

## Methods of Producing Full-Length Antibodies Using *E. coli*

### CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority and the benefit of U.S. Provisional Application No. 63/018,436, filed on April 30, 2020, said provisional application is herein incorporated by reference in its entirety for all purposes.

### SEQUENCE LISTING

[0002] The instant application contains a Sequence Listing which has been submitted electronically in ASCII format and is hereby incorporated by reference in its entirety. Said ASCII copy, created on April 29, 2021, is named 091200-1246043-006810WO\_SL.txt and is 81,355 bytes in size.

### BACKGROUND OF THE INVENTION

[0003] Full-length antibodies play major roles in treating a wide variety of human diseases. As the demand for antibody therapeutics increases, the development of host systems for enhanced and less expensive production of the full-length antibodies has also become more important. Almost all therapeutic antibodies approved to date are predominantly produced in mammalian hosts; these production methods have drawbacks such as high production cost and long-term cultivation. Alternative production systems such as bacteria and yeasts have been considered. Still, these systems have their own limitations. For example, the proteins produced in these systems often require refolding and secretion across several biological membranes and lack disulfide bonds that are essential for maintaining the three-dimensional form of the antibodies. Also, it remains a challenge to express the heavy chain and light chain of the same antibody in appropriate ratios to ensure the heterotetramer form of the full-length antibodies can be produced in maximum yield. Thus, there remains a need for producing full-length antibodies in high yield in order to meet the demand for commercial production.

## SUMMARY OF INVENTION

[0004] This disclosure provides a method to produce a full-length antibody comprising a heavy chain (HC) and a light chain (LC) in high yield. The method comprises culturing *E. coli* cells expressing a coding sequence for the HC and a coding sequence for the LC under conditions permissible for producing the HC and LC, and the full-length antibody can be produced in an amount of at least 200 mg/L. In some embodiments, the full-length antibody is produced at a wet weight percentage in a range from 0.05% to 20%, from 0.05% to 10%, from 0.05% to 5%, or from 0.05% to 1% relative to the weight of wet cell pellet produced from the *E. coli* cells in the culture. In some embodiments, at least some of the *E. coli* cells contain an oxidative cytoplasm. In some embodiments, the method comprises transforming the plasmid to the *E. coli* strain. In some embodiments, the HC and LC are produced in the cytoplasm of the *E. coli*. In some embodiments, the method further comprises assembling the produced HC and LC under non-reducing conditions to form the full-length antibody. In some embodiments, the molar ratio of the produced HC and the produced LC from the *E. coli* is 1:1 to 1:3.

[0005] In some embodiments, expression of HC, LC, or both is controlled by a promoter, for example, a T7 promoter or a promoter that has substantially similar promoter strength as the T7 promoter. In some embodiments, the plasmid comprises a bicistronic operon, wherein the bicistronic operon comprises the coding sequence for the HC and the coding sequence for the LC. The bicistronic operon may comprise a promoter that drives expression of both the HC and LC, and the promoter is a T7 promoter or a promoter that has substantially similar promoter strength as the T7 promoter. In some embodiments, the bicistronic operon comprises a T7 terminator. In some embodiments, the plasmid encoding the HC and LC comprises a first monocistronic operon for the HC and a second monocistronic operon for LC. In some embodiments, the first operon comprises a ribosomal binding site that comprises a sequence selected from any of SEQ ID NO: 18-20 and the second monocistronic operon comprises a ribosomal binding site comprising the sequence of SEQ ID NO: 17. In some embodiments, the first monocistronic operon or the second monocistronic operon each independently comprise a T7 promoter or a promoter that has substantially similar promoter strength as the T7 promoter. In some embodiments, the first or second monocistronic operon each independently comprise a T5 promoter, or a promoter that has substantially similar promoter strength as the T5 promoter. In

some embodiments, the first and second monocistronic operons comprise the same promoter. In some embodiments, the first and second monocistronic operons comprise different promoters.

**[0006]** *E. coli* cell may comprise a first ribosome-binding site for translation of the HC and a second ribosomal binding site for translation of the LC, wherein the first and second ribosomal binding sites are selected such that the molar ratio of HC and LC produced from the *E. coli* ranges from 1:1 to 1:3. In some embodiments, the first ribosomal binding site comprises a sequence selected from the group consisting of SEQ ID NO: 17-19, and the second ribosomal binding site comprises a sequence that is selected from the group consisting of SEQ ID NO: 20-23. In some embodiments, the first ribosomal binding site comprises a sequence of SEQ ID NO: 17 or SEQ ID NO: 18.

**[0007]** In some embodiments, the HC and/or the LC of the full-length antibody comprises at least one non-natural amino acid. For example, the non-natural amino acid can be para-methylazido-L-phenylalanine (pAMF), AEK, or pAcF. For example, the coding sequence of the HC and/or the coding sequence of the LC can be modified to have at least one non-natural amino acid codon, and the non-natural amino acid codon that does not result in incorporation of one of the 20 naturally occurring amino acids. The at least one non-natural amino acid may be introduced by charging a tRNA that contains anticodons that are complementary to the at least one non-natural amino acid codon, e.g., the amber codon (TAG). In some embodiments, the codon that is immediately 3' to at least one of the non-natural amino acid is codon optimized to boost the expression of the HC or LC.

**[0008]** In certain embodiments, the full-length antibody that can be produced using the methods is a B10 antibody, an H01 antibody, a 7219 antibody, an anti-PD1 antibody, an anti HER2 antibody (e.g., trastuzumab), an anti-Tim3 antibody, or an anti LAG3 antibody. In some embodiments, the coding sequence for the HC of the B10 antibody contains a mutation relative to SEQ ID NO: 1, wherein the mutation results in a codon for a natural amino acid is being substituted with a non-natural amino acid codon, and the natural amino acid being substituted is one or more amino acids selected from F404, Y180, and F241. A non-natural amino acid can be introduced to the HC or the LC by charging a tRNA that is complementary to the non-natural amino acid codon. In some embodiments, the coding sequence for the LC of the B10 antibody or the anti HER2 antibody contains a mutation relative to SEQ ID NO: 2, wherein the mutation

results in a codon for a natural amino acid is being substituted with a non-natural amino acid codon, wherein the natural amino acid is K42 or E161.

[0009] In some embodiments, the method of producing a full length FolRa-B10 IgG antibody comprises transforming an *E. coli* host strain with a construct expressing FolRa-B10 IgG, where the S181 codon of the HC encoding sequence is AGC or AGT.

[0010] In some embodiments, the method further comprises covalently linking a warhead moiety to the non-natural amino acid on the HC or the LC of the full-length antibody via a linker.

#### BRIEF DESCRIPTION TO THE DRAWINGS

[0011] FIG. 1 shows the comparison of titer in cell lysate from Shuffle and Snuggle strains co-expressing the heavy chain (HC) and the light chain (LC) for the FoRa B10 wt IgG, indicating additional mutations present in Snuggle, including e.g., the deletion of the *trx4* gene and the presence of the wt *ahpC* gene, are essential for high-level IgG synthesis.

[0012] FIG. 2 shows the comparison of titer in cell lysate for FolRa-B10 IgGs produced with 0, 2, or 4 pAMF non-natural amino acids and the indicated S181 codons. "DAR0" represents the B10 IgG that does not contain any non-natural amino acids.

[0013] FIG. 3 shows the comparison of assembly of FolRa-B10 IgGs produced with 0, 2, or 4 pAMF non-natural amino acids after overnight reassembly by FlowerPlate determined by capillary bioanalyzer under non-reducing conditions.

[0014] FIG 4 shows the expression of B10-F404pAcF and B10-F404AEK in the presence and absence of the appropriate non-natural amino acid in the culture media. The low titers in the absence of non-natural amino acid (NNAA) indicate that full-length IgG production requires NNAA for amber suppression.

[0015] FIG. 5 shows the results of deconvoluted LCMS spectra for Fc-fragment of B10 F404pAMF that is conjugated to the DBCO-maytansine drug-linker SC236. The predominant species observed has a mass of 25257.97 Da compared with the theoretical mass of 25258.04 Da for the conjugate. No species were observed near 23974.14 Da, the theoretical mass for the unconjugated Fc fragment. This conjugate has a calculated conjugation efficiency of 99.997%.

[0016] FIG. 6 shows the results of LCMS of purified trastuzumab LC K42pAcPhe E161pAcPhe run in positive ionization mode. There is one clean peak corresponding to the expected mass for the 2x pAcPhe LC product.

[0017] FIG. 7 shows the results of the LCMS of purified trastuzumab LC K42pAcPhe E161pAcPhe after oxime ligation with PEG8-alkoxy amine, run in positive ionization mode. Conjugation efficiency was calculated using peaks at 23692, 24073 and 24455 AMUs which correspond to the expected mass for the 0, 1 and 2 PEG8 products respectively. A) shows LC before oxime ligation and B) after oxime ligation. The calculated conjugation efficiency was around 98%.

[0018] FIG. 8A-8I show exemplary non-natural amino acids that can be used in the methods disclosed herein.

## DETAILED DESCRIPTION OF THE INVENTION

### I. Introduction

[0019] This disclosure provides methods of producing a full-length antibody by expressing both the heavy chain (HC) and the light chain (LC) of the antibody in *E. coli* cells. The method of the present invention utilizes a combination of features that allow the production of full-length antibodies in high yield, e.g., at a concentration of greater than 200 mg of protein per liter of culture medium.

[0020] The *E. coli* cells typically have an oxidative cytoplasm, which helps the formation of the disulfide bonds that are required to maintain the three-dimensional structure and stability of the antibody. The methods employ a strong promoter to drive the transcription of the HC, the LC, or both. In some embodiments, the promoter is a T7 promoter or a T5 promoter. In some cases, the strength of the promoter is substantially similar to that of the T7 promoter or the T5 promoter. In some cases, the methods employ optimized ribosomal binding sites to ensure efficient translation and/or assembly of the HC and LC of the antibodies.

[0021] The inventors have found that maintaining the proportional expression of the heavy chain (HC) and light chain (LC), i.e. the ratio of the amount of expression of the HC and LC within an optimal range, can increase the yield of the full-length and properly assembled

antibody. For example, it is desirable to maintain the molar ratio of the produced HC and LC from the *E. coli* culture within a suitable range, e.g., from 1:1 to 1:3, from 1:1 to 1:2.5, or from 1:1 to 1:2. In some approaches, maintaining the expression of the HC and LC within the optimal range ratio is achieved by using one or two promoters that can control the transcription level of HC and LC. In some approaches, the proportional expression of the HC and LC in the *E. coli* culture is achieved by using modified ribosomal binding sites to control the translation efficiency of the HC and LC. Exemplary ribosomal binding sequences are provided herein.

[0022] In some approaches, at least one of the HC and the LC coding sequence is modified to comprise at least one non-natural amino acid codon (e.g., an amber codon), provided that the non-natural amino acid codon does not result in incorporation of one of the 20 naturally occurring amino acids. These non-natural amino acids may serve as potential reactive groups for conjugating with one or more biologically active adducts in a site-specific manner.

[0023] This disclosure incorporates the entire content of the International Application No., PCT/US2019/060345 by reference.

## II. Definition

[0024] Unless defined otherwise, technical and scientific terms used herein have the same meaning as commonly understood by a person of ordinary skill in the art. See, e.g., Lackie, DICTIONARY OF CELL AND MOLECULAR BIOLOGY, Elsevier (4th ed. 2007); Sambrook *et al.*, MOLECULAR CLONING, A LABORATORY MANUAL, Cold Springs Harbor Press (Cold Springs Harbor, NY 1989); Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley and Sons (Hoboken, NY 1995).

[0025] As used herein the singular forms “a,” “and,” and “the” include plural referents unless the context clearly dictates otherwise. Thus, for example, reference to “a cell” includes a plurality of such cells and reference to “the protein” includes reference to one or more proteins and equivalents thereof known to those skilled in the art, and so forth. All technical and scientific terms used herein have the same meaning as commonly understood to one of ordinary skill in the art to which this invention belongs unless clearly indicated otherwise.

[0026] The term “about” denotes a range of  $\pm 10\%$  of a reference value, e.g., “about 10,” means  $10 \pm 10 \times 10\% = 9-11$ .

[0027] Unless explicitly stated otherwise, the terms “yield” or “titer” refer to the amount of antibody produced (including the total amount of HC and LC) relative to the culture medium volume. For example, 200 mg/L refers to 200 mg of antibody produced per liter of culture medium.

[0028] The term “*E. coli* strain” refers to a subtype of *E. coli*, the cells of which have a certain biological form and share certain genetic makeup. The term “*E. coli* strain having oxidative cytoplasm” refers to that some or all of the cells derived from the strain each have an oxidative cytoplasm.

[0029] The term “amino acids” refers to naturally occurring and non-naturally occurring amino acids, as well as amino acids such as proline, amino acid analogs, and amino acid mimetics that function in a manner similar to naturally occurring amino acids.

[0030] The term “naturally occurring amino acids” or “naturally encoded amino acids” refer to the proteinogenic amino acids known to those of skill in the art. They include the 20 common amino acids (alanine, arginine, asparagine, aspartic acid, cysteine, glutamine, glutamic acid, glycine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine, and valine) and the less common pyrrolysine and selenocysteine. Naturally encoded amino acids include post-translational variants of the 22 naturally occurring amino acids such as prenylated amino acids, isoprenylated amino acids, myrisoylated amino acids, palmitoylated amino acids, N-linked glycosylated amino acids, O-linked glycosylated amino acids, phosphorylated amino acids, and acylated amino acids.

[0031] The term “non-natural amino acid” refers to an amino acid that is not a proteinogenic amino acid or a post-translationally modified variant thereof. In particular, the term refers to an amino acid that is not one of the 20 common amino acids, pyrrolysine, selenocysteine, or post-translationally modified variants thereof.

