with disease states.

In one embodiment, the present invention provides engineered FlgM coding sequences, or ten-codon-fragments thereof, that, when appended to a heterologous polynucleotide sequence encoding a desired protein, increases or decreases the translation efficiency of that protein which results in modulated production of a polypeptide of interest.

In one embodiment, the present invention further utilizes a bacterial secretion system (such as the flagellar type III secretion (T3S) system (T3SS)) to secrete the desired protein into culture media and therefore provide a protein production method that is more efficient than a method requiring cell lysis for protein recovery.

The present invention also provides a recombinant nucleic acid molecule comprising a mutant triacontanucleotide sequence having ten codons, wherein at least one of the sixth, seventh, eighth, ninth and tenth codons, in the 5' to 3' direction, are synonymous codons. In certain embodiments, the recombinant nucleic acid molecule consists of a mutant triacontanucleotide sequence. In another embodiment, the mutant triacontanucleotide sequence encodes the amino acid sequence MSIDRTSPLK (SEQ ID NO:1). In other embodiments, the mutant triacontanucleotide sequence is part of a sequence that encodes the full-length FlgM amino acid sequence: MSIDRTSPLKPVSTVQTRETS DTPVQKTRQEK TSAATSASVTLSDAQAKLMQPGVSDINMERVEALKTAIRNGELKMDTGK IADSLIREAQSYLQSK (SEQ ID NO: 3).

In one embodiment, the mutant triacontanucleotide sequence is operably linked to a polynucleotide sequence that encodes a polypeptide of interest. In another embodiment, the mutant triacontanucleotide sequence is operably linked, in the 5' to 3' direction, to a polynucleotide sequence that encodes a polypeptide of interest. In certain embodiments, the mutant triacontanucleotide sequence encodes the amino acid sequence SEQ ID NO: 3 and is operably linked, in the 5' to 3' direction, to a polynucleotide sequence that encodes a polypeptide of interest.

In one embodiment, the mutant triacontanucleotide sequence is further operably linked to a cleavage site nucleotide sequence. In one embodiment, the cleavage site encodes a Tobacco Etch Virus (TEV) protease cleavage site or an Enterokinase (ETK) cleavage site. In another embodiment, the mutant triacontanucleotide sequence is operably linked to a cleavage site nucleotide sequence and a polynucleotide sequence that encodes a polypeptide.

- 3 -

In one embodiment, the mutant triacontanucleotide sequence replaces the ten 5'-most codons of a polynucleotide sequence that encodes a polypeptide of interest. In one embodiment, the polypeptide of interest is heterologous to the mutant triacontanucleotide sequence.

In one embodiment, the mutant triacontanucleotide sequence is a ribonucleic acid sequence.

5 In another embodiment, the mutant triacontanucleotide sequence is a deoxyribonucleic acid sequence.

In one embodiment, the mutant triacontanucleotide sequence is a synthetic sequence.

In one embodiment, the sixth and eighth codons of the mutant triacontanucleotide sequence, in the 5' to 3' direction, are synonymous codons.

10 In one embodiment, a vector comprises a mutant sequence of the invention. In another embodiment, the vector is an expression vector.

In one embodiment, a vector comprises the mutant triacontanucleotide sequence operably linked to a polynucleotide sequence that encodes a polypeptide of interest and is operably linked to a nucleotide sequence that encodes a cleavage site. In another embodiment, the cleavage site  
15 nucleotide sequence is between the mutant triacontanucleotide sequence and the polynucleotide sequence of interest. In another embodiment the vector further comprises a nucleic acid sequence encoding a purification tag wherein the nucleic acid sequence is operably linked to the mutant triacontanucleotide sequence.

In one embodiment, the mutant triacontanucleotide sequence is operably linked to a  
20 polynucleotide sequence encoding a polypeptide of interest, a nucleic acid sequence encoding a purification tag, and, optionally, a nucleotide sequence encoding a cleavage site.

In one embodiment, the mutant triacontanucleotide sequence or polynucleotide sequence is operably linked to a heterologous promoter. In another embodiment, the promoter is an inducible promoter, a constitutive promoter, or a tissue specific promoter. An inducible promoter of the  
25 present invention may be an arabinose inducible promoter such as the *ParaBAD* (*ΔaraBAD*) promoter or a salicylate inducible promoter such as the *Psal* promoter. In one embodiment, a host cell comprises the recombinant molecule. In another embodiment, the host cell is a bacterial, fungal, yeast, viral, plant, insect, or mammalian cell. In another embodiment, the host cell is a *Salmonella* or *Escherichia coli* cell. In another embodiment, the host cell is a *Salmonella enterica*  
30 cell. In one embodiment, the mutant triacontanucleotide sequence operably linked to the

- 4 -

polynucleotide sequence, which is optionally further operably linked to a cleavage site, is operatively integrated into the host cell genome.

In one embodiment, a host cell comprises a vector of the invention.

5 In one embodiment, the polynucleotide sequence within the vector has been codon optimized for expression within a host cell.

The present invention provides a recombinant molecule comprising a nucleic acid sequence having ten codons, wherein at least one of the sixth, seventh, eighth, ninth and tenth codons, in the 5' to 3' direction, are synonymous codons. In one embodiment, the nucleic acid sequence is operably linked to a polynucleotide sequence.

10 The present invention provides a recombinant nucleic acid molecule having a modulated translation speed, comprising the decacodon sequence 5'-AUGAGCAUUGACCGUACCUCACCUUUGAAA-3' (SEQ ID NO:2), wherein at least one of the sixth, seventh, eighth, ninth and tenth codons, in the 5' to 3' direction, are synonymous codons. In certain embodiments, the nucleic acid molecule comprises this nucleic acid sequence and encodes the amino acid sequence SEQ ID NO: 3.

15 In one embodiment, the sixth codon is the synonymous codon ACU, ACA, or ACG and wherein the recombinant molecule has increased translation speed as compared to a molecule comprising the decacodon sequence. In one embodiment, the sixth codon is a synonymous codon and the molecule has between about a 5- to about 15-fold increase in translation speed as compared to a molecule comprising the native decacodon sequence. In another embodiment, the molecule has between about a 9- to about 15-fold increase in translation speed as compared to a molecule comprising the native decacodon sequence.

20 In one embodiment, the eighth codon is the synonymous codon CCG and wherein the recombinant molecule has decreased translation speed as compared to a molecule comprising the native decacodon sequence. In one embodiment, the eighth codon is a synonymous codon and the recombinant molecule has between about a 0.01 to about 0.10-fold decrease in translation speed as compared to a molecule comprising the native decacodon sequence. In another embodiment, the molecule has between about a 0.03- to about 0.07-fold decrease in translation speed as compared to a molecule comprising the native decacodon sequence.

30 In one embodiment, the sixth codon is the synonymous codon ACU, the eighth codon is the synonymous codon CCG, and wherein the recombinant molecule has increased translation speed

- 5 -

as compared to a molecule comprising the native decacodon sequence. In one embodiment, the sixth codon is a synonymous codon, the eighth codon is a synonymous codon, and wherein the recombinant molecule has between about a 10- to 50-fold increase in translation speed as compared to a molecule comprising the native decacodon sequence. In another embodiment, the molecule has between about a 20- to about 50-fold increase in translation speed. In another embodiment, the molecule has between about a 30- to about 50-fold increase in translation speed. In another embodiment, the molecule has between about a 40- to about 50-fold increase in translation speed. In another embodiment, the molecule has between about a 27.5- to 40.5-fold increase in translation speed as compared to a molecule comprising the native decacodon sequence.

In one embodiment, the decacodon sequence is operably linked to a polynucleotide sequence that encodes a polypeptide of interest.

The present invention provides a method of modulating protein production within a host cell, comprising culturing a host cell of the invention under conditions sufficient for protein expression, wherein a recombinant nucleic acid molecule has been stably introduced into the host cell. In one embodiment, the recombinant nucleic acid molecule comprises a mutant triacontanucleotide sequence having ten codons, wherein at least one of the sixth, seventh, eighth, ninth and tenth codons, in the 5' to 3' direction, are synonymous codons. In one embodiment, the recombinant molecule comprises a nucleic acid sequence having ten codons, wherein at least one of the sixth, seventh, eighth, ninth and tenth codons, in the 5' to 3' direction, are synonymous codons. In another embodiment, the nucleic acid sequence comprising the synonymous codons encodes the amino acid sequence SEQ ID NO: 3.

In one embodiment, the sixth codon is the synonymous codon ACU, ACA, or ACG and wherein protein production is increased as compared to a corresponding host cell utilizing wild type triacontanucleotide sequence. In one embodiment, the sixth codon is a synonymous codon and the molecule has between about a 5- to about 15-fold increase in protein production as compared to a corresponding host cell utilizing a wild type triacontanucleotide sequence. In another embodiment, the molecule has between about a 9- to about 15-fold increase in protein production as compared to a corresponding host cell utilizing a wild type triacontanucleotide sequence.

In one embodiment, the eighth codon is the synonymous codon CCG and wherein protein production is decreased as compared to a corresponding host cell utilizing a wild type

- 6 -

triacontanucleotide sequence. In one embodiment, the eighth codon is a synonymous codon and the recombinant molecule has between about a 0.01 to about 0.10-fold decrease in protein production as compared to a corresponding host cell utilizing a wild type triacontanucleotide sequence. In another embodiment, the molecule has between about a 0.03- to about 0.07-fold decrease in protein production as compared to a corresponding host cell utilizing a wild type triacontanucleotide sequence.

In one embodiment, the sixth codon is the synonymous codon ACU, the eighth codon is the synonymous codon CCG, and wherein protein production is increased as compared to a corresponding host cell utilizing wild type triacontanucleotide sequence.

In one embodiment, the sixth codon is a synonymous codon, the eighth codon is a synonymous codon, and wherein the recombinant molecule has between about a 10- to 50-fold increase in protein production as compared to a corresponding host cell utilizing a wild type triacontanucleotide sequence. In another embodiment, the molecule has between about a 20- to about 50-fold increase in protein production as compared to a corresponding host cell utilizing a wild type triacontanucleotide sequence. In another embodiment, the molecule has between about a 30- to about 50-fold increase in protein production as compared to a corresponding host cell utilizing a wild type triacontanucleotide sequence. In another embodiment, the molecule has between about a 40- to about 50-fold increase in protein production as compared to a corresponding host cell utilizing a wild type triacontanucleotide sequence. In another embodiment, the molecule has between about a 27.5- to 40.5-fold increase in protein production as compared to a corresponding host cell utilizing a wild type triacontanucleotide sequence.

The present invention provides a method of increasing the translation speed of a polynucleotide sequence, comprising providing a polynucleotide sequence that encodes a protein and is operably linked to a wild type triacontanucleotide sequence wherein the triacontanucleotide sequence comprises ten codons and encodes MSIDRTSPLK (SEQ ID NO:1), and mutating the triacontanucleotide sequence so that at least one of the sixth, seventh, eighth, ninth and tenth codons, in the 5' to 3' direction, are synonymous codons. In further embodiments, the wild type triacontanucleotide sequence and the mutant triacontanucleotide sequence encode the amino acid sequence SEQ ID NO: 3.

In one embodiment, the sixth codon of the resulting mutant triacontanucleotide sequence comprises the synonymous codon ACU, ACA, or ACG and wherein translation speed of the

- 7 -

polynucleotide sequence is increased compared to translation under the control of the wild type triacontanucleotide sequence.

5 In one embodiment, the eighth codon of the resulting mutant triacontanucleotide sequence comprises the synonymous codon CCG and wherein translation speed of the polynucleotide sequence is decreased compared to translation under the control of the wild type triacontanucleotide sequence.

In one embodiment, the sixth codon is the synonymous codon ACU, the eighth codon is the synonymous codon CCG, and wherein translation speed of the polynucleotide sequence is increased compared to translation under the control of the wild type triacontanucleotide sequence.

10 The present invention provides a method of producing polypeptides of interest comprising culturing a host cell that contains a mutant triacontanucleotide sequence operably linked to a polynucleotide sequence encoding the polypeptide of interest. In one embodiment, the mutant triacontanucleotide sequence and polynucleotide sequence are operably linked to a nucleotide sequence encoding a cleavage site and/or a nucleic acid sequence encoding a purification tag. In one embodiment, the peptide is purified from the culture media. In one embodiment, the purification tag is utilized to purify the peptide from the culture media. In one embodiment, the triacontanucleotide sequence encodes the FlgM protein having the amino acid sequence SEQ ID NO: 3 and the peptide is isolated from the FlgM protein. In another embodiment, the cleavage site is utilized to isolate the polypeptide from the FlgM protein.

20 The present invention provides a method of producing polypeptides of interest comprising culturing one or more bacterial cells having a Type III Secretion System and also contain a mutant triacontanucleotide sequence operably linked to a heterologous polynucleotide sequence that encodes a desired polypeptide, and wherein the polypeptide is secreted into the culture media. In one embodiment, the bacterial cells are either *Salmonella* or *Escherichia* cells. In one embodiment, the triacontanucleotide sequence encodes a full-length FlgM peptide. In one embodiment, the triacontanucleotide sequence and polynucleotide sequence are operably linked to a nucleotide sequence encoding a purification tag and/or a cleavage site. In one embodiment, the method utilizes is a continuous flow manufacturing system for production of the polypeptide of interest. In one embodiment, the continuous flow manufacturing system produces insulin, MaSp1, or MaSp2 protein.

25

30

- 8 -

## BRIEF DESCRIPTION OF THE DRAWINGS/FIGURES

Figure 1 depicts the *Salmonella enterica* flagellum gene regulation coupled to assembly.

Figure 2 depicts the effect of synonymous changes at codons Thr6 and Pro8 of *flgM* on FlgM anti- $\sigma^{28}$  activity in exponential cultures. Data is presented as levels of FlgM anti- $\sigma^{28}$  activity.

5 Wild-type codon (ACC) exhibited the lowest FlgM activity/luciferase activity for each of the 15 combinations of synonymous Thr6 and Pro8 codons assayed.

Figure 3 depicts the effects of synonymous codon mutagenesis on suppression of the Pro8 CCG translation-defective allele of *flgM*. Mutants that showed wild-type FlgM activity on indicator medium are shown above the FlgM sequence. Mutants with slightly increased FlgM  
10 activity are Ser2 (ACG to AGT), Thre6 (ACC to stop), Val12 (GTT to GTG and GTT to GTC). The remaining mutants depicted below the FlgM sequence (ten isolates) had no phenotype. The number in parenthesis denotes the number of repeats.

Figure 4 depicts the detection of FlgM::6xHis::ETK::Insulin-glargine chimeric proteins within secreted fractions (top) and cellular fractions (bottom) of engineered *Salmonella* cells via  
15 Western blot using anti-6xHis antibodies.

## DETAILED DESCRIPTION OF THE INVENTION

*Definitions*

To facilitate an understanding of the present invention, a number of terms and phrases are defined below. Art-recognized synonyms or alternatives of the following terms and phrases, even  
20 if not specifically described, are contemplated.

As used in the present disclosure and claims, the singular forms "a," "an," and "the" include plural forms unless the context clearly dictates otherwise; i.e., "a" means "one or more" unless indicated otherwise.

The terms "about" or "approximately" mean roughly, around, or in the regions of. The terms  
25 "about" or "approximately" further mean within an acceptable contextual error range for the particular value as determined by one of ordinary skill in the art, which will depend in part on how the value is measured or determined, i.e. the limitations of the measurement system or the degree of precision required for a particular purpose, e.g. the amount of a nutrient within a feeding formulation. When the terms "about" or "approximately" are used in conjunction with a numerical

- 9 -

range, it modifies that range by extending the boundaries above and below the numerical values set forth. For example "between about 5.5 to 6.5 g/l" means the boundaries of the numerical range extend below 5.5 and above 6.5 so that the particular value in question achieves the same functional result as within the range. For example, "about" and "approximately" can mean within  
5 one or more than one standard deviation as per the practice in the art. Alternatively, "about" and "approximately" can mean a range of up to 20%, preferably up to 10%, more preferably up to 5%, and more preferably up to 1% of a given value.

The term "and/or" as used in a phrase such as "A and/or B" is intended to include "A and B," "A or B," "A," and "B." Likewise, the term "and/or" as used in a phrase such as "A, B, and/or C" is  
10 intended to encompass each of the following embodiments: A, B, and C; A, B, or C; A or C; A or B; B or C; A and C; A and B; B and C; A (alone); B (alone); and C (alone).

Unless specified otherwise, all of the designations "A%-B%," "A-B%," "A% to B%," "A to B%," "A%-B," "A% to B" are given their ordinary and customary meaning. In some embodiments, these designations are synonyms.

The terms "substantially" or "substantial" mean that the condition described or claimed functions in all important aspects as the standard described. Thus, "substantially free" is meant to encompass conditions that function in all important aspects as free conditions, even if the numerical values indicate the presence of some impurities or substances. "Substantial" generally means a value greater than 90%, preferably greater than 95%, most preferably greater than 99%.  
15 Where particular values are used in the specification and in the claims, unless otherwise stated, the term "substantially" means with an acceptable error range for the particular value.

"Optional" or "optionally" means that the subsequently described event, circumstance, or material may or may not occur or be present, and that the description includes instances where the event, circumstance, or material occurs or is present and instances where it does not occur or is not  
20 present.

The terms "transformation" and "transfection" mean the introduction of a nucleic acid molecule (e.g., an expression vector) into a recipient cell. Transformation or transfection techniques are well known by the art. It may be specified that a nucleic acid sequence (e.g., a coding sequence carried by an expression vector) is introduced (integrated) into the genomic  
30 (chromosomal DNA) of said recipient (host) cell.

- 10 -

"FlgM" and "*flgM*" refers to the protein and polynucleotide sequence, respectively, that negatively regulates flagellin synthesis in Type III Secretion Systems (Gillen and Hughes, *Molecular Characterization of flgM, a Gene Encoding a Negative Regulator of Flagellin Synthesis in Salmonella typhimurium*, J. Bacteriology 173(20): 6453-6459 (1991)).

5 The term "recombinant" or "mutant" means that the referenced product has been altered with respect to its naturally occurring counterpart.

The term "triacontanucleotide sequence" means a nucleotide sequence that is at least thirty nucleotides long. For example, a triacontanucleotide sequence of the present invention may consist of thirty nucleotides or may be a part of a larger sequence, encoding for example, a full-length  
10 FlgM protein.

The term "codon" refers to the three consecutive nucleotides that together encode one amino acid.

The term "decacodon" means a nucleotide sequence comprising at least five codons (i.e., at least fifteen nucleotides). For example, a decacodon sequence of the present invention may consist  
15 of 10 codons or may comprise about 97 codons and encode a full-length FlgM peptide.

The term "synonymous codon" means a codon that is altered as compared to wild-type but that encodes the same amino acid as does the wild-type codon.

The terms "nucleic acid," "polynucleotide," and "nucleotide" are used synonymously herein. The term "polynucleotide," when used in singular or plural, generally refers to any a  
20 polyribonucleotide or polydeoxyribonucleotide, which can be unmodified RNA or DNA or modified RNA or DNA. Thus, for instance, polynucleotides as defined herein include, without limitation, single- and double-stranded DNA, DNA including single- and double-stranded regions, single- and double-stranded RNA, and RNA including single- and double-stranded regions, hybrid  
25 molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or include single- and double-stranded regions. Thus, DNAs or RNAs with backbones modified for stability or for other reasons are "polynucleotides" as that term is intended herein. Moreover, DNAs or RNAs comprising unusual bases, such as inosine, or modified bases, such as tritiated bases, are included within the term "polynucleotides" as defined herein. In general, the  
30 term "polynucleotide" embraces all chemically, enzymatically and/or metabolically modified forms of unmodified polynucleotides. Polynucleotides can be made by a variety of methods,

- 11 -

including in vitro recombinant DNA-mediated techniques and by expression of DNAs in cells and organisms

The terms "protein," "polypeptide," and "peptide" are used synonymously herein. As used herein, the term "polypeptide" is intended to encompass a singular "polypeptide" as well as plural "polypeptides," and refers to a molecule composed of monomers (amino acids) linearly linked by amide bonds (also known as peptide bonds). The term "polypeptide" refers to any chain or chains of two or more amino acids, and does not refer to a specific length of the product. Thus, peptides, dipeptides, tripeptides, oligopeptides, "protein," "amino acid chain," or any other term used to refer to a chain or chains of two or more amino acids, are included within the definition of "polypeptide," and the term "polypeptide" can be used instead of, or interchangeably with any of these terms. The term "polypeptide" is also intended to refer to the products of post-expression modifications of the polypeptide, including without limitation glycosylation, acetylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, or modification by non-naturally occurring amino acids. A polypeptide can be derived from a natural biological source or produced by recombinant technology, but is not necessarily translated from a designated nucleic acid sequence. It can be generated in any manner, including by chemical synthesis. Certain embodiments of the present invention comprise a fusion peptide comprising at least a FlgM peptide and a polypeptide wherein the polypeptide and FlgM peptide do not naturally associate with each other (i.e., the polypeptide is heterologous to the FlgM peptide). Such a polypeptide may be referred to as a "heterologous polypeptide," a "target polypeptide," a "desired polypeptide" or a "polypeptides of interest."

An "isolated" biological component (such as a nucleic acid molecule or protein) has been substantially separated or purified away from other biological components in the cell of the organism in which the component naturally occurs, i.e., other chromosomal and extra-chromosomal DNA and RNA, proteins and organelles. Nucleic acids and proteins that have been "isolated" include nucleic acids and proteins purified by standard purification methods. The term also embraces nucleic acids and proteins prepared by recombinant expression in a host cell as well as chemically synthesized nucleic acids.

The term "heterologous," "foreign" or "non-native" refers to two structures, for example, that are not naturally associated with each other. For example, wherein a nucleotide sequence is operably linked to a heterologous promoter, the nucleotide sequence is not naturally associated

- 12 -

with the promoter even though the nucleotide sequence and the promoter sequence may originate from the same organism.

The terms "ribonucleic acid" and "RNA" are used herein synonymously.

The terms "deoxyribonucleic acid" and "DNA" are used herein synonymously.

5 A "vector" as used herein may be a cloning vector, an expression vector, a plasmid including a BAC or a YAC vector, or other nucleic acid molecule capable of transporting a polynucleotide sequence into a cell. A vector may be designed for storage of the polynucleotide sequence, cloning, and/or designed for expression of the polynucleotide sequence. "Plasmid" and "vector" may be used interchangeably, as a plasmid is a commonly used form of a vector. The term  
10 "expression vector" includes any vector, (e.g., a plasmid, cosmid or phage chromosome) containing a gene construct in a form suitable for expression by a cell (e.g., linked to a transcriptional control element). "Construct" or "gene construct" as used herein does not include an entire chromosomal and/or mitochondrial genome. Expression vectors suitable for use within a specified host cell are known and readily identified using routine experimentation and known  
15 techniques (see, e.g., U.S. Patent Nos. 7,785,830; 8,663,980; 8,628,954, incorporated herein by reference in their entirety).

A "host cell" as used herein refers to a cell into which a molecule is or has been introduced. Generally, a host cell herein refers to a cell into which a foreign (heterologous, non-native) molecule is or has been introduced. In certain embodiments, the host cell herein is a bacterial cell.  
20 In certain embodiments, the host cell is a gram-negative bacterial cell. In certain embodiments, the host cell is a bacterial cell comprising a secretion system such as a Type III Secretion (T3S) System (T3SS). In certain embodiments, a host cell herein is selected from the group consisting of a *Salmonella*, *Shigella*, *Chlamydia*, *Yersinia*, *Pseudomonas*, and *Escherichia* cell. In certain  
25 embodiments, a host cell herein is selected from the group consisting of a *Salmonella enterica* serovar *Typhimurium*, *Shigella flexneri*, *Chlamydia trachomatis*, *Yersinia pseudotuberculosis*, *Pseudomonas aeruginosa*, and *Escherichia coli* cell.

A "purification tag" as used herein refers to a ligand that aids protein purification with, for example, size exclusion chromatography, ion exchange chromatography, and/or affinity chromatography. Purification tags and their use are well known to the art (see, e.g., Thermo  
30 Scientific Protein Purification Handbook 2010) and may be, for example, poly-histidine, glutathione S-transferase (GST), Myc, HA, FLAG, or maltose binding protein (MBP). A step of

- 13 -

purifying, collecting, obtaining, or isolating a protein may therefore include size exclusion chromatography, ion exchange chromatography, or affinity chromatography. In certain embodiments, a step of purifying a FlgM peptide, or a fusion (chimeric) protein comprising FlgM, utilizes affinity chromatography and, for example, a  $\sigma^{28}$  affinity column or an affinity column comprising an antibody that binds FlgM or another member of the fusion protein. In one embodiment, a step of purifying a fusion protein comprising at least a FlgM peptide and a purification tag utilizes affinity chromatography and, for example, an affinity column that binds the purification tag.