[0032] The term “aminoacylation” or “aminoacylate” refers to the complete process in which a tRNA is charged with its correct amino acid that is a result of adding an aminoacyl group to a compound. As it pertains to this invention, a tRNA that undergoes aminoacylation or has been aminoacylated is one that has been charged with an amino acid, and an amino acid that

undergoes aminoacylation or has been aminoacylated is one that has been charged to a tRNA molecule.

**[0033]** The term “aminoacyl-tRNA synthetase” or “tRNA synthetase” or “synthetase” or “aaRS” or “RS” refers to an enzyme that catalyzes a covalent linkage between an amino acid and a tRNA molecule. This results in an aminoacylated tRNA molecule, which is a tRNA molecule that has its respective amino acid attached via an ester bond.

**[0034]** The term “charged” in the context of tRNA refers to the aminoacylation of a tRNA with an amino acid, both natural and non-natural, where the aminoacylation permits a ribosome to incorporate the amino acid into a polypeptide being translated from mRNA.

**[0035]** The term “biologically active adduct” refers to a chemical, molecule or reagent that can perform a function in a cell or an organism. For example, the function may include cell proliferation, apoptosis, post-translational modification (*e.g.*, phosphorylation), cell signaling activation, cell signaling inactivation, cell death, or cell labeling.

**[0036]** The term “selective incorporating” in the context of protein translation refers to including or introducing a specific amino acid (*e.g.*, a specific non-natural amino acid) in a predetermined, desired amino acid position in the sequence of the protein without disturbing the desired function of the protein.

**[0037]** The term “preferentially aminoacylates” refers to the preference of a tRNA synthetase to aminoacylate (charge) a particular tRNA molecule with a predetermined amino acid molecule compared to another amino acid molecule. In other words, the tRNA synthetase can selectively aminoacylate a non-natural amino acid (nnAA) over a naturally occurring amino acid. For example, the tRNA synthetase can aminoacylate a specific nnAA at a frequency of greater than 90%, (*e.g.*, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100%), compared to any or all other natural amino acids.

**[0038]** The term “naturally occurring amino acid” refers to any one of the 20 amino acids encoded by the genetic code, such as arginine, Arg, R; histidine, His, H; lysine, Lys, K; aspartic acid, Asp, D; glutamic acid, Glu, E; serine, S, Ser; threonine, Thr, T; asparagine, Asn, N; glutamine, Gln, Q; cysteine, Cys, G; glycine, Gly, G; proline, Pro, P; alanine, Ala, A; isoleucine,

Ile, I; leucine, Leu, L; methionine, Met, M; phenylalanine, Phe, F; tryptophan, Trp, W; tyrosine, Tyr, Y; and valine, Val, V, that are precursors to proteins.

**[0039]** Amino acids may be referred to herein by either the commonly known three-letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes.

**[0040]** The term “nucleic acid” or “polynucleotide” refers to deoxyribonucleic acids (DNA) or ribonucleic acids (RNA) and polymers thereof in either single- or double-stranded form. Unless specifically limited, the term encompasses nucleic acids containing known analogues of natural nucleotides that have similar binding properties as the reference nucleic acid and are metabolized in a manner similar to naturally occurring nucleotides. Unless otherwise indicated, a particular nucleic acid sequence also implicitly encompasses conservatively modified variants thereof (e.g., degenerate codon substitutions), alleles, orthologs, SNPs, and complementary sequences as well as the sequence explicitly indicated. Specifically, degenerate codon substitutions may be achieved by generating sequences in which the third position of one or more selected (or all) codons is substituted with mixed-base and/or deoxyinosine residues (Batzer *et al.*, *Nucleic Acid Res.* 19:5081 (1991); Ohtsuka *et al.*, *J. Biol. Chem.* 260:2605-2608 (1985); and Rossolini *et al.*, *Mol. Cell. Probes* 8:91-98 (1994)). The term nucleic acid or polynucleotide is used interchangeably with gene, cDNA, and mRNA encoded by a gene.

**[0041]** The term “peptide,” “protein,” and “polypeptide” are used herein interchangeably and refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical mimetic of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers and non-naturally occurring amino acid polymers. As used herein, the terms encompass amino acid chains of any length, including full-length proteins and truncated proteins, wherein the amino acid residues are linked by covalent peptide bonds.

**[0042]** The term “substitution at amino acid position” refers to a change of an amino acid residue at a specific position of the amino acid sequence of a protein. For example, the term “X20Y” refers to the replacement of the wild-type (reference) amino acid X at amino acid position 20 of the protein with the amino acid Y.

[0043] The term “suppression codon” refers to a nucleotide triplet that is introduced into a polynucleotide at a predetermined location and is recognized by a specific tRNA that can recognize a stop codon (e.g., an amber, ochre or opal stop codon) and allows translation to read through the codon to produce the protein, thereby suppressing the stop codon.

[0044] The term “antibody” refers to a protein functionally defined as a binding protein and structurally defined as comprising an amino acid sequence that is recognized by one of skill as being derived from the framework region of an immunoglobulin encoding gene of an animal producing antibodies. An antibody can consist of one or more polypeptides substantially encoded by immunoglobulin genes or fragments of immunoglobulin genes. The recognized immunoglobulin genes include the kappa, lambda, alpha, gamma, delta, epsilon and mu constant region genes, as well as myriad immunoglobulin variable region genes. Light chains are classified as either kappa or lambda. Heavy chains are classified as gamma, mu, alpha, delta, or epsilon, which in turn define the immunoglobulin classes, IgG, IgM, IgA, IgD and IgE, respectively.

[0045] The term “full-length antibody,” used interchangeably with “full-length IgG,” refers to an immunoglobulin (antibody) structural unit comprising a tetramer, each is composed of two identical pairs of polypeptide chains, each pair having one "light" (about 25 kD) and one "heavy" chain (about 50–70 kD). The N-terminus of each chain defines a variable region of about 100 to 110 or more amino acids primarily responsible for antigen recognition. The terms variable light chain (VL) and variable heavy chain (VH) refer to these light and heavy chains respectively.

[0046] Antibodies exist as intact immunoglobulins or as a number of well characterized fragments produced by digestion with various peptidases. Thus, for example, pepsin digests an antibody below the disulfide linkages in the hinge region to produce F(ab)'<sub>2</sub>, a dimer of Fab which itself is a light chain joined to VH-CH1 by a disulfide bond. The F(ab)'<sub>2</sub> may be reduced under mild conditions to break the disulfide linkage in the hinge region thereby converting the (Fab)'<sub>2</sub> dimer into a Fab' monomer. The Fab' monomer is essentially a Fab with part of the hinge region (see, Fundamental Immunology, W.E. Paul, ed., Raven Press, N.Y. (1993), for a more detailed description of other antibody fragments). While various antibody fragments are defined in terms of the digestion of an intact antibody, one of skill will appreciate that such Fab' fragments may be synthesized de novo either chemically or by utilizing recombinant DNA

methodology. Thus, the term antibody, as used herein also includes antibody fragments either produced by the modification of whole antibodies or synthesized de novo using recombinant DNA methodologies. Antibodies also include single chain antibodies (antibodies that exist as a single polypeptide chain), and single chain Fv antibodies (sFv or scFv) in which a variable heavy and a variable light chain are joined together (directly or through a peptide linker) to form a continuous polypeptide. The single chain Fv antibody is a covalently linked VH-VL heterodimer which may be expressed from a nucleic acid including VH- and VL- encoding sequences either joined directly or joined by a peptide-encoding linker. Huston, *et al.* (1988) Proc. Nat. Acad. Sci. USA, 85: 5879-5883. While the VH and VL are connected to each as a single polypeptide chain, the VH and VL domains associate non-covalently. The first functional antibody molecules to be expressed on the surface of filamentous phage were single-chain Fv's (scFv); however, alternative expression strategies have also been successful. For example, Fab molecules can be displayed on phage if one of the chains (heavy or light) is fused to g3 capsid protein and the complementary chain exported to the periplasm as a soluble molecule. The two chains can be encoded on the same or on different replicons; the important point is that the two antibody chains in each Fab molecule assemble post-translationally, and the dimer is incorporated into the phage particle via linkage of one of the chains to g3p (*see, e.g.*, U.S. Patent No: 5733743). The scFv antibodies and a number of other structures converting the naturally aggregated, but chemically separated light and heavy polypeptide chains from an antibody V region into a molecule that folds into a three-dimensional structure substantially similar to the structure of an antigen-binding site are known to those of skill in the art (*see, e.g.*, U.S. Patent Nos. 5,091,513, 5,132,405, and 4,956,778). Antibodies also include all those that have been displayed on phage (e.g., scFv, Fv, Fab and disulfide linked Fv (Reiter *et al.* (1995) Protein Eng. 8: 1323-1331). Antibodies can also include diantibodies, miniantibodies and scFv-Fc fusions.

[0047] The term “reductase” refers to a thioredoxin reductase (TrxB), glutathione or glutathione reductase (Gor) or any other enzyme that can reduce members of the thioredoxin or glutaredoxin systems.

[0048] The term “thioredoxin” includes thioredoxin 1 (TrxA) and thioredoxin 2 (TrxC), as described in Rietsch and Beckwith (1998) Ann. Rev. Genet. 32: 163. Thioredoxins are small proteins characterized by the presence of the motif Cys-Xaa-Xaa-Cys (where Xaa denotes any

amino acid) in their active site. Thioredoxin is re-reduced by thioredoxin reductase (encoded by the *trxB* gene) and NADPH. In a *trxB* mutant, thioredoxin accumulates in an oxidized form. TrxA is encoded by the *trxA* gene and TrxB is encoded by the *trxB* gene.

[0049] The term “*gor*” refers to the glutathione oxidoreductase gene and the term “GOR” refers to glutathione oxido-reductase.

[0050] “DsbC” is a protein encoded by the gene *dsbC*, which catalyzes disulfide bond isomerization. DsbC null mutants have a defect in the folding of proteins with multiple disulfide bonds.

[0051] The term “glutathione” refers to  $\gamma$ -L-glutamyl-L-cysteinyl-glycine (GSH), which is a highly conserved low molecular weight thiol found in many organisms including cyanobacteria, proteobacteria, a few strains of gram-positive bacteria, and in all eukaryotes having mitochondria and chloroplasts. Glutathione is synthesized by the action of two enzymes: glutamate-cysteine ligase (*gsh4*) and glutathione synthetase (*gshB*). Glutamate-cysteine ligase catalyzes the reaction between glutamic acid and cysteine to form  $\gamma$ -glutamyl cysteine, which is subsequently conjugated to glycine by glutathione synthetase to form GSH.

[0052] A nucleic acid is “operably linked” to another nucleic acid when it is placed into a functional relationship with another nucleic acid sequence. For example, a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosomal binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, “operably linked” means that the DNA sequences being linked are contiguous and, in the case of a secretory leader, contiguous and in reading phase. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, the synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice.

[0053] AhpC is one of the two subunits of the alkyl hydrogen peroxide reductase AhpCF. The other subunit of AhpCF is flavoenzyme AhpF. Tarataglia et al, J. Biol. Chem., Volume 265, 10535-10540, 1990; Smillie et al, Genbank submission NCBL gi; 216542, 1993). These two proteins act together; AhpF uses NADH or NADPH as electron donor to AhpC, which reduces physiological lipid peroxides such as linoleic acid hydroperoxide and thymine hydroperoxide and nonphysiological alkyl hydroperoxides to their respective nontoxic alcohol forms. This

enzymatic complex (or system) scavenges oxygen and its derivatives. AhpC has been demonstrated to act as specific alkyl hydroperoxide-scavenging enzyme for protection against oxygen radical damage, though elimination of reactive nitrogen intermediates also has been demonstrated to occur. AhpF is related to thioredoxin reductases possessing an extended additional N-terminal fragment essential to specifically reduce AhpC.

**[0054]** The term “cytosolic” when used to describe a protein refers to that the protein is present in the cytosol of a cell.

**[0055]** A “heterologous protein or polypeptide” refers to a protein or polypeptide which is not normally produced in the host cell. A heterologous polypeptide can be from the same species and type as the host cell provided that it is expressed from a nucleic acid which has been introduced into the host cell.

**[0056]** An “exogenous polypeptide” refers to a polypeptide that is not naturally produced in a cell and is expressed or present in the cell after being introduced into the cell.

**[0057]** A “null mutation” refers to a mutation in a gene that result in a nonfunctional gene. The null mutation can cause complete lack of production of associated gene product or a product that does not function properly.

**[0058]** The term “protein disulfide isomerase,” used interchangeably with the term “disulfide isomerase” or “PDI,” refers to an enzyme that catalyzes disulfide bond formation and isomerization. PDI has been implicated in the catalysis of disulfide bond formation and rearrangement through in vitro data. (Creighton et al. (1980) *J. Mol. Biol.* 142:43; Feedman et al. (1989) *Biochem. Soc. Symp.* 5:167; and Bardwell and Beckwith (1993) *Cell* 74:899. Yeast mutants in PDI have been shown to have a defect in the formation of disulfide bonds in carboxypeptidase Y (LaMantia and Lennarz (1993) *Cell* 74:899). Use of PDI for expression of heterologous proteins in host cells is further described in PCT application having publication No. WO 93/25676; WO 94/08012; and EP 509,841.

**[0059]** The term “prolyl isomerase,” used interchangeably with “peptidylprolyl isomerase” or “PPlase,” refers to an enzyme found in both prokaryotes and eukaryotes that interconverts the *cis* and *trans* isomers of peptide bonds with the amino acid proline. Proteins with prolyl isomerase activity include, but are not limited to, cyclophilin (e.g., accession # Q13427), FKBP (e.g.,

accession # Q02790), parvulin (e.g., accession # Q9Y237), Tig (e.g., accession #P0A850 ), SlyD (e.g., accession # P0A9K9), and yCpr6 (e.g., accession #S000004206).