"Cleavage site" as used herein refers to a sequence that is recognized and/or cut by a chemical or protein (a restriction enzyme or protease, for example). Cleavage sites are well known by the art (*see, e.g.*, US PG PUB NO. 2015/0037868, incorporated herein by reference in its entirety). Because cleavage sites may be polynucleotide sequences (e.g., restriction enzyme sites) or amino acid sequences (e.g., peptidase sites), a person with ordinary skill in the art will recognize, based on the context within which it is recited, whether "cleavage site" refers to a polynucleotide sequence or an amino acid sequence. Exemplary cleavage sites include those recognized by a protease selected from the group consisting of an alanine carboxypeptidase, *Armillaria mellea* astacin, bacterial leucyl aminopeptidase, cancer procoagulant, cathepsin B, clostripain, cytosol alanyl aminopeptidase, elastase, endoproteinase Arg-C, enterokinase, gastricsin, gelatinase, Gly-X carboxypeptidase, glycyl endopeptidase, human rhinovirus 3C protease, hypodermin C, Iga-specific serine endopeptidase, leucyl aminopeptidase, leucyl endopeptidase, lysC, lysosomal pro-X carboxypeptidase, lysyl aminopeptidase, methionyl aminopeptidase, myxobacter, nardilysin, pancreatic endopeptidase E, picornain 2A, picornain 3C, proendopeptidase, prolyl aminopeptidase, proprotein convertase I, proprotein convertase II, russellysin, saccharopepsin, semenogelase, T-plasminogen activator, thrombin, tissue kallikrein, tobacco etch virus (TEV), togavirin, tryptophanyl aminopeptidase, U-plasminogen activator, V8, venombin A, venombin AB, and Xaa-pro aminopeptidase. *See* US PG PUB NO. 2015/0037868, incorporated herein by reference in its entirety. In some embodiments, the cleavage site is a protease cleave site. In some embodiments, the cleavage site is a Tobacco Etch Virus (TEV) protease or Enterokinase (ETK) cleavage site. The components of a fusion (chimeric) peptide may be separated (released) via cutting at a cleavage site before, during, or after a step of fusion peptide purification. For example, wherein a fusion peptide comprises a FlgM protein fused to a

- 14 -

heterologous polypeptide and a cleavage site, and the fusion peptide has been secreted into culture media, the FlgM protein and the polypeptide may be separated by the application of a protease to the culture media before a step of purification from culture media or the fusion peptide may be purified from the culture media and then a protease may be applied to the fusion peptide to release the FlgM protein from the polypeptide. In certain embodiments, a step of fusion peptide decomposition (cleavage) is concurrent with a step of purification. For example, a fusion peptide may be passed over an affinity column, the heterologous polypeptide or the purification tag may bind the affinity column and, while bound, a cutting agent may be applied such that the cleavage site is cut and the fusion peptide decomposes.

As is conventional, the designation "NH<sub>2</sub>" or "N-" refers to the N-terminus of an amino acid sequence and the designation "COOH" or "C-" refers to the C-terminus of an amino acid sequence.

The term "naturally-occurring", "native" or "wild-type" as used herein as applied to an object refers to the fact that an object can be found in nature. For example, a polypeptide or polynucleotide sequence that is present in an organism (including viruses) that can be isolated from a source in nature and which has not been intentionally modified by man in the laboratory or otherwise is wild-type.

A "mutation" as used herein refers to a variation in a nucleotide or amino acid sequence as compared to the wild type sequence. A mutation or "alteration" herein is usually a manufactured nucleic acid change within a coding polynucleotide sequence that results in an alternate but synonymous codon (i.e., the encoded amino acid residue does not change as a result of the mutation). The production and identification of such mutant, synonymous codon optimized, sequences are as further described herein.

The term "control sequence" as used herein refers to polynucleotide sequences which are necessary to effect the expression and processing of coding sequences to which they are ligated.

The term "control sequences" is intended to include, at a minimum, all components whose presence is essential for expression and processing, and can also include additional components whose presence is advantageous, for example, leader sequences and fusion partner sequences. A control sequence may be a "Transcription Control Element (TCE)", the nature of which differs depending upon the host organism. A person with ordinary skill in the art will recognize that in prokaryotes, such TCEs generally include promoter, ribosomal binding site, and transcription termination sequences while in eukaryotes, generally, such TCEs include promoters and

- 15 -

transcription termination sequences. Control sequences may be regulatable (inducible, for example) or constitutive.

According to the invention any promoter may be used. Promoter usually refers to the nucleotide sequence upstream (5') to the coding sequence and controls the expression of the coding  
5 sequence by providing the recognition of the RNA polymerase and other factors which are necessary for the correct transcription. The promoter used according to the invention may comprise a minimal promoter which specify the transcription start site to which regulator elements are attached for expression control.

The term "operably linked" or "operatively linked" as used herein refers to the positioning of  
10 components such that they are in a relationship permitting them to function in their intended manner. A person with ordinary skill in the art will recognize that under certain circumstances and depending on the nature of the components (e.g., a cleavage site or purification tag), two or more components "operably linked" together are not necessarily contiguously linked or contiguously associated. A control sequence "operably linked" to a coding sequence is ligated in such a way that  
15 expression of the coding sequence is achieved under conditions compatible with the control sequences. If two polynucleotide sequences are said to be operably linked to a control sequence such as a transcription control element (TCE) (or more than one transcription control element), a person with ordinary skill in the art will recognize that a variety of configurations are functional and encompassed. For example, a person with ordinary skill in the art will recognize that at least  
20 all the following configurations are encompassed: NH<sub>2</sub>-the transcription control element-the first polynucleotide sequence-the second polynucleotide sequence-COOH; NH<sub>2</sub>-the transcription control element-the second polynucleotide sequence-the first polynucleotide sequence-COOH; NH<sub>2</sub>-the first transcription control element-the first polynucleotide sequence-the second transcription control element-the second polynucleotide sequence-COOH; NH<sub>2</sub>-the second  
25 transcription control element-the second polynucleotide sequence-the first transcription control element-the first polynucleotide sequence-COOH; NH<sub>2</sub>-the first transcription control element-the second polynucleotide sequence-the second transcription control element-the first polynucleotide sequence-COOH; NH<sub>2</sub>- the second transcription control element-the first polynucleotide sequence- the first transcription control element-the second polynucleotide sequence- COOH; and  
30 combinations thereof. Further, for example, if a mutant triacontanucleotide sequence is said to be operatively linked to a polynucleotide sequence encoding a polypeptide, a nucleic acid sequence

- 16 -

encoding a purification tag, and a cleavage site, a person with ordinary skill in the art will recognize that a variety of configurations are functional and encompassed. For example, such a person will recognize that at least the following configurations are encompassed (from the 5' to the 3' end): 5'-(the mutant triacontanucleotide sequence)-(the cleavage site)-(the polynucleotide sequence encoding a polypeptide)-3'; 5'-(the nucleic acid sequence encoding a purification tag)-(the mutant triacontanucleotide sequence)-(the cleavage site)-(the polynucleotide sequence encoding a polypeptide)-3'; 5'-(the mutant triacontanucleotide sequence)-(the cleavage site)-(the polynucleotide sequence encoding a polypeptide)-(the nucleic acid sequence encoding a purification tag)-3'; and combinations thereof.

The similarity between two nucleic acid sequences, or two amino acid sequences, is referred to as "sequence identity." Sequence identity is frequently measured in terms of percentage identity (or similarity); the higher the percentage, the more similar the two sequences are. Homologs or orthologs of human polypeptides, and the corresponding cDNA or gene sequence(s), will possess a relatively high degree of sequence identity when aligned using standard methods. This sequence identity will be more significant when the orthologous proteins or genes or cDNAs are derived from species that are more closely related (e.g., human and chimpanzee sequences), compared to species more distantly related (e.g., human and *C. elegans* sequences).

Methods of alignment of sequences for comparison are well known in the art. Various programs and alignment algorithms are described in: Smith & Waterman *Adv. Appl. Math.* 2: 482, 1981; Needleman & Wunsch *J. Mol. Biol.* 48: 443, 1970; Pearson & Lipman *Proc. Natl. Acad. Sci. USA* 85: 2444, 1988; Higgins & Sharp *Gene*, 73: 237-244, 1988; Higgins & Sharp *CABIOS* 5: 151-153, 1989; Corpet et al. *Nuc. Acids Res.* 16, 10881-90, 1988; Huang et al. *Computer Appls. in the Biosciences* 8:155-65, 1992; and Pearson et al. *Meth. Mol. Bio.* 24:307-31, 1994. Altschul et al. *J. Mol. Biol.* 215:403-410, 1990, presents a detailed consideration of sequence alignment methods and homology calculations.

The NCBI Basic Local Alignment Search Tool (BLAST) (Altschul et al. *J. Mol. Biol.* 215:403-410, 1990) is available from several sources, including the National Center for Biotechnology Information (NCBI, Bethesda, Md.) and on the Internet, for use in connection with the sequence analysis programs blastp, blastn, blastx, tblastn and tblastx. By way of example, for comparisons of amino acid sequences of greater than about 30 amino acids, the Blast 2 sequences function is employed using the default BLOSUM62 matrix set to default parameters, (gap

- 17 -

existence cost of 11, and a per residue gap cost of 1). When aligning short peptides (fewer than around 30 amino acids), the alignment is performed using the Blast 2 sequences function, employing the PAM30 matrix set to default parameters (open gap 9, extension gap 1 penalties).

5 An alternative indication that two nucleic acid molecules are closely related is that the two molecules hybridize to each other under stringent conditions. Stringent conditions are sequence-dependent and are different under different environmental parameters. Generally, stringent conditions are selected to be about 5°C to 20°C lower than the thermal melting point (T<sub>m</sub>) for the specific sequence at a defined ionic strength and pH. The T<sub>m</sub> is the temperature (under defined ionic strength and pH) at which 50% of the target sequence remains hybridized to a perfectly  
10 matched probe or complementary strand. Conditions for nucleic acid hybridization and calculation of stringencies can be found in Sambrook et al. (In Molecular Cloning: A Laboratory Manual, CSHL, New York, 1989) and Tijssen (Laboratory Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes Part I, Chapter 2, Elsevier, New York, 1993).

15 Nucleic acid sequences that do not show a high degree of sequence identity may nevertheless encode similar amino acid sequences, due to the degeneracy of the genetic code. It is understood that changes in nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid molecules that all encode substantially the same protein.

The phrase "synonymous codon and codon context effects on translation" may be referred to using the acronym CEOT (codon effects on translation).

20 The phrase "synonymous codon mutant" (SCM) herein refers to a coding sequence that has been mutagenized so as to comprise one or more synonymous codons as compared to the wild type coding sequence. The phrase "SCM *flgM*" refers to a *flgM* coding sequence that has been mutagenized and comprises one or more synonymous codons as a result of that mutagenesis (i.e., "SCM *flgM*" refers to a mutant *flgM* coding sequence). As used herein, "synonymous codon  
25 mutagenesis/alteration" (SCA) refers to the step of mutating (i.e., altering) a coding sequence such that the resulting mutant/alterated sequence comprises one or more synonymous codons as compared to a corresponding control sequence.

#### *Methods of assaying translation speed*

30 The present invention relates to an assay for determining translation speed that takes advantage of the characterization of the regulatory mechanism called transcriptional attenuation

- 18 -

in the histidine biosynthetic (*his*) operon of *Salmonella* (see Chevance *et al.*, PLOS Genetics 10(6):1-14 (2014)). The transcription of the *his* structural genes is dependent on levels of charged histidyl- tRNA in the cell and independent of mRNA or leader peptide stability. The 5'-region of the *his* operon has a 16 amino acid open reading frame with seven consecutive His codons.

5 When charged His-tRNA levels are high, attenuation occurs and RNA polymerase ceases transcription of the *his* operon. If charged His-tRNA is low the ribosome stalls within the stretch of seven His codons and the attenuation mechanism is thwarted so that RNA polymerase continues into the *his* structural genes in response to limiting His-tRNA. A system was developed to measure the rate of translation through individual codons, pairs of codons, triplets, etc., and measure *in vivo* how fast the ribosome can translate through different codons for the same amino acid (so-called silent codons or synonymous codons). In this way, an assay was developed for translation speed and determining how one or more synonymous codons effect translation speed.

15 An exemplary system was produced by replacing the *his* structural genes with the *lac* operon as a reporter for His attenuation. With the *his* attenuation system controlling *lac* operon transcription, this system may take advantage of color indicator media to provide simple, qualitative color-readouts of attenuation. Based on colony-color phenotypes, the levels of *his-lac* constructs that are either attenuated or non-attenuated are determined. Using the *lac* operon reporter system, the His5 position of the leader peptide is substituted with all 63 codons and the effect of individual codons on the degree of ribosomal stalling in the leader peptide region was measured (*Id.*). Levels of de-attenuation that resulted from the stalling of the ribosome due to the codon change at His5 was measured. Stop codons (UAA, UAG and UGA) exhibited the highest levels of de-attenuation while the arginine codons AGA and AGG showed high levels of de-attenuation(*Id.*). Comparing third positions NNU versus NNC that are read by a single tRNA species, translation of NNU showed a higher level of de-attenuation compared to NNC (*Id.*). This is consistent with *in vitro* studies showing that NNG is translated faster than the cognate NNU by a single tRNA species.

25 The use of this translation speedometer assay for measuring the effects of codon context on translation speed was further demonstrated by measuring the effects of each codon in both 5' and 30 3' context with the UCA codon. The UCA codon is particularly useful because it is translated by only one essential tRNA species, the SerT tRNA. *his-lac* expression on Mac-Lac indicator media

- 19 -

for all UCA-NNN and NNN-UCA codon pairs placed at His4-His5 in the *his* operon leader peptide was determined. 22 of the 64 codons exhibited codon context effects. In other words, translation speeds were different for the UCA-NNN and NNN-UCA codon pair for 22 of the 64 codons. The most dramatic context effects were obtained with tyrosine codons. The UCA-  
5 UAU and UCA-UAC constructs exhibited levels of de-attenuation comparable to those that were produced by stop codons. However, in the reverse orientation, UAU-UCA and UAC-UCA, de-attenuation is low. The UCA-UAU and UCA-UAC codon pairs are translated at a slow rate whereas the reverse orientations UAU-UCA and UAC-UCA are translated at fast rates. The codon context effect with the CGU codon is also dramatic. Both UCA-CGU and CGU-UCA  
10 exhibits a fast translation rate, yet CAU(His)-CGU at the His4- His-5 positions is translated at a very slow rate.

#### *Methods of modulating translation speed*

Bacteria comprise a variety of secretion systems with varying structures and functions.  
15 Gram-negative bacteria, for example, comprise at least six different secretion systems (Types I-VI) spanning one or both of the outer and inner membranes and playing roles in environmental response, adhesion, or pathogenicity processes (*see, e.g., Costa et al., Secretion systems in Gram-negative bacteria: structural and mechanistic insights*, Nature Reviews, Microbiology 13:343-359 (2015)). For flagella assembly, for example, bacteria may utilize a secretion system (T3SS, e.g.,)  
20 by which proteins that are required for flagellum structure and assembly are exported from the cytoplasm, pass the outer membrane, and through the growing flagellar structure.

The *Salmonella enterica* flagellum contains a complex motor structure, known as the hook-basal body (HBB) complex, that is embedded in the cell wall and membranes. The HBB is connected to a long, external filament that extends about 10 microns from the cell surface. The  
25 external filament, comprising up to 10,000 FliC subunits, extends from the HBB. Each filament represents about 1% of the total cell protein. The proton motive force powers the rotation of the motor-filament to propel the bacterium through liquid environments and across surfaces. The HBB also contains a type III secretion system that directs the secretion of substrates through a central channel of the structure. Subunits travel to the tip of the growing organelle where they self-  
30 assemble into place.

- 20 -

There are over 60 genes in the flagellar regulon, which includes the genes of the chemosensory system. The transcription of these genes is coupled to flagellum assembly as shown in Figure 1. The flagellar regulon is organized into a transcriptional hierarchy of three promoter classes. The Class 1 operon is often called the master operon and encodes the transcriptional regulators FlhD and FlhC. The FlhD and FlhC proteins form a complex, FlhD<sub>4</sub>C<sub>2</sub>, which directs  $\sigma^{70}$  RNA polymerase-dependent transcription from flagellar Class 2 promoters. Genes expressed from Class 2 promoters encode the proteins necessary for the structure and assembly of the HBB complex. In addition to HBB genes, several regulatory proteins are transcribed from Class 2 promoters; prominent among these are the  $\sigma^{28}$  structural gene, *FliA*, and the anti- $\sigma^{28}$  gene, *FlgM*.  $\sigma^{28}$  and FlgM are regulatory proteins that couple transcription of the flagellar class 3 promoters to completion of the HBB. The  $\sigma^{28}$  protein is a flagellar-specific transcription factor that directs RNA polymerase to transcribe from the flagellar class 3 promoters. Class 3 genes include the structural genes of the flagellar filament and genes of the chemosensory signal transduction system that controls the direction of flagellar rotation according to changing concentrations of extracellular ligands.

During HBB assembly, the flagellar type III secretion (T3S) system is specific for rod-hook secretion substrates. Upon HBB completion, the FlhB component of the T3S apparatus undergoes a conformational change and the secretion specificity switches to Late or Filament-type secretion. Prior to HBB completion, FlgM inhibits  $\sigma^{28}$ -dependent transcription from flagellar Class 3 promoters. FlgM is a late secretion substrate and upon HBB completion and the secretion specificity switch, FlgM is secreted from the cell, releasing  $\sigma^{28}$  to transcribe from Class 3 promoters.

The FlgM protein comprises a secretion signal at its N-terminus that is not cleaved during the secretion process. Substrate secretion through the T3SS is often facilitated by chaperone-assisted delivery of substrates to the secretion apparatus. In this way, FlgM secretion is greatly enhanced by the binding of the secretion chaperone  $\sigma^{28}$  to the C-terminal half of FlgM.

Because FlgM is a small T3S substrate and not part of the final flagellar structure, it can be used as a vehicle to direct secretion of foreign proteins into the periplasm or into the extracellular milieu. The inventors have shown that *FlgM* polynucleotide sequences operably linked to a heterologous coding sequence may be expressed in various organisms and secreted by the flagellar T3S system (see U.S. PG PUB NO. 2015/0225466, incorporated herein by reference in its

- 21 -

entirety). Specifically, a chimeric protein comprising a FlgM peptide fused to a heterologous protein is expressed and secreted through the flagellar T3SS structures and into culture media (see U.S. PG PUB NO. 2015/0225466, incorporated herein by reference in its entirety). The inventors have further shown that expression of a FlgM fusion peptide may be controlled using inducible and/or overexpression control sequences, it may be manipulated by the salt concentration (e.g., NaCl and KCl) within media, and the fusion peptide or its components may be isolated from media using a variety of means (*Id.*).

The examples herein further demonstrate the use of a T3SS for the modulated production and secretion of chimeric proteins into culture media. In one embodiment, the polypeptide is insulin.

As used herein, the term "insulin" means the active principle of the pancreas that affects the metabolism of carbohydrates in the animal body and which is of value in the treatment of diabetes mellitus. The term includes synthetic and biotechnologically derived products that are the same as, or similar to, naturally occurring insulins in structure, use, and intended effect and are of value in the treatment of diabetes mellitus.

The term "insulin" or "insulin molecule" is a generic term that designates the 51 amino acid heterodimer comprising the A-chain peptide having the amino acid sequence shown in SEQ ID NO: 1 and the B-chain peptide having the amino acid sequence shown in SEQ ID NO: 2, wherein the cysteine residues at positions 6 and 11 of the A chain are linked in a disulfide bond, the cysteine residues at position 7 of the A chain and position 7 of the B chain are linked in a disulfide bond, and the cysteine residues at position 20 of the A chain and 19 of the B chain are linked in a disulfide bond. The term "insulin" also includes precursors of insulin molecules.

The term "insulin analog" as used herein includes any heterodimer analog that comprises one or more modification(s) of the native A-chain peptide and/or B-chain peptide. Modifications include but are not limited to substituting an amino acid for the native amino acid at a position selected from A4, A5, A8, A9, A10, A12, A13, A14, A15, A16, A17, A18, A19, A21, B1, B2, B3, B4, B5, B9, B10, B13, B14, B15, B16, B17, B18, B20, B21, B22, B23, B26, B27, B28, B29, and B30; and/or deleting any or all of positions B1-4 and B26-30. Insulin analogs include molecules having one to 10 amino acids at the N or C terminus of the A-chain peptide and/or B-chain peptide. Insulin analogs further include molecules amidated at the C-terminus of the A-chain peptide and/or B-chain peptide. Examples of insulin analogs include but are not limited to the heterodimer analogs disclosed in published international application WO20100080606,

- 22 -

WO2009/099763, and WO2010080609, the disclosures of which are incorporated herein by reference. Insulin glargine (Gly(A21), Arg(B31), Arg(B32)-human insulin), insulin lispro (Lys(B28), Pro(B29)-human insulin, insulin glusiline (Lys(B3), Glu(B29)-human insulin), and insulin detemir (Lys-myristic acid(B29)-human insulin) are examples of commercially available insulin analogs.

The term "insulin analogs" further includes heterodimer polypeptide molecules that have little or no detectable activity at the insulin receptor but which have been modified to include one or more amino acid modifications or substitutions to have an activity at the insulin receptor that has at least 1%, 10%, 50%, 75%, or 90% of the activity at the insulin receptor as compared to native insulin. In particular aspects, the insulin analog is a partial agonist that has from 2x to 100x less activity at the insulin receptor as does native insulin. In other aspects, the insulin analog has enhanced activity at the insulin receptor.

Further described herein are expression systems utilizing synonymous codon mutants within a flagellar secretion system gene and methods of modulating heterologous polynucleotide translation when operably appended to such mutants. The synonymous codon mutagenesis of the present invention differs from traditional, codon optimization that relies on the codon usage (bias) of the host cell (organism). Codon usage bias relies on the relative amount that a particular codon is used by a particular organism. The art has long recognized that, between organisms, one synonymous codon for an amino acid is more prevalent (used with higher frequency) as compared to other synonymous codons for that same amino acid (i.e., an organism "prefers" or has a "bias" toward the use of a particular synonymous codon). Divergent organisms may prefer different synonymous codons and therefore have a different codon usage bias.

A table of codon usage frequency in *Escherichia coli* and *Salmonella enterica* is shown below. Therein, a higher usage frequency value corresponds to a greater bias (preference) toward that synonymous codon (*available at* [WorldWideWeb.sci.sdsu.edu/~smaloy/MicrobialGenetics/topics/rev-sup/wobble.html](http://WorldWideWeb.sci.sdsu.edu/~smaloy/MicrobialGenetics/topics/rev-sup/wobble.html), *last visited* November 15, 2015). *See also* The Codon Bias Database ([WorldWideWeb.homepages.luc.edu/~cputonti/cbdb/genera/shigella.html](http://WorldWideWeb.homepages.luc.edu/~cputonti/cbdb/genera/shigella.html)) and Hilterbrand, *et al.* (*CBDB: The codon bias database*, BMC Bioinformatics 13(62):1-7 (2012)).

30

Codon preference in *E. coli* and *S. typhimurium*.<sup>1</sup>

	U			C			A			G			
<b>U</b>	UUU	Phe	19	UCU	Ser	9	UAU	Tyr	15	UGU	Cys	4	<b>U</b>
	UUC	Phe	17	UCC	Ser	10	UAC	Tyr	12	UGC	Cys	6	<b>C</b>
	UUA	Leu	11	UCA	Ser	6	UAA	STOP	2	UGA	STOP	0.8	<b>A</b>
	UUG	Leu	12	UCG	Ser	8	UAG	STOP	0.2	UGG	Trp	12	<b>G</b>
<b>C</b>	CUU	Leu	10	CCU	Pro	7	CAU	His	11	CGU	Arg	23	<b>U</b>
	CUC	Leu	10	COC	Pro	5	CAC	His	10	CGC	Arg	23	<b>C</b>
	CUA	Leu	4	CCA	Pro	7	CAA	Gln	13	CGA	Arg	3	<b>A</b>
	CUG	Leu	55	CCG	Pro	15	CAG	Gln	31	CGG	Arg	5	<b>G</b>
<b>A</b>	AUU	Ile	27	ACU	Thr	9	AAU	Asn	17	AGU	Ser	7	<b>U</b>
	AUC	Ile	27	ACC	Thr	25	AAC	Asn	24	AGC	Ser	16	<b>C</b>
	AUA	Ile	4	ACA	Thr	6	AAA	Lys	36	AGA	Arg	2	<b>A</b>
	AUG	Met	26	ACG	Thr	15	AAG	Lys	12	AGG	Arg	1	<b>G</b>
<b>G</b>	GUU	Val	17	GCU	Ala	16	GAU	Asp	33	GGU	Gly	24	<b>U</b>
	GUC	Val	16	GCC	Ala	25	GAC	Asp	22	GGC	Gly	33	<b>C</b>
	GUA	Val	12	GCA	Ala	16	GAA	Glu	43	GGA	Gly	6	<b>A</b>
	GUG	Val	26	GCG	Ala	37	GAG	Glu	20	GGG	Gly	10	<b>G</b>

<sup>1</sup> The numbers represent the average frequency of codon usage per 1000 codons based upon the DNA sequence of about 450,000 genes in *E. coli* and *S. typhimurium*. Some additional codon frequencies can be found in Miller (1992).

Generally, the presence of a host cell's preferred synonymous codons for, for example, the expression of a heterologous polynucleotide sequence within that host cell results in increased translation efficiency. Traditional codon optimization techniques utilize codon usage bias by mutagenizing a coding sequence such that the host cell's preferred codons are present to increase translation efficiency or absent to decrease translation efficiency.