**[0060]** The term "deaggregase" refers to a protein chaperone that aids in deaggregating and/or solubilizing proteins of interest that are produced, for example, in a bacterial free translation system. Such chaperones are particularly helpful at high concentrations because their mechanism of action is stoichiometric rather than catalytic and is believed to work by stabilizing hydrophobic patches of the newly synthesized protein while the protein is folding. Non-limiting examples of deaggregases include Skp (e.g., accession # P0AEU7), GroEL(e.g., accession #P0A6F5), GroES (e.g., accession #P0A6F9), DnaK (e.g., accession #P0A6Y8 ), DnaJ (e.g., accession #P08622 ), or GrpE (e.g., accession #P09372).

**[0061]** When a protein is described to be in a "reduced state," it refers to the protein having more electrons than its oxidized form.

**[0062]** The term "oxidative cytoplasm" refers to the cytosol of a cell in which a substrate is more likely to become oxidized than reduced.

**[0063]** The term "thioredoxin reductase activity" refers to the ability of thioredoxin reductase (TRXB) to maintain thioredoxin 1 in the reduced state.

**[0064]** The term "thioredoxin 1 activity" refers to the ability of thioredoxin1 (TRXA) to maintain ribonucleotide reductase at reduced state.

**[0065]** The term "peroxyreductase activity" refers to the ability of AhpC to reduce physiological lipid oxide.

**[0066]** The term "glutathione reductase activity" refers to the ability of catalyzing the reduction of glutathione disulfide (GSSG) to the sulfhydryl form glutathione (GSH). For example, glutathione reductase (GOR) possesses the glutathione reductase activity.

**[0067]** The term "recombinant" or "recombinantly" refers to a biomolecule, e.g., a gene or protein, that (1) has been removed from its naturally occurring environment, (2) is not associated with all or a portion of a polynucleotide in which the gene is found in nature, (3) is operatively linked to a polynucleotide which it is not linked to in nature, or (4) does not occur in nature. The term "recombinant" can be used in reference to cloned DNA isolates, chemically synthesized

polynucleotide analogs, or polynucleotide analogs that are biologically synthesized by heterologous systems, as well as proteins and/or mRNAs encoded by such nucleic acids.

[0068] The term “culture titer” refers to the concentration of the protein in the culture media. For example, a culture titer of “136 mg/l” refers to 136 mg protein per 1 liter of cell culture media.

[0069] The term “lysate titer” refers to the concentration of the protein in the cell lysate. For example, a lysate titer of “0.655 g/l” refers to 0.655 g protein per 1 liter of cell lysate.

[0070] The term “weight percentage,” when used to describe the titer or yield of protein produced from a cell culture, refers to the wet weight percentage, i.e., the ratio of the weight of a protein (e.g., IgG) to the weight of cell pellet (undried), from which the protein is extracted. The cell pellet is collected from a cell culture expressing the protein and lysed to release the protein using standard techniques (e.g., as described in Example 6). As an illustrative example, if 100 g cell pellet is lysed to produce 1 liter of lysate, which is determined to contain 50 mg IgG, the weight percentage of the IgG is  $50 \text{ mg} / 100 \text{ g} = 0.05\%$ .

### III. Host cells- *E. coli* strains having an oxidative cytoplasm

[0071] The *E. coli* strain in this disclosure can be any *E. coli* strain that is known to one of skill in the art. In some embodiments, the *E. coli* strain is a A (K-12), B, C or D strain. In some embodiments, each of at least some *E. coli* cells from the *E. coli* strain contains an oxidative cytoplasm such that the disulfide bonds in the antibody can form and be maintained. The oxidative status of the cytosol is typically assessed by measuring the redox potential of the cytoplasm of the *E. coli* cell. Methods for measuring cellular redox status are well known, e.g., as described in Gilbert et al. (1990) Adv. Enzymol. Rel. Areas Mol. Biol. 63:69; Holmgren and Fgestedt (1982) J. Biol. Chem. 257: 6926; and Hwang et al. (1992) Science 257: 1496. In some cases, the cytosol of the *E. coli* strain has redox potential in the range of -400 mV to -100 mV, e.g., -350 mV to -150 mV, -300 mV to -200 mV, or -291 mV to -250 mV.

#### Producing an *E. coli* strain having oxidative cytoplasm by genetic engineering

[0072] An *E. coli* strain having oxidative cytoplasm may be one that displays reduced or abolished expression or activity of at least one reductase as compared to a corresponding wild

type *E. coli* strain. The reductase may be one or more reductases selected from the group consisting of thioredoxin reductase, glutathione reductase, and glutathione. In some approaches, the *E. coli* strain contains a null mutation in at least one of the reductase disclosed above. In some embodiments, the *E. coli* strain lacks the activity of a thioredoxin reductase encoded by *trxB*; lacks the activity of a thioredoxin 1 encoded by *trxA*; and/or lacks the activity of a glutathione reductase encoded by *gor*. In some embodiments, the *E. coli* strain expresses a mutated AhpC protein, wherein the mutated AhpC protein has glutathione reductase activity. In one illustrative embodiment, the *E. coli* strain is the Snuggle strain, which comprises null mutations in *trxA*, *trxB*, and *gor* and lacks the thioredoxin reductase activity (TrxB), thioredoxin 1 activity (TrxA), and glutathione reductase activity (GOR). The Snuggle strain also overexpresses a DsbC without its signal sequence and overexpresses a variant of the *ahpC* gene (*ahpC\**) that encoded an enzyme that lacks peroxireductase activity but has glutathione reductase activity. Methods for confirming that expression of these enzymes are reduced are well known and for example, as disclosed below.

#### Methods of introducing mutations to *E. coli*

[0073] In some embodiments, the gene modifications, e.g., the knock-outs of *trxA* and *trxB*, can be performed with a site-specific recombination. Site-specific recombination uses enzymes possessing both endonuclease activity and ligase activity and the enzymes recognize a certain part of DNA sequences and replace it with any other corresponding DNA sequences, see Yang W. and Mizuuchi K., Structure, 1997, Vol. 5, 1401-1406(9). Site-specific recombination systems are well known in the art, e.g., Int/att system from bacterio  $\lambda$  phage, Cre/LoxP system from PI bacteriophage and FLP-FRT system from yeast.

[0074] Non-limiting examples of methods of introducing site-specific recombination to various proteins disclosed herein include the Cre/Lox and Flp/Frt recombination systems. Both systems are well-known in the art. For instance, site-specific integration into bacterial chromosomes has been reported (see, e.g., Sauer et al., Proc. Natl. Acad. Sci. 85, 5166-5170 (1988); Fukushige et al., Proc. Natl. Acad. Sci. , 89, 7905-7907 (1992); Baubonis et al. , Nucleic Acids Research. 21, 2025-2029 (1993); Hasan et al., Gene, 150, 51-56 (1994); Golic et al., Cell. 5\_9, 499-509 (1989); Sauer, Mol. Cell. Biolo. 1\_, 2087-2096 (1987); Sauer et al., Methods: Companion to Methods in Enzymol. 4., 143-149 (1992); Sauer et al., The New Biologist. 2.,

441-449 (1990); Sauer et al., *Nucleic Acids Res.* 17, 147-161 (1989); Qin et al., *Proc. Natl. Acad. Sci.* 91, 1706-1710 (1994); Orban et al., *Proc. Natl. Acad. Sci.*, 89, 6861-6865 (1992)). Specific deletions of chromosomal sequences and rearrangements and the excision of foreign DNA as a plasmid from  $\lambda$  vectors are also known (see, e.g., Barinaga, *Science*, 265, 27-28 (1994); Sauer, *Methods in Enzymol.* 225, 890-900 (1993); Sauer et al., *Gene*, 70, 331-341 (1988); Brunelli et al., *Yeast*, 1309-1318 (1993); Invitrogen (San Diego, CA) 1995 Catalog, 35; Clontech (Palo Alto, CA) 1995/1996 Catalog, 187-188). Cloning schemes have been generated so that recombination either reconstitutes or inactivates a functional transcription unit by either deletion or inversion of sequences between recombination sites (see, e.g., Odell et al., *Plant Physiol.* 106, 447-458 (1994); Gu et al., *Cell*, 73, 1155-1164 (1993); Lakso et al., *Proc. Natl. Acad. Sci.* 89, 6232-6236 (1992); Fiering et al., *Proc. Natl. Acad. Sci.* 90, 8469-8473 (1993); O'Gorman et al., *Science*, 251, 1351-55 (1991); Jung et al., *Science*, 259, 984-987 (1993)).

**[0075]** Genes encoding the Cre or FLP recombinases can be provided in trans under the control of either constitutive, inducible or developmentally-regulated promoters, or purified recombinase has been introduced (see, e.g., Baubonis et al., *supra*; Dang et al., *Develop. Genet.* 13, 367-375 (1992); Chou et al., *Genetics*, 131, 643-653 (1992); Morris et al., *Nucleic Acids Res.* 19, 5895-5900 (1991)).

**[0076]** In some embodiments, the genomic manipulations disclosed herein are performed with a modified site-specific recombination protocol from Kirill A. Datsenko and Barry L. Wanner *Proc Natl Acad Sci U S A.* 2000 Jun 6; 97(12): 6640-6645. In one embodiment, knocking out a gene for example, *trxA*, can be performed as follows: A PCR amplicon was generated comprising an antibiotic resistance gene flanked by two FRT sites and homology extensions, (H1 and H2), which are homologous to the two ends of the gene to be knocked out. After transforming cells with this PCR product, the gene to be knocked out is then replaced by the antibiotic resistance gene through Red-mediated recombination in these flanking homology regions. After selection, the resistance gene can be eliminated using a helper plasmid expressing the FLP recombinase, which acts on the directly repeated FRT (FLP recognition target) sites flanking the resistance gene. The Red and FLP helper plasmid can be simply cured by growth at 37 °C because they are temperature-sensitive replicons. Knocking-in a gene, such as *dsbC*, can

be performed by standard molecular cloning techniques that are well known for one skilled in the art.

[0077] In some embodiments, the inactivation including deletion of a gene is performed using a CRISPR/Cas system. The CRISPR/Cas system uses a Cas protein and at least one to two ribonucleic acids that are capable of directing the Cas protein to a sequence in a target gene, e.g., *gor*, to remove the gene. Methods of using CRISPR/Cas system to eliminate gene expression are well known and described in e.g., US. Pat. Pub. No. 2014/0170753, the disclosure of which hereby is incorporated by reference in its entirety.

[0078] Additional methods of knocking out a target gene include, but are not limited to, homologous recombination technology, transcription activation of the effector nuclease (Transcription Activator-Like Effector Nuclease, TALEN) technology, a zinc finger nuclease (Zinc-Finger Nuclease, ZFN). These methods are also well known in the art.

#### Confirming modified expression of the reductases

[0079] Various methods can be used to determine protein expression level of the various modified genes in the *E. coli*, and/or confirm whether a gene has been knocked out or knocked in. For example, expression of a gene can be determined by conventional Northern blotting to quantitate the transcription of mRNA. Various labels may be employed, most commonly radioisotopes. However, other techniques may also be employed, such as using biotin-modified nucleotides for introduction into a polynucleotide. The biotin then serves as the site for binding to avidin or antibodies, which may be labeled with a wide variety of labels, such as radionuclides, fluorescers, enzymes, or the like.

[0080] In some embodiments, the expressed protein can be purified and quantified using gel electrophoresis (e.g., PAGE), Western analysis or capillary electrophoresis (e.g., Caliper LabChip). Protein synthesis in cell-free translation reactions may be monitored by the incorporation of radiolabeled amino acids, typically, <sup>35</sup>S-labeled methionine or <sup>14</sup>C-labeled leucine. Radiolabeled proteins can be visualized for molecular size and quantitated by autoradiography after electrophoresis or isolated by immunoprecipitation. The incorporation of recombinant His tags affords another means of purification by Ni<sup>2+</sup> affinity column chromatography. Protein production from expression systems can be measured as soluble protein yield or by using an assay of enzymatic or binding activity.

[0081] In some embodiments, if the protein to be quantified possesses defined biological activity, for example, enzymatic activity (such as alkaline phosphatase) or growth inhibition activity, the expression of the protein of interest can be confirmed by assaying its activity by incubating with proper substrates.

#### Measuring enzymatic activities

[0082] In some embodiments, a mutation introduced in one or more of the genes, e.g., *trx4*, does not abolish protein expression or mRNA expression, but results in a mutein that lacks the activity that the corresponding wild type protein possesses, e.g., thioredoxin activity of TrxA. It is understood by one of skill in the art that sometimes knocking out a gene does not require completely abolishing its activity; thus for purpose of this disclosure, lacking an activity means the mutant protein (“mutein”) loses 85–100% of the activity of the control protein, e.g., the wild type protein. The various muteins generated can be tested to confirm that they lack the activity of the wild type protein. For example, each of the coding sequences for the muteins can be separately expressed in a host strain, and the muteins are purified and tested for their activities, as described below.

#### Confirming the loss of thioredoxin reductase activity.

[0083] In one embodiment, the thioredoxin reductase activity can be measured by its activity in reducing 5,5-Dithiobis(2-nitrobenzoic acid) (DTNB) in the presence of NADPH. The reaction is typically started by mixing DTNB with thioredoxin reductase (TrxR), thioredoxin (Trx), and NADPH. The increase in absorbance at 412 nm are observed over time. The activity of the enzyme can be defined as the rate of absorbance increase. An embodiment of detecting thioredoxin reductase activity is disclosed in U.S. Pat. No. 8,592,468, the relevant portion of the disclosure is herein incorporated by reference.

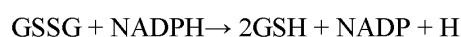
#### Confirming the loss of thioredoxin activity

[0084] Methods for determining thioredoxin activity are also well known. In one embodiment, the assay is an insulin precipitation assay, such as described by Sung-Jong Jeon et al., European Journal of Biochemistry, Vol. 269, No. 22. Thioredoxins are known to possess an activity as disulfide reductases of insulin; and reduction of insulin disulfide bonds can be measured by the increase in turbidity due to precipitation of the free insulin B-chain. In one illustrative example, a standard assay mixture contains 0.1M potassium phosphate (pH 7.0), 1 mM EDTA, and 0.13

mM bovine insulin in the absence or in the presence of the recombinant protein, and the reaction was initiated upon the addition of 1mM dithiothreitol. An increase of the absorbance at 650 nm was monitored at 30°C.

Confirming the loss of glutathione reductase activity

[0085] The glutathione reductase activity of the AhpC\* or the loss of the glutathione reductase activity of the mutein GOR can also be monitored. In some embodiments, the glutathione reductase activity is measured by its activity in reducing cysteine. For example, cysteine were incubated with reduction solution containing the candidate protein, e.g., AhpC\* or the mutein GOR, in the presence of cofactors. Preferably the cofactor is a coenzyme. Preferably the cofactor is nicotinamide adenine dinucleotide phosphate (NADPH) or nicotinamide adenine dinucleotide (NADH). The reaction reduces cystine to cysteine. The schematic reaction is as follows:



The activity can be measured by measuring the production of cysteine. In one particular example, the glutathione reductase activity in terms of reducing cysteine is described in WO2018114576.

Confirming the loss of peroxyreductase activity of AhpC

[0086] The lack of peroxyreductase activity of AhpC\* can be confirmed by incubating the protein with organic hydroperoxides or hydrogen peroxide in the presence of NADH. A functional peroxyreductase would convert these substrates into water in an NADH-dependent mechanism; a lack of evidence of such conversion indicates a lack of peroxyreductase activity.

Maintaining an oxidative cytoplasm by inhibiting reductase activity

[0087] Alternatively, the redox potential of the cytoplasm of the *E. coli* cell can be modulated by contacting the cell with an agent that is capable of inhibiting the activity of the one or more reductases in the cell. The redox potential of the cytoplasm can be adjusted to an oxidative state that can promote disulfide bond formation and protein stability. Agents that are capable of inhibiting the one or more reductases as disclosed above are well known, for example, aurothioglucose, aurothiomalate or CuSO<sub>4</sub>.

[0088] Suitable reductase inhibitors can also be identified by screening a library of compounds in an *in vitro* enzymatic assay. The activities of a target reductase in the presence or absence of a candidate compound can be measured and compared, and the candidate compound can be selected as the reductase inhibitor if it demonstrated an ability to reduce the reductase activity in the assay.

Additional properties of the *E. coli* strain that may promote antibody yield

[0089] In some embodiments, the *E. coli* strain further expresses a cytosolic prokaryotic disulfide isomerase, which can facilitate protein folding and further improve antibody yield. In one embodiment, the disulfide isomerase is DsbC. In one embodiment, the disulfide isomerase is a yeast protein disulfide isomerase (yPDI), e.g., as described in Groff et al., MAbs 6(3): 671-678 (2014). In some embodiments, the *E. coli* strain further expresses a disulfide oxidase to help promote disulfide formation and enhance IgG assembly. In one embodiment, the disulfide oxidase is DsbA. In one embodiment, the disulfide reductase is Qsox. In some embodiments, the protein isomerase is a prolyl isomerase. Suitable prolyl isomerase may be include, but are not limited to, FkpA, cyclophilin, FKBP, parvulin, or slyD. deaggregase skP or groEL/groES, danK, dnaJ, or grpE.

Confirming the cytosol of the *E. coli* strain is at an oxidative state

[0090] Whether the *E. coli* strain has an oxidative cytoplasm can be confirmed by measuring the redox potential of the cytoplasm, as disclosed above. Alternatively, the oxidative status of the cytoplasm can be verified by assessing the biological activity of a test protein having one or more disulfide bonds, for example, an LC that is difficult to express in a wild type *E. coli* strain. Preferred test polypeptides or proteins are those which are normally secreted from cells or which are membrane proteins. In some cases, these polypeptides are modified by the deletion or mutation of the signal sequence, such that the proteins are not exported outside of the cytoplasm of the cell.

[0091] As one illustrative example, a coding sequence for the LC of the B10 antibody (SEQ ID NO: 2) can be engineered into an expression cassette under a suitable promoter and transformed into the modified *E. coli* strain. The soluble protein fraction that contains the LC is measured. A suitable *E. coli* strain will be able to express in a soluble form of at least 1 mg/100 mL of the LC. Methods for preparing a bacterial lysate and measuring the amount of protein expression (e.g.,

the expression of LC) in the lysate are well known. In some embodiments, the *E.coli* cells can be treated with a lysis agent to produce a lysate. Cytoplasmic proteins can be released by treating the lysate with enzymes, such as benzonase and egg white lysozyme. The insoluble protein fraction can be separated from the soluble fraction by, e.g., centrifugation. The soluble protein fraction (containing the LC) can be collected and analyzed by SDS-PAGE. The amount of LC protein in the soluble protein fraction can then be quantified by, e.g., densitometry.

#### IV. Antibodies of Interest

[0092] The methods provided herein can be used for recombinant production of any full-length antibodies in *E. coli* cells. Disulfide bonds are present in antibodies and thus the present system are beneficial to prevent degradation and increase yield of the antibody. Any antibodies can be produced using the method described herein in a yield, referring to the amount of protein per liter of culture medium, that is at least about 200 mg/L, at least about 250 mg/L, at least about 500 mg/L, at least about 750 mg/L, or at least about 1000 mg/L. In some embodiments, the antibody is selected from the group consisting of a B10 antibody, an H01 antibody, a 7219 antibody (an anti-CD74 antibody), an anti-PD1 antibody, an anti-Tim3 antibody, an anti-HER2 antibody (e.g., trastuzumab), and an anti LAG3 antibody. B10 and H01 are different antibodies but both target the Folate receptor. Polynucleotides encoding the HC and LC of the antibodies can be introduced into the host cells to express the HC and LC, which can then be assembled to form full -length antibodies. Non-limiting examples of the antibodies are listed in Table 1.

Table 1. Exemplary antibodies

| Antibody name | HC (SEQ ID NO) | LC (SEQ ID NO) |
|---------------|----------------|----------------|
| B10           | 1              | 2              |
| H01           | 24             | 2              |
| 7209          | 25             | 9              |
| anti-PD1      | 11             | 12             |
| anti-Tim3     | 13             | 14             |

|             |    |    |
|-------------|----|----|
| anti LAG3   | 15 | 16 |
| Trastuzumab | 27 | 2  |

## V. Promoters

[0093] Promoters that can be used in the methods of the invention to drive the transcription of the HC or LC may be any appropriate promoter sequence suitable for *E. coli*. Such promoters may include mutant, truncated, and hybrid promoters, and may be obtained from polynucleotides encoding extracellular or intracellular polypeptides either endogenous (native) or heterologous (foreign) to the cell. A promoter used herein may be a constitutive or an inducible promoter.

[0094] In some embodiments, the promoter is a constitutive promoter. The promoter may be one that has substantially the same promoter strength as T7, i.e., the strength of the promoter is at least 60%, at least 70%, at least 80% of the strength of T7 promoter. In some embodiments, the T7 promoter comprises or consists of a sequence of SEQ ID NO: 19. In some embodiments, the promoter may exhibit a strength that is within the range of 50–200%, e.g., 80%–150% or 90%–140%, of the strength of T7 promoter. Non-limiting examples of promoters that are suitable for driving the transcription of the HC and LC in the *E. coli* cells include a T3 promoter, an SP6 promoter, a pBad, an XylA, or a PhoA promoter.

[0095] The strength of two promoters can be compared by comparing the amount of transcription of a same gene product initiated by the two promoters. For example, host cells containing an expression construct with the promoter to be tested (“test host cells”) and control host cells containing a control expression construct (containing the reference promoter, e.g., T7 promoter) can be grown in culture in replicates. The total RNA of the host cells and controls can be extracted and measured by absorbance at 260 nm. cDNA can then be synthesized from the equal amount of total RNA from the test host cells and the control host cells. RT-PCR can be performed to amplify the cDNA corresponding to the transcript produced from the promoter. An exemplary method is described in De Mey et al. (“Promoter knock-in: a novel rational method for the fine tuning of genes”, BMC Biotechnol 2010 Mar 24; 10:26).

[0096] In some embodiments, the promoters that drive the expression of the HC and LC are the same, for example, both are T7 promoters. In some cases, the HC and LC are directed by

one singular promoter, e.g., one single T7 promoter in a bicistronic operon, as disclosed below. In some embodiments, the promoters for the HC and LC are different promoters. In some embodiments, the promoter strength of the HC and LC are substantially the same. In some embodiments, the promoter strength of the HC and LC are different.

## VI. Vector

[0097] Polynucleotides encoding an LC and an HC of the antibody can be inserted into one or two replicable vectors for expression in the *E. coli*. Many vectors are available for this purpose, and one skilled in the art can readily select suitable vectors for use in the methods disclosed herein. Besides the gene of interest and promoter that drives the expression of the gene, the vector typically comprises one or more of the following: a signal sequence, an origin of replication, or one or more marker genes.

[0098] In some embodiments, a bicistronic vector, which employs a promoter to transcribe both the HC and LC coding sequences. In some embodiments, the HC coding sequence is proximal to the promoter. In some embodiments, the LC coding sequence is proximal to the promoter. Any of promoters disclosed herein can be used in the bicistronic vector. In some embodiments, the promoter is the T7 promoter. In some embodiments, the bicistronic vector further comprises a transcriptional terminator, such as the T7 terminator. The HC and the LC coding sequences may be separated by 15 to 30 nucleotides, e.g., 20 to 24 nucleotides, that is, the two closest nucleotides, one on the HC and the other on the LC, are separated by 15 to 30 nucleotides, e.g., 20 to 24 nucleotides. In some embodiments, the HC and the LC coding sequence are separated by a nucleotide sequence comprising a sequence that encodes a ribosomal binding site.

[0099] In some embodiments, the plasmid encoding the HC and LC comprises two monocistronic operons, one for the HC and the other for the LC production. In some embodiments, the two monocistronic operons share the same promoter, e.g., a T7 promoter or a T5 promoter. In some embodiments, the two monocistronic operons comprise different promoters. i.e., the promoter in the monocistronic operon for LC and the promoter in the monocistronic operon for the HC are of different sequence and/or having different promoter activity. Any of the promoters disclosed herein can be used in the vector comprising the two

monocistronic operons. In some embodiments, each monocistronic operon further comprises a transcriptional terminator, such as the T7 terminator.

**[0100]** In some embodiments, the translation of the HC is modulated by a first ribosomal binding site and the translation of the LC is modulated by a second ribosomal binding site. The ribosomal binding sites are independently selected to ensure the amount of expression of the HC and LC is at a desired ratio, for example, to ensure the molar ratio of the HC to LC is within the range of 1:1 to 1:3, e.g., about 2.5. This approach can boost the yield of the full-length antibody that are formed by the HC and LC produced in the *E. coli* culture. The ratio of the HC to LC can be determined based on the amount of HC and LC produced in the cell lysate obtained from the *E. coli* culture. The amount of the HC and LC can be determined using any methods suitable for protein quantification. In one illustrative embodiment, the *E. coli* culture is collected and lysed to prepare cell lysates. Optionally, the cell lysates are purified by, e.g., chromatography. The cell lysates or purified cell lysates are treated with a reducing agent (e.g., DTT) to separate the HC and LC. The reduced cell lysate sample can then be analyzed on a SDS-PAGE gel, in which HC and LC are separated and resolved by size. The gel is then blotted using an antibody that recognizes the HC and/or LC. The signal associated with binding between the antibody and HC or LC can be quantified. The signal is proportionally related to the amount of HC or LC produced in the culture.

**[0101]** In some embodiments, the first ribosomal binding site or the second ribosomal binding site comprises SEQ ID NO: 28. In some embodiments, the second ribosomal binding site comprises SEQ ID NO: 29. In some embodiments, the first ribosomal binding site comprises SEQ ID NO: 28 and the second ribosomal binding site comprises SEQ ID NO: 29. In some embodiments, the first ribosomal binding site comprises a sequence that is selected from the group consisting of SEQ ID NO: 17-19, and the second ribosomal binding site comprises a sequence that is selected from the group consisting of SEQ ID NO: 20-23; see Example 13. Tables 2 and 3. In some embodiments, the first ribosomal binding site comprises a sequence selected from the group consisting of SEQ ID NO: 20-22, and the second ribosomal binding site comprises SEQ ID NO: 17-19 (e.g., SEQ ID NO: 17); see Example 24. In some embodiments, the second ribosomal binding site comprises SEQ ID NO: 21. In some embodiments, the second ribosomal binding site comprises SEQ ID NO: 22. In some embodiments, the first ribosomal

binding site comprises SEQ ID NO: 17 or SEQ ID NO: 18. Any of the sequences disclosed for the first ribosomal binding site can be used as the second ribosomal binding site so long as the amount of expression of the HC and LC can be maintained at a desired ratio as described above.

[0102] In some embodiments, the antibody is selected from the group consisting of SP7219 IgG, B10 IgG, H01 IgG,  $\alpha$ PD1 IgG,  $\alpha$ Tim3 IgG,  $\alpha$ LAG3 IgG, and variants thereof.

## VII. Antibodies comprising non-natural amino acids

[0103] Antibodies produced using the methods may comprise at least one non-natural amino acid. In some embodiments, the non-natural amino acids are incorporated in the HC or LC at specific sites in the protein. These non-natural amino acids typically have bio-orthogonal reactive chemical side chains which can be used for conjugation with various biologically active adduct. See below. This approach can introduce additional functionality to the antibody, such as fluorescent or radioactive labels, photoactivatable markers, pharmacokinetic modifying PEGs, or chemotherapeutic agents.