Synonymous codon mutations in the *fliA*, *fliC* and *flgM* genes of the *Salmonella* Type III Secretion System have been isolated that either increase or decrease gene expression. For *fliA*, a synonymous codon 13 (Asp) change from GAU to GAC resulted in a ~2-fold increase in  $\sigma^{28}$  protein produced (Barker, C. S., et al., *Assembling Flagella in Salmonella Mutant Strains Producing a Type III Export Apparatus without FliO*, J. Bacteriol. 196(23):4001-4011 (2014)). Similarly, for the *fliC* gene of the *Salmonella* T3SS, a synonymous codon 13 (Leu) change from UUG to CUG produced ~2- fold more FliC protein, whereas synonymous codon changes at codon 14 (Thr) from ACC to either ACA or ACG produced ~2-fold and ~1.5- fold more FliC protein,

- 24 -

respectively (Rosu, V., *et al.*, *Translation Inhibition of the Salmonella fliC Gene by the FliC 5' Untranslated Region, fliC Coding Sequences, and FlgM*, J. Bacteriol. 188(12):4497-4507 (2006)).

Based on codon usage bias as depicted within the table above, for example, the preference of *Salmonella* for the Asp GAU codon (33) over GAC (22) renders the at least about 2-fold increase in protein production using the GAC codon surprising (Barker *et al.*, *supra*). Likewise, the preference for the Thr ACC codon (25) over the ACA (6) and ACG (15) codons renders the at least about 2-fold and 1.5-fold increase in protein production using the ACA and ACG codons, respectively, surprising (Rosu *et al.*, *supra*).

## Examples

### 10 *Bacterial Strains and Media*

The commonly used non- pathogenic strain of *Salmonella enterica* serovar Typhimurium LT2 was used for all experiments. LT2 harbors mutations in the *ropS* and *mviA* genes that render it avirulent (Yarus, *Rates of aminoacyl-tRNA selection at 29 sense codons in vivo*, J. Mol. Biol. 209: 65–77 (1989); Hughes & Roth, *Transitory cis complementation: a method for providing transposition functions to defective transposons*, Genetics 119: 9–12 (1988); Johnston & Roth, *Genetic analysis of the histidine operon control region of Salmonella typhimurium*, J. Mol. Biol. 145: 713–734 (1981)). The *mviA* gene encodes for the regulator of a two-component system involved in *Salmonella* virulence. The *rpoS* gene product (stationary-phase sigma transcription factor) is required for the expression of multiple virulence factors. The two mutations within the virulence pathway of *Salmonella* result in attenuation of virulence.

Cells were cultured in Luria-Bertani (LB) medium and, when necessary, supplemented with ampicillin (100 µg/milliliter) or tetracycline (15 µg/milliliter). The generalized transducing phage of *S. Typhimurium* P22 *HT105/1 int-201* was used in all transductional crosses (Johnston & Roth, *DNA sequence changes of mutations altering attenuation control of the histidine operon of Salmonella typhimurium*, J. Mol. Biol. 145: 735–756 (1981)).

### *Strain Construction*

Targeted chromosomal mutagenesis was carried out via the tetRA insertion and replacement with the L-Red recombinase system as described in Datsenko and Wanner (*One-step inactivation*

- 25 -

*of chromosomal genes in Escherichia coli K-12 using PCR products*, PNAS 97: 6640– 6645 (2000)).

All Primers were synthesized by Integrated DNA Technologies (Coralville, IA). All PCR reactions were performed using a proof reading polymerase (Accuprime Pfx, Invitrogen or  
5 Phusion, Fermenta). Recombinant products mediated by L-Red were PCR-checked using Taq DNA polymerase and further sequenced.

#### *Screen for translation-slowest codon pairs*

Strains containing slow pair or fast pairs were constructed using by L-Red recombination as  
10 described above. After electroporation and plating on anhydrotetracycline plates, the plates were replica printed onto MacConkey-lactose and tetrazolium-lactose indicator plates. White colonies on tetrazolium plates (Lac<sup>+</sup>) were isolated and sent for DNA sequence analysis. The T-N-N codons were specifically avoided to prevent isolation of stop codons. Tz-Lac screening was used to identify translation-slow and translation-fast codon pairs in the His Leader system.

#### *b-Galactosidase assays*

Thirty microliters of an overnight culture were subcultured into 3 ml of fresh LB medium. Tubes were incubated with shaking at 37°C until the contents reached a mid-log-phase density of  
15 OD 0.4. Cultures were put on ice, spun down, and resuspended in 3 ml of cold-buffered saline. Culture samples of 0.5 ml (diluted if necessary) were added to 0.55 ml of complete Z-buffer (Z-buffer plus 5 ml of 10% sodium dodecyl sulfate and 100 ml of chloroform) (Stanley R. Maloy, EXPERIMENTAL TECHNIQUES IN BACTERIAL GENETICS. (Jones and Bartlett Publishers, Inc. 1990)). The assay was continued as described previously (*Id.*). For each strain, assays were performed for three independent biological replicates.

#### *Identification of Synonymous Codon Pairs in a coding sequence associated with a disease or condition*

A search for synonymous codon mutations within a polynucleotide sequence (e.g., a coding  
30 sequence) associated with a disease or condition is performed using, for example, a database. The synonymous mutations are independent of any additional mutations (including insertions and deletions). Synonymous mutations are parsed based on the 15 base pair sequences flanking either

- 26 -

side. The effect a synonymous codon mutation has on the expression of the polynucleotide sequence (expression of an encoded protein, for example) is validated by introducing the synonymous codon mutagenized/alterd (SCA) polynucleotide sequence into a host cell and culturing the host cell in conditions sufficient for polynucleotide expression. Any effect is  
5 determined by comparison to wild-type polynucleotide expression within a comparable control cell. Resulting protein production is determined by SDS-PAGE and Western blotting.

#### *SDS-PAGE and Western blotting*

Expressed chimeric proteins comprising FlgM (and optionally a poly-Histidine purification tag) or polypeptides expressed from a synonymous codon mutagenized polynucleotide sequence  
10 are retrieved from whole-cell lysate or cultural supernatant and subjected to SDS polyacrylamide gel electrophoresis and then analyzed by immunoblotting using polyclonal anti-FlgM antibodies, monoclonal anti-6xHis (rabbit) antibodies, or anti-polypeptide antibodies for detection. Antigen-antibody complexes are visualized by chemiluminescent or infrared detection using the LI-COR  
15 Odyssey or Bio-Rad ChemiDoc imaging system. For chemiluminescent development, secondary goat anti-rabbit antibodies (Bio-Rad) conjugated with horseradish peroxidase (HRP) and an ECL detection kit (Amersham Biosciences) are used. For infrared detection, secondary anti-rabbit IRDye690 (LI-COR) are used. Densitometric measurements of protein bands are performed using  
20 ImageJ 1.45s for Mac OS X (Abramoff *et al.*, *Image processing with ImageJ*, Biophotonics Int. 11:36–42(2004)).

#### *Recombinant expression and purification of protein fusions*

Cells expressing optimized fusion proteins are picked from a fresh single colony and grown in 10 ml LB overnight. The overnight cultures are diluted 1:100 into 1 liter of fresh medium in a  
25 baffled flask and grown in a shaking incubator at 200 rpm for 6-12 hours. If appropriate, expression is induced after the first 2 hours by the addition of 0.2% arabinose or Na-Salicylate. Cells are pelleted by centrifugation (7,000 rpm), and the supernatant containing the recombinant polypeptide of interest is passed through a 0.22- $\mu$ m polyethersulfone filter (Corning, NY), a low-protein-binding membrane for removal of residual bacteria. For further purification, a gravity flow  
30 column (Bio-Rad) packed with 3 g Ni-IDA resin (Protino Ni-IDA; Machery- Nagel) is used, and

- 27 -

affinity-tagged proteins are eluted under native conditions at pH 7.5 with a buffer containing 250 mM imidazole.

#### *Secretion assay*

5 Overnight cultures are diluted 1:100 in LB and grown for 2 h at 37°C before inducing the expression of the respective FlgM fusions by adding 0.2% L-arabinose. Cells are kept at 37°C for an additional 4-12 hours, while the fusion proteins are expressed. Afterwards, the optical density at 600 nm (OD<sub>600</sub>) is determined for all strains.

10 Two-milliliter aliquots of the resulting cell culture are centrifuged for 10 min at 4°C and 7,000 rpm to obtain, for each aliquot, a pellet and supernatant. The supernatant is filtered through a low-protein-binding filter with a 0.2- $\mu$ m pore size (Acrodisk syringe filter; Pall Life Sciences) to remove the remaining cells. Alternatively, aliquots are centrifuged twice at maximum speed to remove residual cells. Secreted proteins in the filtered or twice-centrifuged supernatant are precipitated by the addition of TCA (10% final concentration). The supernatant samples are  
15 resuspended in 2 $\mu$ L SDS sample buffer (100 mM Tris [pH 6.8], 4% SDS, 10% glycerol, 2%B-mercaptoethanol, 25 mM EDTA, 0.04% bromophenol blue) and adjusted to 20 OD<sub>600</sub> units/ $\mu$ L. The cellular pellet fraction is suspended in 2% SDS sample buffer, whose volume is adjusted to yield 20 OD<sub>600</sub> units/ $\mu$ L.

#### 20 ***Example 1: Effects of synonymous codon mutants on translation***

An allele of the SerT tRNA that is defective in translating the UCA for amino acid Ser7 of the flagellar anti- $\sigma^{28}$  gene, *flgM* was isolated (Chevance, F. F., *et al.*, *J Bacteriol* 188:297-304 (2006)). Despite there being many genes with UCA codons for serine in the *Salmonella* flagellar system, the *serT* tRNA mutant allele only effected *flgM* translation. Without being  
25 bound by theory, it is believed that the effect of the *serT* mutant allele on translation of the UCA (Ser7) codon of *flgM* is a codon-context effect.

Transcription from a  $\sigma^{28}$ -dependent promoter is inhibited by FlgM. Using the bacterial luciferase operon (*lux*) as a reporter, the anti- $\sigma^{28}$  activity of FlgM was measured. When FlgM activity is high,  $\sigma^{28}$ -dependent transcription of P<sub>*motA*</sub>-*lux* is low. Synonymous mutations were  
30 introduced adjacent to position Ser7 and the effects of flanking synonymous mutations on *flgM* mRNA translation was determined.

- 28 -

All 16 combinations of synonymous changes of the Thr6 and Pro8 codons flanking the UCA Ser7 codon of *flgM* were made. The FlgM inhibition of  $\sigma^{28}$ -dependent  $P_{motA-lux}$  transcription was measured. Results were as shown at Figure 2.

The synonymous changes at Thr6 to ACU, ACA or ACG produce an about 11- to 13-fold increase in FlgM inhibitory activity. The synonymous Pro8 codon change CCU to CCG results in an about 20-fold reduction in FlgM inhibitory activity. The combination of a ACC to ACU change at Thr6 with a CCU to CCG change at Pro8 results in an about 34-fold increase in FlgM inhibitory activity. These results demonstrate that the efficiency of codon translation is influenced by up to two adjacent codons (see Figure 2).

***Example 2: Modulation of polynucleotide translation using synonymous mutagenesis of FlgM***

Starting with a *flgM* sequence comprising the CCU to CCG change at Pro8, the region of the *flgM* gene that encodes amino acids 2 through 25 was modified using doped oligonucleotide mutagenesis designed so that, on average, each oligonucleotide had a synonymous codon change for amino acids 2 through 25. Results were as shown in Figure 3.

Only synonymous changes in the two codons preceding Pro8 and the Leu9 codon following Pro8 CCG resulted in restoration to wild- type FlgM inhibitory activity. These results demonstrate that efficiency of translation of a specific codon is dependent on the flanking two codons.

Synonymous variations coding for amino acids Thr6-Ser7-Pro8-Leu9 of FlgM result in an about ~2-fold range in FlgM protein levels. However, since FlgM is a regulatory protein that inhibits  $\sigma^{28}$ -dependent *fliC* transcription, the about 2-fold range in FlgM protein levels produces more than an about 1000-fold range in *fliC* transcription activity. For example, a synonymous codon change at codon 8 of *flgM* from CCU to CCG (encoding Proline) resulted in an about 2-fold lower FlgM protein level. The about 2-fold reduction in FlgM protein levels resulted in an about 20-fold increase in transcription of the  $\sigma^{28}$ -dependent *fliC* promoter.

***Example 3: Modulation of heterologous polynucleotide translation using synonymous mutagenesis of FlgM***

- 29 -

The first ten codons of *lacZ* were replaced with the first ten codons of *flgM*, and the *flgM* sequence was mutated to comprise a CCU to CCG synonymous codon change at Pro8 (codon number 8, in the 5' to 3' direction, of *flgM* that encodes proline). Consistent with the results above, expression of the heterologous *lacZ* sequence under the control of the context-codon-optimized *flgM* sequence resulted in reduced  $\beta$ -galactosidase activity.  $\beta$ -galactosidase activity was reduced about 210 to 150 units.

These results demonstrate that synonymous changes in one or more of Thr6-Ser7-Pro8-Leu9 of *FlgM* effect heterologous protein production (here LacZ) levels when a polynucleotide sequence encoding that polypeptide (here *lacZ*) is operably linked to the context-codon-optimized *FlgM* sequence (SCM *FlgM*) and expressed and secreted through a bacterial secretion system (here an engineered T3SS).