[0104] The one or more non-natural amino acids can be located at selected site-specific positions in the heavy chain, or light chain, or both, of the antibody. The site-specific position can be in any domain of the antibody, including any variable domain and any constant domain. The number of non-natural amino acids present in the LC, HC, or both may vary, ranging from 1-10, e.g., from 2-8, from 3-6, from 5-10, or from 3-7. Heavy chains comprising one or more non-natural amino acids

[0105] In certain embodiments, the antibodies provided herein comprise one or more non-natural amino acids, each replacing a natural amino acid at a position selected from the group consisting of heavy chain residues HC-F404, HC-K121, HC-Y180, HC-F241, HC-221, HC-S136, HC-S25, HC-A40, HC-S119, HC-S190, HC-K222, HC-R19, HC-Y52, or HC-S70. In these designations, HC indicates a heavy chain residue. Light chains comprising one or more non-natural amino acids

[0106] In certain embodiments, the antibodies provided herein comprise one or more non-natural amino acids, each replacing a natural amino acid at a position selected from LC-K42,

LC-E161, LC-T22, LC-S7, LC-N152, LC-K42, LC-E161, or LC-D170. In these designations, LC indicates a light chain residue.

[0107] An antibody disclosed herein may also contain non-natural amino acids at any combination of the positions in HC and LC as described above.

Illustrative antibodies comprising non-natural amino acids

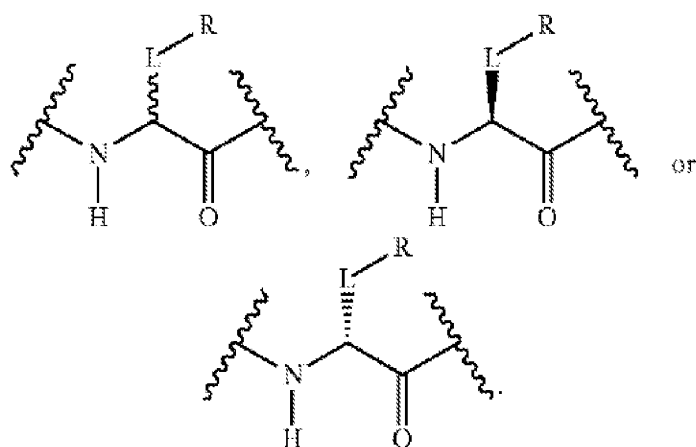
[0108] In one illustrative embodiment, the antibody is a B10 antibody comprising an HC having one or mutations selected from positions HC-F404, HC-Y180, HC-F241 relative to SEQ ID NO: 1, and positions LC-K42 relative to SEQ ID NO: 5. The mutations result in the amino acid codons at these positions being replaced with one or more non-natural amino acid codons. In one embodiment, the non-natural amino acid codon is the amber codon. In some embodiments, the B10 antibody comprises an HC having a sequence selected from the group consisting of SEQ ID NO: 3, 4, 6, and 7; and/or an LC having a sequence of SEQ ID NO: 5. In another illustrative embodiment, the antibody is trastuzumab, which comprises an LC, in which at least one of LC-E42 and LC-E161 is replaced by a non-natural amino acid.

Non-natural amino acids

[0109] Suitable non-natural amino acids that can be incorporated in the antibodies include, for example, those disclosed in U.S. Pat. No. 10,179,909; U. S. Pat. No. 9,938,516; U. S. Pat. No. 9,682,934; U. S. Pat. No. 10,596,270; and U.S. Pat. No. 10,610,571, the entire contents of which are herein incorporated by reference.

[0110] The non-natural amino acid may comprise a reactive group useful for forming a covalent bond to a linker or a biologically active adduct (aka a payload), as described below. In certain embodiments, the reactive group is selected from the group consisting of amino, carboxy, acetyl, hydrazino, hydrazido, semicarbazido, sulfanyl, azido and alkynyl. The non-natural amino acids may be L-amino acids, or D-amino acids, or racemic amino acids. In certain embodiments, the non-natural amino acids described herein include D-versions of the natural amino acids and racemic versions of the natural amino acids.

[0111] In certain embodiments, the non-natural amino acid is according to any of the following formulas:



[0112] In the above formulas, the wavy lines indicate bonds that connect to the remainder of the polypeptide chains of the antibodies. These non-natural amino acids can be incorporated into polypeptide chains just as natural amino acids are incorporated into the same polypeptide chains. In certain embodiments, the non-natural amino acids are incorporated into the polypeptide chain via amide bonds as indicated in the formulas. In the above formulas, R designates any functional group without limitation, so as long as the amino acid residue is not identical to a natural amino acid residue. In certain embodiments, R can be a hydrophobic group, a hydrophilic group, a polar group, an acidic group, a basic group, a chelating group, a reactive group, a therapeutic moiety or a labeling moiety. In the above formulas, each L represents a linker (e.g., a divalent linker), as further described below.

[0113] In some embodiments, the non-naturally encoded amino acids include side chain functional groups that react efficiently and selectively with functional groups not found in the 20 common amino acids (including but not limited to, azido, ketone, aldehyde and aminooxy groups) to form stable conjugates. For example, antigen-binding polypeptide that includes a non-naturally encoded amino acid containing an azido functional group can be reacted with a polymer (including but not limited to, poly(ethylene glycol)) or, alternatively, a second polypeptide containing an alkyne moiety to form a stable conjugate resulting for the selective reaction of the azide and the alkyne functional groups to form a Huisgen [3+2] cycloaddition product.

[0114] Exemplary non-naturally encoded amino acids that may be suitable for use in the present invention and that are useful for reactions with water soluble polymers, including but are not limited to, those with carbonyl, aminoxy, hydrazine, hydrazide, semicarbazide, azide and alkyne reactive groups. In some embodiments, non-naturally encoded amino acids comprise a saccharide moiety. Examples of such amino acids include N-acetyl-L-glucosaminyl-L-serine, N-acetyl-L-galactosaminyl-L-serine, N-acetyl-L-glucosaminyl-L-threonine, N-acetyl-L-glucosaminyl-L-asparagine and O-mannosaminyl-L-serine. Examples of such amino acids also include examples where the naturally-occurring N- or O-linkage between the amino acid, and the saccharide is replaced by a covalent linkage not commonly found in nature-including but not limited to, an alkene, an oxime, a thioether, an amide and the like. Examples of such amino acids also include saccharides that are not commonly found in naturally-occurring proteins such as 2-deoxy-glucose, 2-deoxygalactose and the like.

[0115] Many of the non-natural amino acids suitable for use in the present invention are commercially available, e.g., from Sigma (USA) or Aldrich (Milwaukee, Wis., USA). Those that are not commercially available are optionally synthesized as provided herein or as provided in various publications or using standard methods known to those of skill in the art. For organic synthesis techniques, see, e.g., Organic Chemistry by Fessenden and Fessenden, (1982, Second Edition, Willard Grant Press, Boston Mass.); See, also, U.S. Patent Application Publications 2003/0082575 and 2003/0108885, which is incorporated by reference herein; Advanced Organic Chemistry by March (Third Edition, 1985, Wiley and Sons, New York); and Advanced Organic Chemistry by Carey and Sundberg (Third Edition, Parts A and B, 1990, Plenum Press, New York). Additional publications describing the synthesis of non-natural amino acids include, e.g., WO 2002/085923 entitled "In vivo incorporation of Non-natural Amino Acids;" Matsoukas et al., (1995) J. Med. Chem., 38, 4660-4669; King, F. E. & Kidd, D. A. A. (1949) A New Synthesis of Glutamine and of  $\gamma$ -Dipeptides of Glutamic Acid from Phthylated Intermediates. J. Chem. Soc., 3315-3319; Friedman, O. M. & Chatterji, R. (1959) Synthesis of Derivatives of Glutamine as Model Substrates for Anti-Tumor Agents. J. Am. Chem. Soc. 81, 3750-3752; Craig, J. C. et al. (1988) Absolute Configuration of the Enantiomers of 7-Chloro-4 [[4-(diethylamino)-1-methylbutyl]amino]quinoline (Chloroquine). J. Org. Chem. 53, 1167-1170; Azoulay, M., Vilmont, M. & Frappier, F. (1991) Glutamine analogues as Potential Antimalarials, Eur. J. Med. Chem. 26, 201-5; Koskinen, A. M. P. & Rapoport, H. (1989) Synthesis of 4-Substituted Prolines

as Conformationally Constrained Amino Acid Analogues. *J. Org. Chem.* 54, 1859-1866; Christie, B. D. & Rapoport, H. (1985) Synthesis of Optically Pure Pipecolates from L-Asparagine. Application to the Total Synthesis of (+)-Apovincamine through Amino Acid Decarbonylation and Iminium Ion Cyclization. *J. Org. Chem.* 1989:1859-1866; Barton et al., (1987) Synthesis of Novel  $\alpha$ -Amino-Acids and Derivatives Using Radical Chemistry: Synthesis of L- and D- $\alpha$ -Amino-Adipic Acids, L- $\alpha$ -aminopimelic Acid and Appropriate Unsaturated Derivatives. *Tetrahedron Lett.* 43:4297-4308; and, Subasinghe et al., (1992) Quisqualic acid analogues: synthesis of beta-heterocyclic 2-aminopropanoic acid derivatives and their activity at a novel quisqualate-sensitized site. *J. Med. Chem.* 35:4602-7. See also, patent applications entitled "Protein Arrays," filed Dec. 22, 2003, Ser. No. 10/744,899 and Ser. No. 60/435,821 filed on Dec. 22, 2002.

**[0116]** Many non-natural amino acids are based on natural amino acids, such as tyrosine, glutamine, phenylalanine, and the like and are suitable for use in the present invention. Tyrosine analogs include, but are not limited to, para-substituted tyrosines, ortho-substituted tyrosines, and meta substituted tyrosines, where the substituted tyrosine comprises, including but not limited to, a keto group (including but not limited to, an acetyl group), a benzoyl group, an amino group, a hydrazine, an hydroxyamine, a thiol group, a carboxy group, an isopropyl group, a methyl group, a C6-C20 straight chain or branched hydrocarbon, a saturated or unsaturated hydrocarbon, an O-methyl group, a polyether group, a nitro group, an alkynyl group or the like. In addition, multiply substituted aryl rings are also contemplated. Glutamine analogs that may be suitable for use in the present invention include, but are not limited to,  $\alpha$ -hydroxy derivatives,  $\gamma$ -substituted derivatives, cyclic derivatives, and amide substituted glutamine derivatives. Example phenylalanine analogs that may be suitable for use in the present invention include, but are not limited to, para-substituted phenylalanines, ortho-substituted phenylalanines, and meta-substituted phenylalanines, where the substituent comprises, including but not limited to, a hydroxy group, a methoxy group, a methyl group, an allyl group, an aldehyde, an azido, an iodo, a bromo, a keto group (including but not limited to, an acetyl group), a benzoyl, an alkynyl group, or the like. Specific examples of non-natural amino acids that may be suitable for use in the present invention include, but are not limited to, an azidoethoxycarbonyl lysine (AEK), a p-acetyl-L-phenylalanine, an O-methyl-L-tyrosine, an L-3-(2-naphthyl)alanine, a 3-methyl-phenylalanine, an O-4-allyl-L-tyrosine, a 4-propyl-L-tyrosine, a tri-O-acetyl-GlcNAc $\beta$ -serine, an

L-Dopa, a fluorinated phenylalanine, an isopropyl-L-phenylalanine, a p-azido-L-phenylalanine, a p-azido-methyl-L-phenylalanine, a p-acyl-L-phenylalanine, a p-benzoyl-L-phenylalanine, an L-phosphoserine, a phosphoserine, a phosphotyrosine, a p-iodo-phenylalanine, a p-bromophenylalanine, a p-amino-L-phenylalanine, an isopropyl-L-phenylalanine, and a p-propargyloxy-phenylalanine, and the like. Examples of structures of a variety of non-natural amino acids that may be suitable for use in the present invention are provided in, for example, WO 2002/085923 entitled "In vivo incorporation of non-natural amino acids." See also Kiick et al., (2002) Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation, PNAS 99:19-24, for additional methionine analogs.

[0117] Particular examples of useful non-natural amino acids include, but are not limited to, p-acetyl-L-phenylalanine, O-methyl-L-tyrosine, L-3-(2-naphthyl)alanine, 3-methyl-phenylalanine, O-4-allyl-L-tyrosine, 4-propyl-L-tyrosine, tri-O-acetyl-GlcNAc b-serine, L-Dopa, fluorinated phenylalanine, isopropyl-L-phenylalanine, p-azido-methyl-L-phenyl alanine, p-azido-L-phenylalanine, p-acyl-L-phenylalanine, p-benzoyl-L-phenylalanine, L-phosphoserine, phosphoserine, phosphotyrosine, p-iodo-phenylalanine, p-bromophenylalanine, p-amino-L-phenylalanine, isopropyl-L-phenylalanine, and p-propargyloxy-phenylalanine. Further useful examples include N-acetyl-L-glucosaminyl-L-serine, N-acetyl-L-galactosaminyl-L-serine, N-acetyl-L-glucosaminyl-L-threonine, N-acetyl-L-glucosaminyl-L-asparagine and O-mannosaminyl-L-serine.

[0118] In particular embodiments, the non-natural amino acids are selected from p-acetyl-phenylalanine, p-ethynyl-phenylalanine, p-propargyloxyphenylalanine, p-azido-methyl-phenylalanine, and p-azido-phenylalanine. In one embodiment, the non-natural amino acid is p-azido phenylalanine. This amino acid residue is known to those of skill in the art to facilitate Huisgen [3+2] cycloaddition reactions (so-called "click" chemistry reactions) with, for example, compounds bearing alkynyl groups. This reaction enables one of skill in the art to readily and rapidly conjugate to the antibody at the site-specific location of the non-natural amino acid.