Exemplary polypeptides that may be recombinantly produced using a secretion system of the present invention are listed in Table 1 below.

Table 1:

Protein	Exemplary sequence information
Insulin (human)	A chain: GIVEQCCTSICSLYQLENYCN B chain: FVNQHLCGSHLVEALYLVCGERGFFYTPKPT
Insulin glargine (human)	A chain: GIVEQCCTSICSLYQLENYCG B chain: FVNQHLCGSHLVEALYLVCGERGFFYTPKTRR
Insulin pre-proinsulin (human)	See UniProtKB Accession No. P01308 and GenBank Accession No. CAA49913 (Chekhranova et al., Mol. Biol. 26(3): 596-600 (1992)):  MALWMRLRLPLLALLALWGPDPAAAFVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN  An exemplary mRNA sequence is available at GenBank Accession No. X70508 (Chekhranova et al., Mol. Biol. 26(3): 596-600 (1992)).
Insulin lisopro	A chain: GIVEQCCTSICSLYQLENYCN B chain: FVNQHLCGSHLVEALYLVCGERGFFYTKPT
MaSp1 (Major Ampullate Spidroin 1; spider silk protein) (including,	See Gaines and Marcotte, Insect Mol. Biol. 17(5): 465-474 (2008); US Pat. No. 8,642,734 (incorporated herein by reference in its entirety); US Pat. No. 7,521,228 (incorporated herein by reference in its entirety); US PG PUB NO. 2014/0093965 (incorporated herein by reference in its entirety).

e.g., MaSp1A or MaSp1B from <i>Nephila</i> spiders)	
MaSp2 (Major Ampullate Spidroin 2; spider silk protein)	US Pat. No. 8,642,734 (incorporated herein by reference in its entirety); US Pat. No. 7,521,228 (incorporated herein by reference in its entirety); US PG PUB NO. 2014/0093965 (incorporated herein by reference in its entirety) incorporated herein by reference in.

**Example 4: Modulation of heterologous protein production using a T3SS**

A nucleic acid molecule comprising a full-length, wild type *FlgM* polynucleotide sequence operably linked to a human insulin glargine nucleic acid sequence, a poly-histidine nucleotide  
5 sequence, an enterokinase cleavage site, and a ParaBAD promoter was introduced into and expressed within a *Salmonella enterica* serovar *Typhimurium* cell. The *Salmonella* cell was modified within several T3S genes as previously described by the inventors (see U.S. PG PUB NO. 2015/0225466, incorporated herein by reference in its entirety) and specifically comprised:  
10 *ParaBAD1::flgM*-His6-ETK-insulin glargine  $\Delta flgMN7753 \Delta flgKL7770$  PflhDC7793 *fliA*\*5225  $\Delta fliB$ -T7771 *fliB*<sub>enx</sub> vh2 (referred to as Strain TS1). A FlgM::Insulin chimeric protein was secreted by the cell and detected via Western blot using anti-6xHis antibodies. Results are shown in Figure 4.

**Example 5: Modulation of heterologous protein production using a T3SS and synonymous mutagenesis**

Secretion of human insulin from an engineered *Salmonella* T3SS may be further optimized by introducing and expressing within a *Salmonella* cell a *FlgM* nucleotide sequence (either a wild  
20 type or a synonymous codon mutated *FlgM* sequence) that is operably linked to a poly-histidine purification tag nucleic acid sequence, a cleavage site (e.g., an enterokinase cleave site), and polynucleotide sequences encoding the A and B chains of human insulin wherein the human insulin B chain sequence comprises a synonymous codon mutation. For example, the two adjacent arginine residues within the human insulin B chain may be context codon optimized to increase  
25 insulin expression and secretion through the T3SS. In particular, wild type adjacent arginine codons CGC and CGG (in the 5' to 3' direction) may be mutated to CGU and CGU; CGU and CGC; CGC and CGU; CGC and CGC; or CGG and CGC. An exemplary construct comprising

- 31 -

nucleic acid sequences encoding, in the N-terminal to C-terminal direction, a poly-histidine purification tag, a cleavage site, and human insulin is provided by SEQ ID NO: 16. The sequence provided by SEQ ID NO: 16 may be appended to a FlgM sequence as described elsewhere herein and by the inventors at U.S. PG PUB NO. 2015/0225466 (incorporated herein by reference in its entirety). It is expected that the resulting pair of synonymous codon optimized arginine codons will cause increased human insulin expression and secretion. A chimeric protein comprising FlgM, a poly-histidine purification tag, a cleavage site, and insulin may be detected and purified as described above. The insulin protein may be isolated from the chimeric protein as described above.

10 Table 2:

Sequence Identifier	Description of sequence	Sequence (written in the 5' to 3' direction or N-terminus to C-terminus direction)
1	The first ten amino acid residues of wild-type FlgM protein (corresponds to the first 10 residues of SEQ ID NO: 3).	MSIDRTSPLK
2	The first ten codons (i.e., 30 nucleotides) of the wild-type <i>FlgM</i> mRNA sequence.	ATGAGCATTGACCGTACCTCACCTTTGAAA
3	The full-length, 97 residue wild-type FlgM amino acid sequence (Gillen and Hughes, J. Bacteriology 173(20): 6453-6459 (1991); publically available at, e.g., GenBank Accession No. AAA27075 and UniProtKB Accession No. P26477). The first ten residues (i.e., SEQ ID NO: 1) underlined.	<u>MSIDRTSPLK</u> PVSTVQTRETS <del>DT</del> PVQKTRQE KTSAAATSASVTLSDAQAKLMQPGVSDINMER VEALKTAIRNGELKMDTGKIADSLIREAQSY LQSK
4	Wild type <i>FlgM</i> gene sequence (Gillen and Hughes, J. Bacteriology 173(20): 6453-6459 (1991); publically available at, e.g., GenBank Accession No. M74222). Coding region (i.e., SEQ ID NO: 5) underlined. First ten codons thereof (i.e., SEQ ID NO: 2) in bold.	AATATTCTTATTAACCTATAATTGTGTAAAGATT TTGTCGCGGCTGCCGATGAGATATTCAACCATGATG GTAGCTGGCCGCTACAACGTAACCCCTCGATGAGG ATAAATAA <b>ATGAGCATTGACCGTACCTCACCTTT</b> <b>GA</b> <u>AACCCGTTAGCACTGTCCAGACGCGGAAACCAG</u> <u>CGACACGCCGGTACAAAAACGCGTCAGGAAAA</u> AC <u>GTCCGCCGCGACGAGCGCCAGCGTAACGTTAAGC</u> <u>GACGCGCAAGCGAAGCTCATGCAGCCAGGCGTCA</u> GC GACATTAATATGGAACGCGTCGAAGCATTAAAAA

		<p>CGGCTATCCGTAACGGTGAGTTAAAAATGGATAC  GG  GAAAAATAGCAGACTCGCTCATTTCGCGAGGCGCA  GAGCTACTTACAGAGTAAATAAGCGTATGACTCG  TT  TGTCAGAAATACTTGACCAGATGACCACCGTCCT  GAATGACCTGAAGACGGTGATGGACGCCGAGCAA  CA  ACAGCTTTCCGTAGGCCAGATTAACGGCAGCCAG  CTACAGCGTATTACAGAAGAAAAAAGCTCGTTGC  TG  GCGACGCTGGATTATCTGGAACAACAGCGCCGTC  TGGAGCAGAATGC</p>
5	Full-length, wild type <i>FlgM</i> coding region and stop codon (corresponds to nucleotides 113 to 406 of SEQ ID NO: 4, stop codon underlined).	<p>ATGAGCATTGACCGTACCTCACCTTTGAAAC  CCGTTAGCACTGTCCAGACGCGGAAACCAG  CGACACGCCGGTACAAAAACGCGTCAGGAA  AAAACGTCCGCCGCGACGAGCGCCAGCGTAA  CGTTAAGCGACGCGCAAGCGAAGCTCATGCA  GCCAGGCGTCAGCGACATTAATATGGAACGC  GTCGAAGCATTAAAAACGGCTATCCGTAACG  GTGAGTTAAAAATGGATACGGGAAAAATAGC  AGACTCGCTCATTTCGCGAGGCGCAGAGCTAC  TTACAGAGTAAATA<u>AA</u></p>
6	Full-length, wild type <i>FlgM</i> mRNA (first ten codons (i.e., mRNA sequence corresponding to SEQ ID NO: 2) in bold, stop codon underlined).	<p><b>AUGAGCAUUGACCGUACCUCACCUUUGAAAC</b>  CCGUUAGCACUGUCCAGACGCGGAAACCAG  CGACACGCCGGUACAAAAACGCGUCAGGAA  AAAACGUCCGCCGCGACGAGCGCCAGCGUAA  CGUUAAAGCGACGCGCAAGCGAAGCUCAUGCA  GCCAGGCGUCAGCGACAUUAAUUGGAACGC  GUCGAAGCAUUAAAAACGGCUAUCCGUAAACG  GUGAGUUAAAAUUGGAUACGGGAAAAUAGC  AGACUCGCUCAUUCGCGAGGCGCAGAGCUAC  UUACAGAGUAA<u>UAA</u></p>
7	AGCSer2AGT synonymous codon mutant triacontanucleotide <i>flgM</i> sequence. Mutation underlined.	ATGAGTATTGACCGTACCTCACCTTTGAAA
8	ACCThr6ACA synonymous codon mutant triacontanucleotide <i>flgM</i> sequence. Mutation underlined.	ATGAGCATTGACCGT <u>ACAT</u> CACCTTTGAAA
9	ACCThr6ACG synonymous codon mutant triacontanucleotide <i>flgM</i> sequence. Mutation underlined.	ATGAGCATTGACCGT <u>ACGT</u> CACCTTTGAAA
10	ACCThr6ACT synonymous codon mutant triacontanucleotide <i>flgM</i> sequence. Mutation underlined.	ATGAGCATTGACCGT <u>ACTT</u> CACCTTTGAAA
11	TCASer7TCG synonymous codon	ATGAGCATTGACCGTACCT <u>TCGC</u> TTTGAAA

	mutant triacontanucleotide <i>flgM</i> sequence. Mutation underlined.	
12	TCASer7TCC synonymous codon mutant triacontanucleotide <i>flgM</i> sequence. Mutation underlined.	ATGAGCATTGACCGTACCT <u>CCC</u> CTTTGAAA
13	TCASer7TCT synonymous codon mutant triacontanucleotide <i>flgM</i> sequence. Mutation underlined.	ATGAGCATTGACCGTACCT <u>CTC</u> CTTTGAAA
14	CCTPro8CCG synonymous codon mutant triacontanucleotide <i>flgM</i> sequence. Mutation underlined.	ATGAGCATTGACCGTACCTC <u>ACC</u> GTTGAAA
15	TTGLeu9TTA synonymous codon mutant triacontanucleotide <i>flgM</i> sequence. Mutation underlined.	ATGAGCATTGACCGTACCTCACCT <u>TTA</u> AAA
16	Human insulin glargine construct comprising a poly-histidine purification tag sequence (in bold), a cleavage site (in underline), and pre-pro insulin sequence (B chain in bold and underline, A chain in bold and double underline).	<u>ATG</u> <b>CAT</b> <b>CAT</b> <b>CAT</b> <b>CAT</b> <b>CAT</b> <b>CAT</b> <u>GGT</u> <u>GGC</u> <u>CGC</u> <u>TTT</u> <u>GTG</u> <u>AAC</u> <u>CAA</u> <u>CAC</u> <u>CTG</u> <u>TGC</u> <u>GGC</u> <u>TCA</u> <u>CAC</u> <u>CTG</u> <u>GTG</u> <u>GAA</u> <u>GCT</u> <u>CTC</u> <u>TAC</u> <u>CTA</u> <u>GTG</u> <u>TGC</u> <u>GGG</u> <u>GAA</u> <u>CGA</u> <u>GGC</u> <u>TTC</u> <u>TTC</u> <u>TAC</u> <u>ACA</u> <u>CCC</u> <u>AAG</u> <u>ACC</u> <u>CGC</u> <u>CGG</u> <u>GAG</u> <u>GCA</u> <u>GAG</u> <u>GAC</u> <u>CTG</u> <u>CAG</u> <u>GTG</u> <u>GGG</u> <u>CAG</u> <u>GTG</u> <u>GAG</u> <u>CTG</u> <u>GGC</u> <u>GGG</u> <u>GGC</u> <u>CCT</u> <u>GGT</u> <u>GCA</u> <u>GGC</u> <u>AGC</u> <u>CTG</u> <u>CAG</u> <u>CCC</u> <u>TTG</u> <u>GCC</u> <u>CTG</u> <u>GAG</u> <u>GGG</u> <u>TCT</u> <u>CTG</u> <u>CAG</u> <u>GCG</u> <u>CGT</u> <u>GGC</u> <u>ATT</u> <u>GTG</u> <u>GAA</u> <u>CAA</u> <u>TGC</u> <u>TGT</u> <u>ACC</u> <u>AGC</u> <u>ATC</u> <u>TGC</u> <u>TCC</u> <u>CTC</u> <u>TAC</u> <u>CAG</u> <u>CTG</u> <u>GAG</u> <u>AAC</u> <u>TAC</u> <u>TGC</u> <u>GGC</u> <u>TAG</u>

\* \* \*

5 It is to be appreciated that the Detailed Description section, and not the Summary and Abstract sections, is intended to be used to interpret the claims. The Summary and Abstract sections sets forth one or more, but not all, exemplary embodiments of the present invention as contemplated by the inventor(s), and thus, are not intended to limit the present invention and the appended claims in any way. All of the patents, patent applications and references cited herein are incorporated by reference in their entirety.