[0119] In certain embodiments, the first reactive group is an alkynyl moiety (including but not limited to, in the non-natural amino acid p-propargyloxyphenylalanine, where the propargyl group is also sometimes referred to as an acetylene moiety) and the second reactive group is an azido moiety, and [3+2] cycloaddition chemistry can be used. In certain embodiments, the first

reactive group is the azido moiety (including but not limited to, in the non-natural amino acid p-azido-L-phenylalanine) and the second reactive group is the alkynyl moiety.

**[0120]** The non-natural amino acids used in the methods and compositions described herein have at least one of the following four properties: (1) at least one functional group on the sidechain of the non-natural amino acid has at least one characteristics and/or activity and/or reactivity orthogonal to the chemical reactivity of the 20 common, genetically-encoded amino acids (i.e., alanine, arginine, asparagine, aspartic acid, cysteine, glutamine, glutamic acid, glycine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine, and valine), or at least orthogonal to the chemical reactivity of the naturally occurring amino acids present in the polypeptide that includes the non-natural amino acid; (2) the introduced non-natural amino acids are substantially chemically inert toward the 20 common, genetically-encoded amino acids; (3) the non-natural amino acid can be stably incorporated into a polypeptide, preferably with the stability commensurate with the naturally-occurring amino acids or under typical physiological conditions, and further preferably such incorporation can occur via an in vivo system; and (4) the non-natural amino acid includes an oxime functional group or a functional group that can be transformed into an oxime group by reacting with a reagent, preferably under conditions that do not destroy the biological properties of the polypeptide that includes the non-natural amino acid (unless of course such a destruction of biological properties is the purpose of the modification/transformation), or where the transformation can occur under aqueous conditions at a pH between about 4 and about 8, or where the reactive site on the non-natural amino acid is an electrophilic site. Any number of non-natural amino acids can be introduced into the polypeptide. Non-natural amino acids may also include protected or masked oximes or protected or masked groups that can be transformed into an oxime group after deprotection of the protected group or unmasking of the masked group. Non-natural amino acids may also include protected or masked carbonyl or dicarbonyl groups, which can be transformed into a carbonyl or dicarbonyl group after deprotection of the protected group or unmasking of the masked group and thereby are available to react with hydroxylamines or oximes to form oxime groups.

**[0121]** In further embodiments, non-natural amino acids that may be used in the methods and compositions described herein include, but are not limited to, amino acids comprising a

photoactivatable cross-linker, spin-labeled amino acids, fluorescent amino acids, metal binding amino acids, metal-containing amino acids, radioactive amino acids, amino acids with novel functional groups, amino acids that covalently or non-covalently interact with other molecules, photocaged and/or photoisomerizable amino acids, amino acids comprising biotin or a biotin analogue, glycosylated amino acids such as a sugar substituted serine, other carbohydrate modified amino acids, keto-containing amino acids, aldehyde-containing amino acids, amino acids comprising polyethylene glycol or other polyethers, heavy atom substituted amino acids, chemically cleavable and/or photocleavable amino acids, amino acids with an elongated side chains as compared to natural amino acids, including but not limited to, polyethers or long chain hydrocarbons, including but not limited to, greater than about 5 or greater than about 10 carbons, carbon-linked sugar-containing amino acids, redox-active amino acids, amino thioacid containing amino acids, and amino acids comprising one or more toxic moiety.

**[0122]** In some embodiments, non-natural amino acids comprise a saccharide moiety. Examples of such amino acids include N-acetyl-L-glucosaminyl-L-serine, N-acetyl-L-galactosaminyl-L-serine, N-acetyl-L-glucosaminyl-L-threonine, N-acetyl-L-glucosaminyl-L-asparagine and O-mannosaminyl-L-serine. Examples of such amino acids also include examples where the naturally-occurring N- or O-linkage between the amino acid and the saccharide is replaced by a covalent linkage not commonly found in nature-including but not limited to, an alkene, an oxime, a thioether, an amide and the like. Examples of such amino acids also include saccharides that are not commonly found in naturally-occurring proteins such as 2-deoxy-glucose, 2-deoxygalactose and the like.

**[0123]** In particular embodiments, the non-natural amino acid is one selected from the group of non-natural amino acids shown in FIG. 8A-8I. Such non-natural amino acids may be in the form of a salt or may be incorporated into a non-natural amino acid polypeptide, polymer, polysaccharide, or a polynucleotide and optionally post translationally modified.

**[0124]** In some embodiments, the non-natural amino acid is para-methylazido-L-phenylalanine (pAMF), Azidoethoxycarbonyl lysine (AEK), or p-acetyl-L-phenylalanine (pAcF). In some embodiments, the non-natural amino acid is (S)-2-amino-3-(5-(6-methyl-1,2,4,5-tetrazin-3-ylamino)pyridin-3-yl)propanoic acid.

Incorporation of non-natural amino acids

[0125] Methods for incorporating the non-natural amino acids are well known, e.g., as described in U.S. Pat. No. 10,610,571, U.S. Pat. No. 9,988,619 and U.S. Pat. No. 9,938,516, the entire contents of which are herein incorporated by reference.

[0126] In one approach, the coding sequence of the HC or LC is modified to contain at least one non-natural amino acid codon. The non-natural amino acid codon is one that does not result in the incorporation of any of the 20 natural amino acids. In some embodiments, the non-natural amino acid codon is an amber, opal, or ochre stop codon, which is repurposed to charge a non-natural amino acid to its cognate tRNA by a tRNA synthetase instead of terminating translation.

[0127] A non-natural amino acid can be charged to a tRNA by a tRNA synthetase, which preferentially acetylates the non-natural amino acid as compared to any of the 20 natural amino acids. tRNA synthetases having such function are known, for example, U.S. Pat. No. 9,938,516 discloses tRNA synthetases that selectively incorporate a non-natural amino acid para-methylazido-L-phenylalanine (pAMF). tRNA synthetases that can selectively incorporate other non-natural amino acids, for example, Azidoethoxycarbonyl lysine (AEK) or -acetyl-L-phenylalanine (pAcF), are also well known; *see*, for example, Chen *et al.*, *Angew Chem Int Ed Engl.* 2009; 48(22):4052-5 (doi: 10.1002/anie.200900683); and Li *et al.*, *Proc Natl Acad Sci U S A*, 2003 Jan 7;100(1):56-61. The entire contents of said publications are herein incorporated by reference. Additional exemplary non-natural amino acids that can be incorporated into the antibodies disclosed herein include aralkyl, heterocyclyl, and heteroaralkyl, and lysine-derivative unnatural amino acids. In some embodiments, such non-natural amino acid comprise pyridyl, pyrazinyl, pyrazolyl, triazolyl, oxazolyl, thiazolyl, thiophenyl, or other heterocycle. Such amino acids in some embodiments comprise azides, tetrazines, or other chemical group capable of conjugation to a coupling partner, such as a water soluble moiety.

[0128] tRNA synthetases that are capable of selectively incorporating a non-natural amino acid may be also be obtained by genetically modifying a wild type tRNA synthetase to produce mutant tRNA synthetases. Each of these mutant tRNA synthetases can then be tested for its activity in selectively incorporating the non-natural amino acid using in a reporter gene, which contains the desired non-natural amino acid codon. The activity of the mutant tRNA synthetase variant in the presence of the non-natural amino acid (e.g., pAMF) as compared to the 20

common naturally occurring amino acids can be measured by detecting the presence or absence of the reporter protein. One exemplary method of generating mutant tRNA synthetases for incorporating non-natural amino acids is disclosed in U.S. Pat. No. 9,938,516, the entire content of which is herein incorporated by reference.

[0129] In some embodiments, the non-natural amino acid codon is a synthetic codon, and the unnatural amino acid is incorporated into the antibody using an orthogonal synthetase/tRNA pair. The orthogonal synthetase may be a synthetase that is modified from any of the natural amino acid synthetases. For example, the orthogonal synthetase may be a proline synthetase, a modified serine synthetase, a modified tryptophan synthetase, or a modified phosphoserine synthetase. The orthogonal tRNA may also be modified from any of the natural amino acid tRNA. For example, the orthogonal tRNA may be a modified alanine tRNA, a modified arginine tRNA, a modified aspartic acid tRNA, a modified cysteine tRNA, a modified glutamine tRNA, a modified glutamic acid tRNA, a modified alanine glycine, a modified histidine tRNA, a modified leucine tRNA, a modified isoleucine tRNA, a modified lysine tRNA, a modified methionine tRNA, a modified phenylalanine tRNA, a modified proline tRNA, a modified serine tRNA, a modified threonine tRNA, or a modified tryptophan tRNA. In some embodiments, a modified tyrosine tRNA, a modified valine tRNA, or a modified phosphoserine tRNA.

#### Codon optimization

[0130] Codon optimization may be used to increase the rate of translation of the HC or LC or to produce recombinant RNA transcripts having desirable properties, such as a longer half-life, as compared with transcripts produced using a non-optimized sequence. The HC or LC coding sequences may be optimized to maximize expression efficiency in *E. coli*. In particular, when expressing an HC or an LC polynucleotide having a non-natural amino acid codon in *E. coli*, bases that are in the vicinity of the non-natural amino acid codon may have affect the mRNA conformation in the P site in the ribosome and thus have an impact on the efficiency of incorporating the non-natural amino acid. Thus, it is desirable to optimize the codons for the amino acid in the vicinity of the non-natural amino acid codon to maximize the yield of the antibodies having the non-natural amino acids. In some embodiments, the codon that is immediately 3' to the non-natural amino acid codon is optimized in order to maximize expression. The optimal codon can be selected by comparing the yield of HCs (or LC) produced

from expressing the HC (or LC) coding sequences having different codons for the same amino acid that is located immediately 3' to the non-natural amino acid codon and selecting the coding sequence that produces the highest yield. In one illustrative example, the coding sequence for the HC of the B10 antibody has an amber codon instead of a tyrosine (Y) codon at position 180. Position 181, which is located immediately 3' to the amber codon, is a serine. The HC coding sequence having a serine codon AGC or AGT at the position produced higher titer than the HC coding sequence having a serine codon TCG at the same position. See, Example 11, Figure 2.

### VIII. Methods

[0131] Unless defined otherwise, all technical and scientific terms used herein have the meaning commonly understood by one of ordinary skill in the art to which this invention belongs. Practitioners are particularly directed to Green, M.R., and Sambrook, J., eds., *Molecular Cloning: A Laboratory Manual*, 4th ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2012), and Ausubel, F. M., et al., *Current Protocols in Molecular Biology* (Supplement 99), John Wiley & Sons, New York (2012), which are incorporated herein by reference, for definitions and terms of the art. Standard methods also appear in Bindereif, Schön, & Westhof (2005) *Handbook of RNA Biochemistry*, Wiley-VCH, Weinheim, Germany which describes detailed methods for RNA manipulation and analysis, and is incorporated herein by reference. Examples of appropriate molecular techniques for generating recombinant nucleic acids and instructions sufficient to direct persons of skill through many cloning exercises are found in Green, M.R., and Sambrook, J., (*Id.*); Ausubel, F. M., et al., (*Id.*); Berger and Kimmel, *Guide to Molecular Cloning Techniques, Methods in Enzymology* (Volume 152 Academic Press, Inc., San Diego, Calif. 1987); and *PCR Protocols: A Guide to Methods and Applications* (Academic Press, San Diego, Calif. 1990), which are incorporated by reference herein.

[0132] PCR amplification methods are well known in the art and are described, for example, in Innis et al., *PCR Protocols: A Guide to Methods and Applications*, Academic Press Inc. San Diego, Calif., 1990. An amplification reaction typically includes the DNA that is to be amplified, a thermostable DNA polymerase, two oligonucleotide primers, deoxynucleotide triphosphates (dNTPs), reaction buffer and magnesium. Typically a desirable number of thermal cycles is between 1 and 25. Methods for primer design and optimization of PCR conditions are well known in the art and can be found in standard molecular biology texts such as Ausubel et al., *Short*

Protocols in Molecular Biology, 5<sup>th</sup> Edition, Wiley, 2002, and Innis et al., PCR Protocols, Academic Press, 1990. Computer programs are useful in the design of primers with the required specificity and optimal amplification properties (e.g., Oligo Version 5.0 (National Biosciences)). In some embodiments, the PCR primers may additionally contain recognition sites for restriction endonucleases, to facilitate insertion of the amplified DNA fragment into specific restriction enzyme sites in a vector. If restriction sites are to be added to the 5' end of the PCR primers, it is preferable to include a few (e.g., two or three) extra 5' bases to allow more efficient cleavage by the enzyme. In some embodiments, the PCR primers may also contain an RNA polymerase promoter site, such as T7 or SP6, to allow for subsequent *in vitro* transcription. Methods for *in vitro* transcription are well known to those of skill in the art (see, e.g., Van Gelder et al., Proc. Natl. Acad. Sci. U.S.A. 87:1663-1667, 1990; Eberwine et al., Proc. Natl. Acad. Sci. U.S.A. 89:3010-3014, 1992).

**[0133]** When the antibodies described herein are referred to by name, it is understood that this includes antibodies with similar functions and having HCs and LCs with similar amino acid sequences. For example, the name “B10 antibody” includes the wild-type antibody having an HC of a sequence of SEQ ID NO: 1, and an LC of a sequence of SEQ ID NO: 2, as well as antibodies having an HC and/or light chain having similar amino acid sequences to the SEQ ID NO: 1 or SEQ ID NO: 2. The sequence similarity may be at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%. The sequence identity of a protein is determined using the BLASTP program with the defaults wordlength of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff and Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915-10919, 1992).