10 The present invention has been described above with the aid of functional building blocks illustrating the implementation of specified functions and relationships thereof. The boundaries of these functional building blocks have been arbitrarily defined herein for the convenience of the description. Alternate boundaries can be defined so long as the specified functions and relationships thereof are appropriately performed.

- 34 -

## WHAT IS CLAIMED IS:

1. A recombinant nucleic acid molecule comprising a mutant triacontanucleotide sequence having at least ten codons, wherein at least one of the sixth, seventh, eighth, ninth and tenth codons, in the 5' to 3' direction, are synonymous codons.
- 5 2. A recombinant nucleic acid molecule consisting of a mutant triacontanucleotide sequence having at least ten codons, wherein at least one of the sixth, seventh, eighth, and ninth codons, in the 5' to 3' direction, are synonymous codons.
3. The recombinant molecule of claim 1, wherein the mutant triacontanucleotide sequence encodes the amino acid sequence NH<sub>2</sub>-MSIDRTSPLK-COOH (SEQ ID NO:1).
- 10 4. The recombinant molecule of claim 1, wherein the mutant triacontanucleotide sequence is operably linked to a heterologous polynucleotide sequence that encodes a polypeptide of interest.
5. The recombinant molecule of claim 4, wherein the polypeptide of interest is selected from the polypeptides listed in Table 1.
- 15 6. The recombinant molecule of claim 4, wherein the polypeptide of interest is insulin, spider silk protein MaSp1, or spider silk protein MaSp2.
7. The recombinant molecule of claim 4, wherein the mutant triacontanucleotide sequence is operably linked, in the 5' to 3' direction, to a polynucleotide sequence.
- 20 8. The recombinant molecule of claim 4, wherein the mutant triacontanucleotide sequence replaces the ten amino-terminal-most codons of the polynucleotide sequence and wherein the heterologous polynucleotide sequence encodes the polypeptide of interest.

- 35 -

9. The recombinant molecule of claim 1, wherein the mutant triacontanucleotide sequence encodes SEQ ID NO: 3.
10. The recombinant molecule of claim 1 or claim 4, wherein the mutant triacontanucleotide sequence is a ribonucleic acid sequence.
- 5 11. The recombinant molecule of claim 1 or claim 4, wherein the mutant triacontanucleotide sequence is a deoxyribonucleic acid sequence.
12. The recombinant molecule of claim 1 or claim 4, wherein the polynucleotide sequence is a ribonucleic acid sequence.
- 10 13. The recombinant molecule of claim 1 or claim 4, wherein the polynucleotide sequence is a deoxyribonucleic acid sequence.
14. The recombinant molecule of claim 1 or claim 4, wherein the mutant triacontanucleotide sequence is a synthetic sequence.
- 15 15. The recombinant molecule of claim 4, wherein one or both of the mutant triacontanucleotide sequence and polynucleotide sequence is a synthetic sequence.
16. The recombinant molecule of claim 1, wherein the sixth and eighth codons, in the 5' to 3' direction, are synonymous codons.
17. A vector comprising the recombinant molecule of any one of claims 1 to 16.
- 20 18. The vector of claim 17, wherein the vector is an expression vector and the mutant triacontanucleotide sequence and polynucleotide sequence are operably linked to a heterologous promoter.

- 36 -

19. The vector of claim 18, wherein the heterologous promoter is an inducible promoter, a constitutive promoter, or a tissue specific promoter.
20. The vector of claim 19, wherein the inducible promoter is the *ParaBAD* or *Psal* promoter.
- 5 21. A host cell comprising the recombinant molecule of any one of claims 1 to 16.
22. The host cell of claim 21, wherein the polynucleotide sequence has been codon optimized for expression within the host cell.
23. The host cell of claim 21, wherein the host cell is a *Salmonella* or *Escherichia coli* cell.
24. The host cell of claim 23, wherein the host cell is a *Salmonella enterica* cell.
- 10 25. A host cell comprising the vector of any one of claims 17-19.
26. The host cell of claim 25, wherein the host cell is a *Salmonella* or *Escherichia coli* cell.
27. The host cell of claim 26, wherein the host cell is a *Salmonella enterica* cell.
- 15 28. A recombinant nucleic acid molecule having a modulated translation speed, comprising the sequence 5'- AUGAGCAUUGACCGUACCUCACCUUUGAAA-3' (SEQ ID NO:2), wherein at least one of the sixth, seventh, eighth, ninth and tenth codons, in the 5' to 3' direction, are synonymous codons.
29. The recombinant molecule of claim 28, wherein the sixth codon is the synonymous codon ACU, ACA, or ACG and wherein the recombinant molecule has increased translation speed as compared to a molecule comprising the unmodified sequence.

- 37 -

30. The recombinant molecule of claim 28, wherein the sixth codon is a synonymous codon and the molecule has between about a 9- to 15-fold increase in translation speed as compared to a molecule comprising the unmodified sequence.
31. The recombinant molecule of claim 28, wherein the eighth codon is the synonymous codon CCG and wherein the recombinant molecule has decreased translation speed as compared to a molecule comprising the unmodified sequence.
32. The recombinant molecule of claim 28, wherein the eighth codon is a synonymous codon and the recombinant molecule has between about a .03- to .07-fold decrease in translation speed as compared to a molecule comprising the unmodified sequence.
33. The recombinant molecule of claim 28, wherein the sixth codon is the synonymous codon ACU, the eighth codon is the synonymous codon CCG, and wherein the recombinant molecule has increased translation speed as compared to a molecule comprising the unmodified sequence.
34. The recombinant molecule of claim 28, wherein the sixth codon is a synonymous codon, the eighth codon is a synonymous codon, and wherein the recombinant molecule has between about a 27.5- to 40.5-fold increase in translation speed as compared to a molecule comprising the unmodified sequence.
35. An isolated polynucleotide comprising the recombinant molecule of any one of claims 28-34 operably linked to a nucleic acid sequence that encodes a polypeptide of interest.
36. A vector comprising the recombinant molecule or polynucleotide of any one of claims 28-35.
37. A host cell comprising the recombinant molecule or polynucleotide of any one of claims 28-35.
38. The host cell of claim 37, wherein the host cell is a *Salmonella* or *Escherichia coli* cell.

- 38 -

39. The host cell of claim 37, wherein the host cell is a *Salmonella enterica* cell.
40. A method of modulating protein production within a host cell, comprising culturing a host cell under conditions sufficient for protein expression, wherein the recombinant molecule of any claims 4-9, 12, 13, and 15 has been stably introduced into the host cell.
- 5 41. The method of claim 40, wherein the host cell is a *Salmonella* or *Escherichia coli* cell.
42. The method of claim 40, wherein the host cell is a *Salmonella enterica* cell.
43. The method of claim 40, wherein the sixth codon is the synonymous codon ACU, ACA, or ACG and wherein protein production is increased as compared to a corresponding host cell utilizing wild type triacontanucleotide sequence.
- 10 44. The method of claim 40, wherein the sixth codon is a synonymous codon and a resulting between about 9- to 15-fold increase in protein production as compared to a corresponding host cell utilizing a wild type triacontanucleotide sequence.
45. The method of claim 40, wherein the eighth codon is the synonymous codon CCG and wherein protein production is decreased as compared to a corresponding host cell utilizing a wild  
15 type triacontanucleotide sequence.
46. The method of claim 40, wherein the eighth codon is a synonymous codon and a resulting between about .03- to .07-fold decrease in protein production as compared to a corresponding host cell utilizing a wild type triacontanucleotide sequence.
47. The method of claim 40, wherein the sixth codon is the synonymous codon ACU, the eighth  
20 codon is the synonymous codon CCG, and wherein protein production is increased as compared to a corresponding host cell utilizing wild type triacontanucleotide sequence.

- 39 -

48. The method of claim 40, wherein the sixth codon is a synonymous codon, the eighth codon is a synonymous codon, and a resulting between about 27.5- to 40.5-fold increase in protein production as compared to a corresponding host cell utilizing a wild type triacontanucleotide sequence.

5 49. A method of increasing the translation speed of a polynucleotide sequence, comprising providing a polynucleotide sequence that encodes a protein and is operably linked to a wild type triacontanucleotide sequence comprising ten codons and encoding NH<sub>2</sub>-MSIDRTSPLK-COOH (SEQ ID NO:1), and mutating the triacontanucleotide sequence so that at least one of the sixth, seventh, eighth, ninth and tenth codons, in the 5' to 3' direction, are synonymous codons.

10

50. A method for increasing specific cellular productivity of a polypeptide of interest in a host cell comprising: (a) providing a polynucleotide sequence that encodes the polypeptide of interest and is operably linked to a wild type triacontanucleotide sequence comprising ten codons and encoding NH<sub>2</sub>-MSIDRTSPLK-COOH (SEQ ID NO:1), (b) mutating the triacontanucleotide sequence so that at least one of the sixth, seventh, eighth, ninth and tenth codons, in the 5' to 3' direction, are synonymous codons; (c) introducing the polynucleotide sequence into the host cell; (d) incubating the cell under conditions wherein the introduced polynucleotide is expressed; and (e) isolating the polypeptide of interest.

15

51. The method of claims 49 or 50, wherein the sixth codon of the resulting mutant triacontanucleotide sequence comprises the synonymous codon ACU, ACA, or ACG and wherein translation speed of the polynucleotide sequence is increased compared to translation under the control of the wild type triacontanucleotide sequence.

20

52. The method of claim 51, wherein the method produces a between about 9- to 15-fold increase in translation speed.

25

53. The method of claims 49 or 50, wherein the sixth codon is the synonymous codon ACU, the eighth codon is the synonymous codon CCG, and wherein translation speed of the polynucleotide

- 40 -

sequence is increased compared to translation under the control of the wild type triacontanucleotide sequence.

54. The method of claim 53, wherein the method produces a between about 27.5- to 40.5-fold increase in translation speed.

5

55. The method of any one of claims 40-54, wherein the protein or polypeptide is operably linked to a heterologous polynucleotide sequence that encodes insulin, an insulin molecule, or an insulin analog.

10

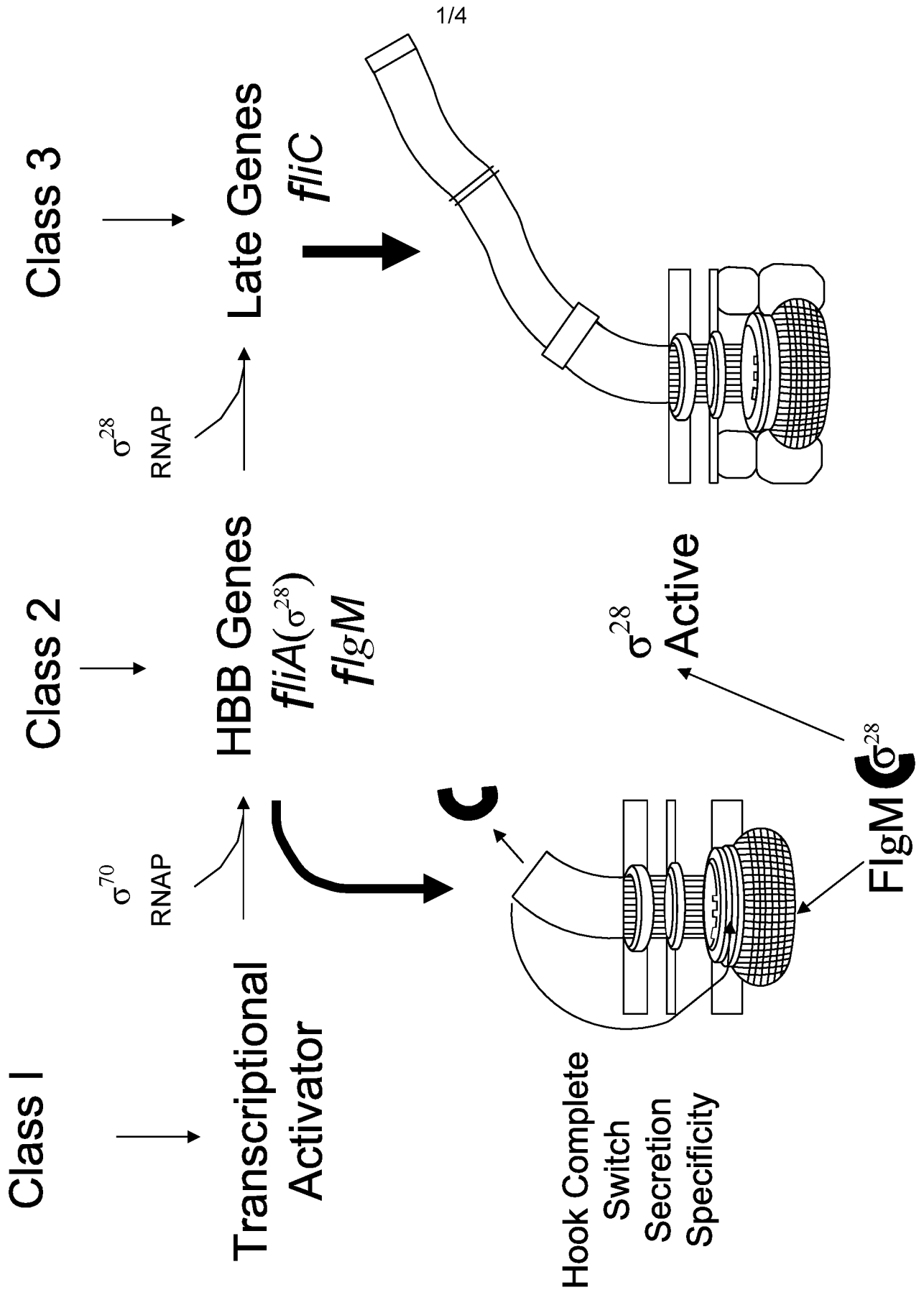
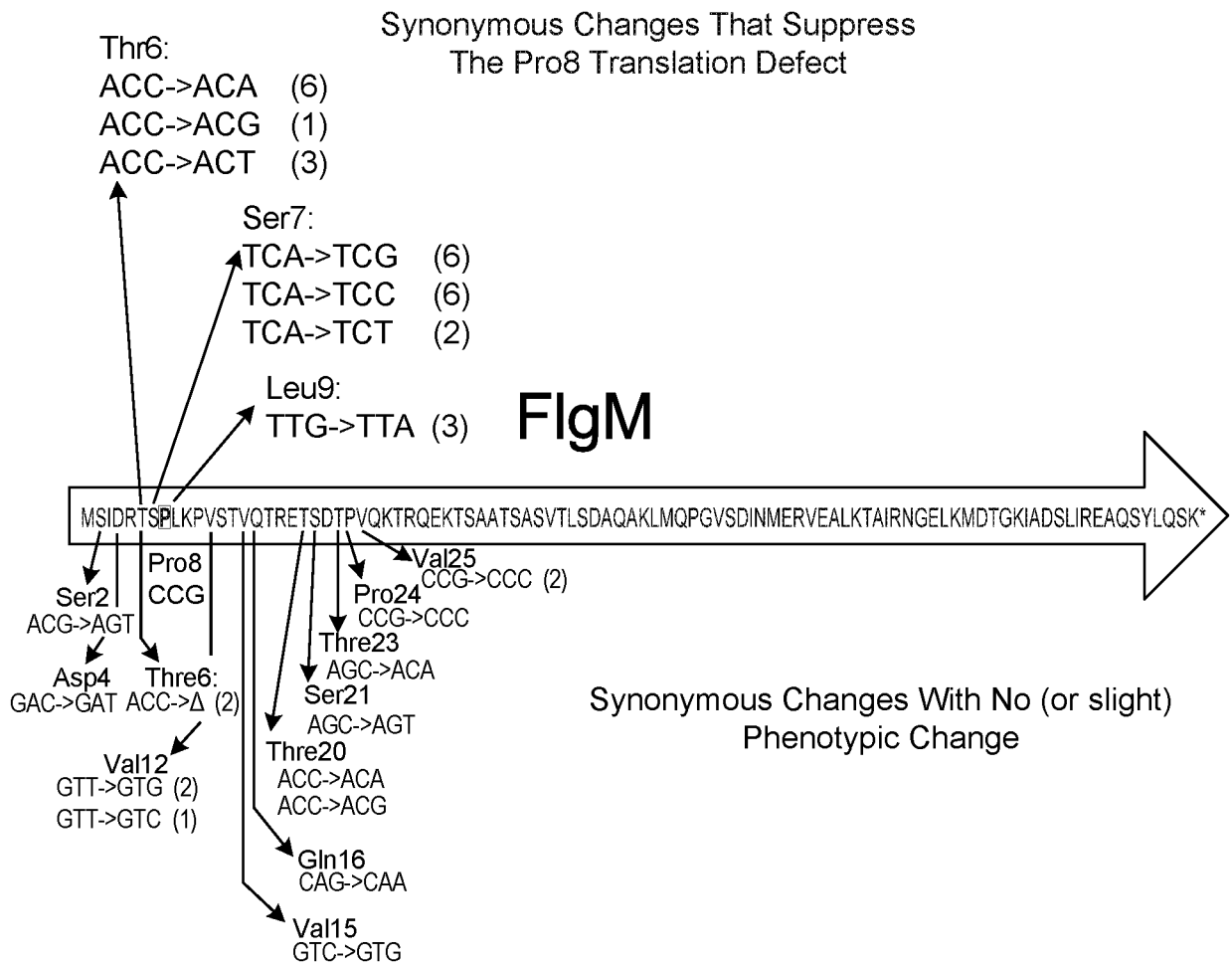


FIG. 1

2/4

Pro8 Thr6	CCU	CCC	CCA	CCG
ACU	11 $\pm$ 1.6	14 $\pm$ 2.7	12 $\pm$ 2.6	34 $\pm$ 6.4
ACC	1.0 $\pm$ 0.15	1.7 $\pm$ 0.3	1.4 $\pm$ 0.2	.053 $\pm$ 0.004
ACA	13 $\pm$ 1.9	14 $\pm$ 2.3	12 $\pm$ 2.3	11 $\pm$ 2.1
ACG	13 $\pm$ 2.1	14 $\pm$ 2.7	11 $\pm$ 2.1	6 $\pm$ 1.3

FIG. 2



**FIG. 3**

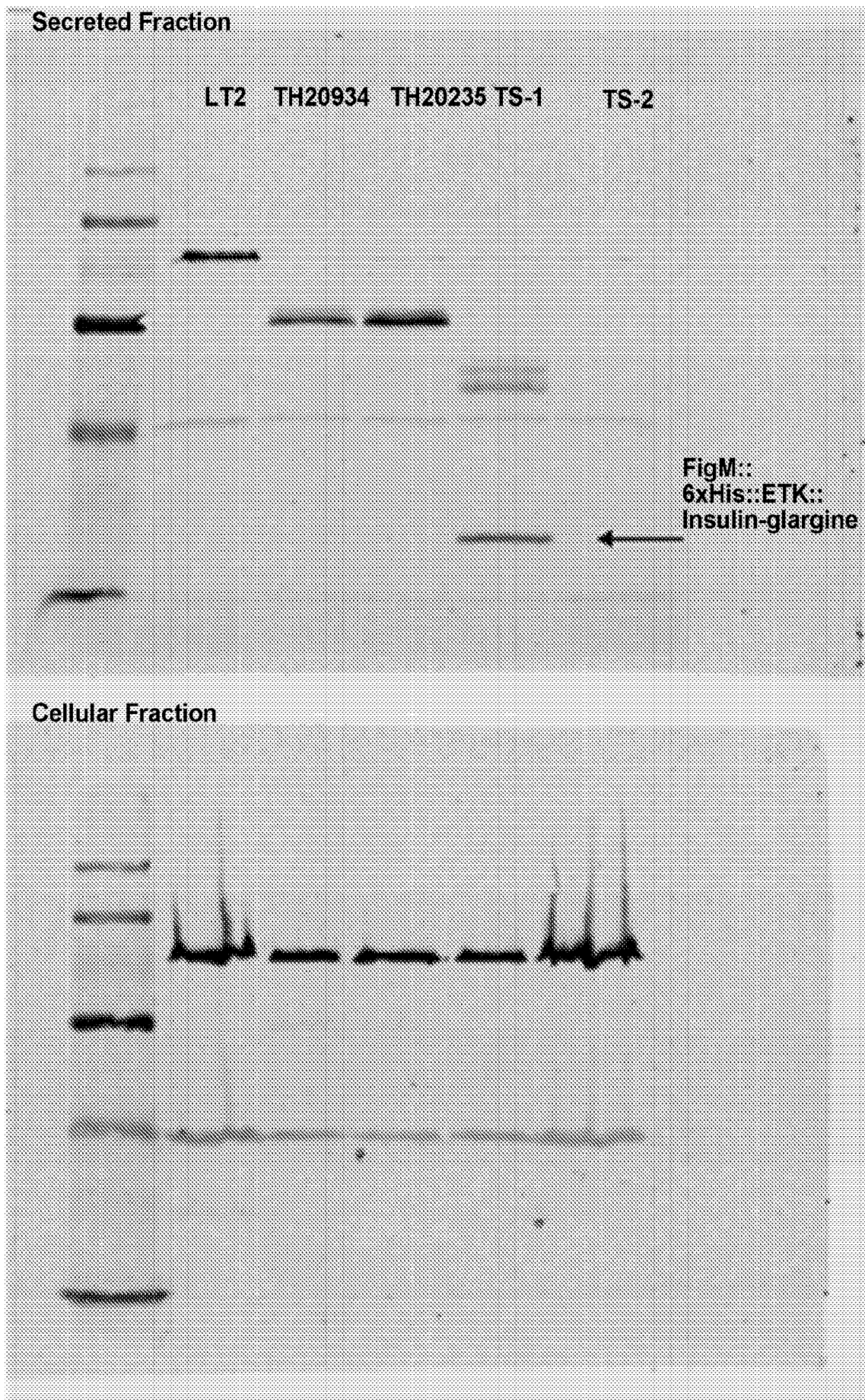


FIG. 4

INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US2016/055162

**A. CLASSIFICATION OF SUBJECT MATTER**  
 IPC(8) - C12N 15/09; C12N 15/11; C12N 15/63; C12N 15/67; C12Q 1/68 (2016.01)  
 CPC - C07K 14/245; C12N 15/63; C12N 15/67 (2016.11)  
 According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)  
 IPC - C12N 15/09; C12N 15/11; C12N 15/63; C12N 15/67; C12Q 1/68  
 CPC - C07K 14/245; C12N 15/63; C12N 15/67

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched  
 USPC - 435/6; 435/69.1; 536/23.1 (keyword delimited)

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)  
 PatBase, Google Patents, ProQuest, PubMed

Search terms used: translation speed efficiency triacontanucleotide FigM synonymous codon fusion heterologous

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2014/0244228 A1 (AGENCY FOR SCIENCE, TECHNOLOGY AND RESEARCH et al) 28 August 2014 (28.08.2014) entire document	1-16, 28-35, 49-54
A	CHEVANCE et al. "The effects of codon context on in vivo translation speed," PLoS Genet. 05 June 2014 (05.06.2014), Vol. 10, e1004392, Pgs. 1-14. entire document	1-16, 28-35, 49-54
A	US 2005/0196865 A1 (FRAZER) 08 September 2005 (08.09.2005) entire document	1-16, 28-35, 49-54
A	US 2010/0290996 A1 (NICKERSON et al) 18 November 2010 (18.11.2010) entire document	1-16, 28-35, 49-54
A	US 2013/0121915 A1 (PAAS et al) 16 May 2013 (16.05.2013) entire document	1-16, 28-35, 49-54
A	US 2015/0225466 A1 (UNIVERSITY OF UTAH RESEARCH FOUNDATION) 13 August 2015 (13.08.2015) entire document	1-16, 28-35, 49-54

Further documents are listed in the continuation of Box C.  See patent family annex.

\* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"E" earlier application or patent but published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 29 December 2016	Date of mailing of the international search report <b>09 JAN 2017</b>
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, VA 22313-1450 Facsimile No. 571-273-8300	Authorized officer Blaine R. Copenheaver PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2016/055162

**Box No. 1** Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

- a.  forming part of the international application as filed:  
 in the form of an Annex C/ST.25 text file.  
 on paper or in the form of an image file.
- b.  furnished together with the international application under PCT Rule 13*ter*. 1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
- c.  furnished subsequent to the international filing date for the purposes of international search only:  
 in the form of an Annex C/ST.25 text file (Rule 13*ter*. 1(a)).  
 on paper or in the form of an image file (Rule 13*ter*. 1(b) and Administrative Instructions, Section 713).

2.  In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

SEQ ID NOs: 1-3 and 17-23 were searched.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2016/055162

**Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)**

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1.  Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
  
2.  Claims Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
  
3.  Claims Nos.: 17-27, 36-48, 55  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

1.  As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2.  As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3.  As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
  
4.  No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

**Remark on Protest**

- The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- No protest accompanied the payment of additional search fees.