#### Transformation

**[0134]** Vectors comprising the HC and LC coding sequences may be transfected into a host cell by standard techniques. The various forms of the term “transfection” are intended to encompass a wide variety of techniques commonly used for the introduction of exogenous DNA into a *E. coli* cell, e.g., electroporation, calcium-phosphate precipitation, DEAE-dextran transfection and the like. *E. coli* cells transformed with the vector can be cultured under suitable growth conditions to allow the expression of HC and LC. Cells are harvested and cell lysates are prepared using standard techniques.

#### Assembling the HC and LC

[0135] The cell lysates prepared from the *E. coli* culture can be adjusted to a condition to allow the two LC and two HC to assemble into a tetrameric form of the full-length antibody, including allowing the dimerization of the two heavy chain and association between the heavy chain and light chain through formation of the disulfide bonds. Typically, cell lysates are adjusted to a suitable pH (e.g., pH 8.0), and incubated for an extended time by shaking to allow the formation of disulfide bonds and the assembling of the full-length antibody. In one illustrative example, the cell lysate is adjusted to pH 8.0 with 100 mM Tris·HCl by incubating for 16 hrs in a FlowerPlate (m2p-labs) at 25°C with shaking at 650 rpm.

#### Purifying the full-length antibody

[0136] The assembled IgGs were purified from cell lysate using standard affinity chromatography methods. Various methods for protein purification, chromatography, electrophoresis, centrifugation, and crystallization are described in Coligan et al. (2000) Current Protocols in Protein Science, Vol. 1, John Wiley and Sons, Inc., New York. Methods for cell-free synthesis are described in Spirin & Swartz (2008) Cell-free Protein Synthesis, Wiley-VCH, Weinheim, Germany. The extent of the assembly of the HC and LC in the antibody can be determined using methods well known in the art, e.g., by a caliper bioanalyzer (Perkin Elmer, Richmond, CA) under non-reducing conditions.

#### Titer/Yield

[0137] The titer of the assembled full-length antibodies can be determined using standard methods such as, but not limited to, SDS-PAGE, western blotting, chromatography, e.g., analytical size exclusion chromatography, liquid chromatography and HPLC, immuno-based assays, and enzyme-linked immunosorbent assays (ELISAs). One illustrative example is shown in Example 22.

[0138] In some embodiments, especially when the *E. coli* cells are grown in high density for IgG production, for example, the OD<sub>600</sub> is in the range of 50–150, the titer of the full-length antibody (in terms of the amount of antibody per liter of culture medium) produced using the methods disclosed herein is greater than about 200 mg/L, about 250 mg/L, about 300 mg/L, about 350 mg/L, about 400 mg/L, about 450 mg/L, about 500 mg/L, about 550 mg/L, about 600 mg/L, about 650 mg/L, about 700 mg/L, about 750 mg/L, about 800 mg/L, about 850 mg/L,

about 900 mg/L, about 950 mg/L, about 1000 mg/L or more. In other embodiments, the protein of interest is produced at a concentration (amount or level) of greater than about 1000 mg/L, e.g., about 1100 mg/L, about 1200 mg/L, about 1300 mg/L, about 1400 mg/L, about 1500 mg/L, about 1600 mg/L, about 1700 mg/L, about 1800 mg/L, about 1900 mg/L, about 2000 mg/L, or more. Typically, high density growth is performed in bioreactors in a volume of 0.5 liter to 5 liters, one illustration of which is shown in Example 6. However, it is contemplated that the methods described herein may be performed at larger scales, including in culture volumes of at least 500 L, at least 1000 L, at least 2000 L, at least 2500 L, at least 3000 L, at least 5000 L, at least 8000L, at least 10,000 L, or greater.

[0139] In some embodiments, the titer of the full-length antibody produced using the methods disclosed herein is at a wet weight percentage in a range from 0.05% to 20%, from 0.05% to 10%, from 0.05% to 5%, from 0.05% to 4%, from 0.05% to 3%, from 0.05% to 2%, from 0.05% to 1%, or from 0.05% to 0.5% relative to the weight of a wet cell pellet produced from the *E. coli* cells in the culture. In terms of the lower end of the range, the titer of the full-length antibody produced is at a wet weight percentage that is at least 0.05%, at least 0.1%, at least 0.2%, at least 0.3%, at least 0.4%, at least 0.5%, or at least 0.6% relative to the weight of the *E. coli* cells in the culture.

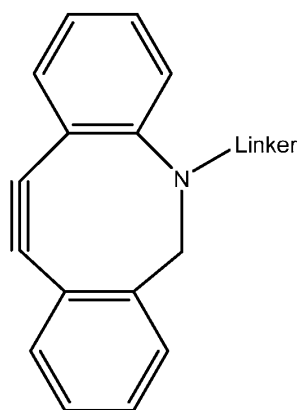
#### **IX. Antibody-drug conjugates**

[0140] Antibodies produced using the methods in this disclosure can be conjugated to a biologically active adduct (aka, a payload) using a chemical reaction such as click chemistry. In some cases, the antibody comprises one or more non-natural amino acids at specific sites in the protein sequence, and the biologically active adduct can be conjugated to these non-natural amino acids. Having antibodies containing the non-natural amino acids at the desired amino acid location, a biologically active adduct can be conjugated to the non-natural amino acid using a chemical reaction such as the click chemistry. For instance, the pAMF containing antibody produced using the methods disclosed herein can be purified by standard procedures. Then, the purified protein is subject to a click chemistry reaction (e.g., copper(I)-catalyzed azide-alkyne 1,3-cycloaddition reaction or copper-free catalyzed azide-alkyne 1,3-cycloaddition reaction) to directly conjugate a biologically active adduct to the pAMF residue.

[0141] Exemplary biologically active adducts for use in the present invention include, but are not limited to, small molecules, oligonucleotides, peptides, amino acids, nucleic acids, sugars, oligosaccharides, polymers, synthetic polymers, chelators, fluorophores, chromophores, other detectable agents, drug moieties, cytotoxic agents, detectable agents, and the like.

[0142] Detailed descriptions of a click chemistry reaction for conjugation of a biologically active adduct are found in, *e.g.*, Baskin *et al.*, *Proc. Natl. Acad. Sci.*, 2007, 104: 16793-16797; Kim *et al.*, *Curr. Opin. Chem. Biol.*, 2013, 14:412-419; and Bundy and Swartz, *Bioconjug. Chem.*, 2010, 21(2):255-263.

[0143] In a click chemistry reaction, alkynes are activated for [3+2] cycloaddition with azides. The biologically active adduct containing (*e.g.*, linked to) a strained alkyne (*e.g.*, cyclooctyne or variant thereof) can undergo strain-promoted alkyne-azide cycloaddition with the non-natural amino acid para-methylazido-L-phenylalanine on the protein of interest, thereby conjugating the biologically active adduct to the protein at the amino acid position of the non-natural amino acid. A preferred strained alkyne reagent is the reagent DBCO, shown below.



#### Linker

[0144] In certain embodiments, the antibodies can be linked to the biologically active adduct with one or more linkers capable of reacting with an antibody amino acid (*e.g.*, an non-natural amino acid) and with the biologically active adduct. The term “linker” is used to refer to groups or bonds that normally are formed as the result of a chemical reaction and typically are covalent linkages. A linker used in this disclosure can be any linkers apparent to those of skill in the art. In certain embodiments, the linker is any divalent or multivalent linker known to those of skill in

the art. Useful divalent linkers include alkylene, substituted alkylene, heteroalkylene, substituted heteroalkylene, arylene, substituted arylene, heteroarylne, and substituted heteroarylene. In certain embodiments, the linker is C1-10 alkylene or C1-10 heteroalkylene. Suitable linkers are also disclosed in U.S. Pat. No. 10,596,270, the entire content of which is herein incorporated by reference.

[0145] The conjugated protein of interest can be purified according to standard methods known in the art including, but not limited to, chromatography (*e.g.*, ion exchange, affinity, hydrophobic, chromatofocusing, and size exclusion), electrophoretic procedures (*e.g.*, preparative isoelectric focusing), differential solubility (*e.g.*, ammonium sulfate precipitation), SDS-PAGE, or extraction (*see, e.g.*, Protein Purification, J. -C. Janson and Lars Ryden, editors, VCH Publishers, New York, 1989) to obtain substantially pure polypeptides.

[0146] The conjugated protein of interest can be quantitated according to standard methods known in the art including, but not limited to, mass spectrometry (*e.g.*, ESI-TOF mass spectrometry and tandem mass spectrometry), microfluidic electrophoresis, gel electrophoresis, Western blotting, immunoassays (*e.g.*, ELISA), and other assays to assess the activity of the conjugated protein.

## **X. Use**

[0147] Antibodies produced by the invention, including those incorporating non-natural amino acids, can be used for one or more of the following purposes or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in form or shape); effecting biorhythms or circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation

and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, the ability to bind antigens or complement); and the ability to act as an antigen in vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

**[0148]** The antibodies produced by the invention can be used for any purpose known to one of skill in the art. Preferred uses include medical uses, including diagnostic uses, prophylactic, and therapeutic uses. For example, the antibodies can be prepared for topical or other type of administration. Accordingly, the proteins produced by the invention are solubilized or suspended in pharmacologically acceptable solutions to form pharmaceutical compositions for administration to a subject. Appropriate buffers for medical purposes and methods of administration of the pharmaceutical compositions are further set forth below. It will be understood by a person of skill in the art that medical compositions can also be administered to subjects other than humans, such as for veterinary purposes.

**[0149]** It is understood that the examples and embodiments described herein are for illustrative purposes only, and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and scope of the appended claims. All publications, sequence accession numbers, patents, and patent applications cited herein are hereby incorporated by reference in their entirety for all purposes.

### **Embodiments**

This disclosure include the following exemplary embodiments:

**[0150]** Embodiment 1. A method to produce a full-length antibody comprising a heavy chain (HC) and a light chain (LC), wherein the method comprises:

culturing *E. coli* cells expressing a coding sequence for the HC and a coding sequence for the LC under conditions permissible for producing the HC and LC in a culture medium, and

wherein the full-length antibody is produced in an amount of at least about 200 mg per liter of the culture medium, or

wherein the weight percentage of the full-length antibody produced relative to the weight of cell pellet produced from the *E. coli* cells is in a range from 0.05% to 20%.

**[0151]** Embodiment 2. The method of embodiment 1, wherein at least some of the *E. coli* cells contain an oxidative cytoplasm.

**[0152]** Embodiment 3. The method of any one of the embodiments above, wherein the method comprises transforming the plasmid to the *E. coli* strain.

**[0153]** Embodiment 4. The method of any one of the embodiments above, wherein the HC and LC are produced in the cytoplasm of the *E. coli*.

**[0154]** Embodiment 5. The method of any one of the embodiments above, wherein the method further comprises assembling the produced HC and LC under non-reducing conditions to form the full-length antibody.

**[0155]** Embodiment 6. The method of any one of the embodiments above, wherein the molar ratio of the produced HC and the produced LC from the *E. coli* is about 1:1 to about 1:3.

**[0156]** Embodiment 7. The method of any one of claims 1-6, wherein expression of HC, LC, or both is controlled by a promoter,

wherein the promoter is a T7 promoter or a promoter that has substantially similar promoter strength as the T7 promoter, or

wherein the promoter is a T5 promoter or a promoter that has substantially similar promoter strength as the T5 promoter.

**[0157]** Embodiment 8. The method of any one of the embodiments above, wherein the plasmid comprises a bicistronic operon, wherein the bicistronic operon comprises the coding sequence for the HC and the coding sequence for LC, or

wherein the plasmid comprises a first monocistronic operon for the HC and a second monocistronic operon for LC.

**[0158]** Embodiment 9. The method of embodiment 8, wherein the bicistronic operon comprises a promoter that drives expression of both the HC and LC, and wherein the promoter is a T7 promoter or a promoter that has substantially similar promoter strength as the T7 promoter, or

wherein the first monocistronic operon or the second monocistronic operon comprise a T7 promoter.

**[0159]** Embodiment 10. The method of any one of the embodiments above, wherein the bicistronic operon comprises a T7 terminator, or

wherein the first monocistronic operon or the second monocistronic operon comprises a T7 terminator.

**[0160]** Embodiment 11. The method of any of claims 1-10, wherein the *E. coli* cells comprises a first ribosome-binding site for translation of the HC and a second ribosomal binding site for translation of LC, wherein the first and second ribosomal binding sites are selected such that the molar ratio of HC and LC produced from the *E. coli* ranges from 1:1 to 1:3.

**[0161]** Embodiment 12. The method of embodiment 11, wherein the first ribosomal binding site is transcribed from a DNA sequence that is selected from the group consisting of SEQ ID NO: 17-19, and the second ribosomal binding sequence is transcribed from a DNA sequence that is selected from the group consisting of SEQ ID NO: 20-23.

**[0162]** Embodiment 13. The method of any one of the embodiments 11-12, wherein the first ribosomal binding site has a sequence of SEQ ID NO: 17 or SEQ ID NO: 18.

**[0163]** Embodiment 14. The method of any one of the embodiments above, wherein the HC and/or the LC of the full-length antibody comprises at least one non-natural amino acid.

**[0164]** Embodiment 15. The method of embodiment 14, wherein the coding sequence of the HC and /or the coding sequence of the LC has been modified to have at least one non-natural

amino acid codon, wherein the non-natural amino acid codon that does not result in incorporation of any naturally occurring amino acids.

**[0165]** Embodiment 16. The method of embodiment 14, wherein the at least one non-natural amino acid is introduced by charging a tRNA that contains anticodons that are complementary to the at least one non-natural amino acid codon.

**[0166]** Embodiment 17. The method of any one of the embodiments 14-16, wherein the at least one non-natural amino acid codon is the amber codon TAG.

**[0167]** Embodiment 18. The method of any one of the embodiments 14-16, wherein the at least one non-natural amino acid is para-methylazido-L-phenylalanine (pAMF), AEK, or pAcF.

**[0168]** Embodiment 19. The method of any one of claims 15-18, wherein the codon that is immediately 3' to at least one of the non-natural amino acid is codon optimized.

**[0169]** Embodiment 20. The method of any one of the embodiments above, wherein the full-length antibody is the B10 antibody, an H01 antibody, a 7219 antibody, an anti-PD1 antibody, an anti-Tim3 antibody, an anti LAG3 antibody, or an anti-Her 2 antibody.

**[0170]** Embodiment 21. The method of embodiment 20, wherein the coding sequence for the HC of the B10 antibody contains a mutation relative to SEQ ID NO: 1, wherein said the mutation results in a codon for a natural amino acid being substituted with a non-natural amino acid codon,

wherein the natural amino acid is one or more amino acid selected from F404, Y180, and F241, and

wherein a non-natural amino acid is introduced to the HC by charging a tRNA that is complementary to the non-natural amino acid codon.

**[0171]** Embodiment 22. The method of any one of the embodiments above, wherein the coding sequence for the LC of the B10 antibody or the coding sequence for the LC of the trastuzumab contains at least one mutation relative to SEQ ID NO: 2, wherein the at least one mutation results in a codon for a natural amino acid being substituted with a non-natural amino acid codon, wherein the wherein the natural amino acid is K42 or E161.

[0172] Embodiment 23. The method of any one of the embodiments above, wherein the method further comprises covalently linking a warhead moiety to the non-natural amino acid on the HC or the LC of the full-length antibody via a linker.

[0173] Embodiment 24. The method of embodiment 23, wherein the warhead is SC236 or Aminoxy-PEG8-methane.

## EXAMPLES

[0174] The following examples are offered to illustrate, but not to limit, the claimed invention.

### EXAMPLE 1. PRODUCTION OF PLASMIDS FOR EXPRESSION OF IGGS IN SNUGGLE

[0175] IgG production requires the concurrent synthesis of heavy chain (HC) and light chain (LC) polypeptides. To produce IgG expression plasmid pJ411-HC-LC B10, the genes for the HC and LC were cloned into a single bicistronic operon with a T7 promoter and T7 terminator. The bicistronic operon comprises SEQ ID NO: 35. This plasmid has a high copy pUC origin of replication and contains a Kanamycin selectable marker. Plasmid sequences were verified by cloning. Both the HC and LC had independent ribosomal binding sites. To optimize the ratio of HC and LC, mutations were made in the ribosomal binding site of either gene. The parent plasmids PJ411, PJ411, PJ 401, PJ 434, which were used to produce various expression plasmids in the examples, were purchased from ATUM (Newark, CA). The same expression plasmid pJ411-HC-LC B10 was used in Examples 1-6.

[0176] For purposes of this disclosure and examples, “ $\alpha$ FolRa B10” and “B10 WT” are used interchangeably. In some cases, additional nomenclature is added to indicate the promoter used in the construct. For example, “B10 WT – T7p” represents a B10 WT construct that uses a T7 promoter.

### EXAMPLE 2. SHAKE FLASK EXAMPLE WT IGG EXPRESSION

[0177] The *E. coli* strain for expression of  $\alpha$ FolR IgG B10 was made by transforming Snuggle, as disclosed in PCT/US2019/060345, with plasmid pJ411-HC-LC B10. This strain was grown overnight at 37° C in terrific broth (TB) with 50 ug/ml kanamycin. In the morning, it was diluted 1:50 into fresh TB + Kan and grown at 37° C until an OD<sub>600</sub> of 1.5. At this point,

expression of the HC and LC genes were induced by the addition of arabinose to a final concentration of 0.2%. The temperature was adjusted to 25° C and protein expression continued for 16 hours. Cells were harvested by centrifugation at 6000xg for 10 minutes.

#### EXAMPLE 3. ANALYSIS OF SHAKE FLASK IN VIVO IGG EXPRESSION

[0178] Cells were harvested by centrifugation at 6000xg for 10 minutes as described in Example 2 and were then frozen. Thawed cells were lysed by sonication in modified S30 buffer (10 mM Tris·HCl pH 8.2, 2 mM Magnesium Acetate, 60 mM Potassium Acetate and 0.1 mg/l hen egg white lysozyme) at 10 mL/g wet cell weight (wcw) and debris was removed by centrifugation at 19,000xg for 30 min. IgGs were purified from the cell lysate using standard protein A affinity chromatography methods, and antibody concentration was determined by UV/Vis spectroscopy.

#### EXAMPLE 4. COMPARISON OF IN VIVO IGG EXPRESSION IN OXIDIZING CYTOPLASM STRAINS SHUFFLE AND SNUGGLE

[0179] Strains Snuggle and Shuffle (NEB, Cambridge MA) were transformed with plasmid pJ411-HC-LC B10. The difference in the genetic composition of Suggle and Shuffle is described in International Application No. PCT/US2019/060345. In brief, the Shuffle strain lacks the thioredoxin reductase activity (TRXB) and glutathione reductase activity (GOR). The Shuffle also overexpresses a DsbC without its signal sequence and a variant of the *ahpC* gene (*ahpC\**) encoding an enzyme that lacks peroxyreductase activity but has glutathione reductase activity. As compared to Shuffle, the Snuggle strain has been further engineered (based on Shuffle) so that it also lacks the thioredoxin 1 activity (TrxA). As a result, the Snuggle strain comprises null mutations in *trxA*, *trxB*, and *gor* and thus lacks the thioredoxin reductase activity (TrxB), thioredoxin 1 activity (TrxA), and glutathione reductase activity (GOR). The Snuggle strain also overexpresses a DsbC without its signal sequence and overexpresses a variant of the *ahpC* gene (*ahpC\**) that encoded an enzyme that lacks peroxireductase activity but has glutathione reductase activity.

[0180] Cell from both strains were grown overnight at 37° C in terrific broth (TB) with 50 ug/ml kanamycin. In the morning, both cultures were diluted 1:50 into fresh TB + Kan and grown at 37° C until both strains reached a final OD<sub>600</sub> of 1.5. IgG expression was induced in SBDG419 by the addition of a final concentration of 0.2% arabinose. IgG expression in Shuffle

cells was induced by the addition of 1 mM IPTG final concentration. The temperature was adjusted to 25° C and protein expression continued for 16 hours.

[0181] **Figure 1** shows a comparison of titer in cell lysate from Shuffle and Snuggle strains co-expressing HC and LC for the FoRa B10 wt IgG, indicating additional mutations present in Snuggle, as described above, are useful for high level IgG synthesis.

#### EXAMPLE 5 CELL BANKING STRAINS FOR IN VIVO PRODUCTION OF WT IGGS IN *E. COLI*

[0182] The post-transformation culture of each of the WT IgG strains (i.e., pJ411-HC-LC B10), described herein, were streaked on a LB plate containing 50 µg/mL of Kanamycin and incubated in a 30° C incubator for colony development. A single colony was then picked and inoculated in a 50 mL bioreactor tube filled with 10 mL of I17-SF Shake Flask Media (Tables 8 and 9 describe the components of I17-SF Shake Flask Media). The tube was then incubated for 15-24 hours at 30° C and 250 RPM in a shaker with an orbital diameter of 25 mm. Once the OD 595 nm of the culture reached between 2-4, the cells were sub-cultured at a seeding density of 10-25% (v/v) in I17-SF Shake Flask Media. The flask was incubated at 37° C and 250 RPM in a shaker with an orbital diameter of 25 mm. Once OD 595 nm of 2-3 was reached, the cells were mixed with 80% pre-sterilized glycerol at a concentration of 20% (v/v) and aliquoted into 1 mL cryovials, which were then flash frozen in liquid nitrogen and placed in a -80°C freezer for long-term storage.

#### EXAMPLE 6 EXPRESSION AND CELL LYSIS FOR WT IGGS PRODUCED IN *E. COLI* WITH HIGH DENSITY FERMENTATIONS

[0183] The fermentation process was initiated by taking a 1 mL vial of the cell bank and inoculating a shake flask containing I17-SF defined Shake Flask Media with 50 µg/ml kanamycin at about 3% (v/v) seeding density. Once an OD 595 nm of 2-4 was reached, the flask culture was used to inoculate a 500 mL bioreactor at a seeding density of 1.5% (v/v) in batched media, which consists of 30 µg/mL of Kanamycin, 0.1% (v/v) P2000 antifoam and 1.2% (v/v) 10x I17 Media in DI H<sub>2</sub>O (Tables 5, 6 and 7 describe the components of 10x I17 Media). The bioreactor temperature, dissolved oxygen, and pH setpoints were 37° C, 30%, and 7, respectively. Once the cells grew to an OD 595 nm between 2-5 in the batch phase, the fed batch phase began by feeding 10x I17 Media at an exponential rate of 0.20 h<sup>-1</sup>. After 10 hours of the

fed-batch phase, the temperature of the bioreactor was decreased to 25°C and the exponential feed rate was decreased to 0.02 h<sup>-1</sup>. An hour later, the induction phase began by adding L-Arabinose to a target concentration of 4 g/L based on the starting volume of the bioreactor. The induction phase took 24-48 hours before the harvest. At the end of the fermentation, the culture was collected and centrifuged at 18,592 xG and 2-8° C for 15 min in a floor centrifuge. The supernatant was discarded and the cell pellets were resuspended and washed with S30 Buffer (Table 10 describes the components of S30 Buffer) at a concentration of 16.67% (w/w) and centrifuged again with the same conditions used in the initial harvest step. After the wash, the supernatant was discarded and the cells were resuspended with S30-5 Buffer (Table 11 describes the components of S30-5 Buffer) at a concentration of 16.67% (w/w). The cell resuspension was then passed through an Avestin Homogenizer (EmulsiFlex-C5) at 17,000 Psi to disrupt the cells and generate the crude lysate. The crude lysate was further clarified by centrifuging at 18,000-20,000 xG and 2-8° C for 30 minutes in a floor centrifuge. The supernatant (clarified lysate) was collected and aliquoted, flash frozen in liquid nitrogen, and stored at -80°C.

#### EXAMPLE 7 PRODUCTION OF PLASMIDS FOR EXPRESSION OF PAMF CONTAINING IGGS IN SNUGGLE

[0184] Nonnatural amino acid (NNAA) containing IgG production requires the concurrent synthesis of Heavy Chain (HC) and Light Chain (LC) polypeptides with either the HC and/or LC containing an NNAA. To produce the plasmid, pJ411-HC F404TAG-LC B10, for expression of the antibody B10 HC F404 pAMF, the genes for the HC and LC were cloned into a single bicistronic operon with a T7 promoter and T7 terminator. The codon for HC residue F404 was mutated to TAG to specify the location of the NNAA. This plasmid has a high copy pUC origin of replication and contains the gene for Kanamycin resistance. Plasmid sequences were verified by cloning. Both the HC and LC had independent ribosomal binding sites. To optimize the ratio of HC and LC, mutations were made in the ribosomal binding site of either gene. pJ411-HC F404TAG-LC B10 is identical to pJ411-HC-LC B10, except for the mutation of the F404 codon to TAG.

[0185] Co-translational pAMF incorporation also requires expression of an amber suppressor tRNA orthogonal to existing aminoacyl tRNA synthetases (AAtRS) and an orthogonal AAtRS that specifically recognizes the amber suppressor tRNA and the pAMF NNAA. The genes for

the pAMF AAtRS and tRNA were cloned as a bicistronic operon into vector pJ434, downstream of the inducible T7 promoter and constitutive promoter Pc0. This vector has a p15A origin of replication and  $\beta$ -lactamase selection marker conferring resistance to carbenicillin. Both the origin and marker are compatible with pJ411. Plasmid sequences were verified by cloning.

#### EXAMPLE 8 SHAKE FLASK EXAMPLE PAMF IGG EXPRESSION

[0186] The *E. coli* strain for expression of the  $\alpha$ FolR IgG with pAMF at residue 404 was made by co-transforming Snuggle strain SBDG419 with plasmid pJ411-HC F404TAG-LC B10 and pJ434 Pc0 pAMF RS-tRNA. This strain was grown overnight at 37° C in terrific broth (TB) containing 50 ug/ml kanamycin and 100 ug/ml carbenicillin. In the morning, it was diluted 1:50 into fresh TB + Kan/Carb and grown at 37° C until an OD<sub>600</sub> of 1.5. At this point, T7 driven transcription of HC, LC, AAtRS, and tRNA genes were induced by the addition of a final concentration of 0.2% arabinose and 2 mM pAMF. The temperature was adjusted to 25° C and protein expression continued for 16 hours. Cells were harvested by centrifugation at 6000xg for 10 minutes.

#### EXAMPLE 9 CELL BANKING STRAINS FOR IN VIVO PRODUCTION OF PAMF-IGGS IN *E. COLI*

[0187] The post-transformation culture of the nnAA IgG strains as described in Example 8 was streaked on a LB plate containing 50  $\mu$ g/mL of Kanamycin and 100  $\mu$ g/mL of Carbenicillin and incubated in a 30° C incubator for colony development. A single colony was then picked and inoculated in a 50 mL bioreactor tube filled with 10 mL of I17-SF Shake Flask Media containing 100  $\mu$ g/mL of Carbenicillin (refer to Table 8 and 9 for the components of I17-SF Shake Flask Media). The tube was then incubated for 15-24 hours at 30° C and 250 RPM in a shaker with an orbital diameter of 25 mm. Once the OD 595 nm of the culture reached between 2-4, the cells were sub-cultured at a seeding density of 10-25% (v/v) in I17-SF Shake Flask Media. The flask was incubated at 37° C and 250 RPM in a shaker with an orbital diameter of 25 mm. Once OD 595 nm of 2-3 was reached, the cells were mixed with 80% pre-sterilized glycerol at a concentration of 20% (v/v) and aliquoted into 1 mL cryovials which were then flash frozen in liquid nitrogen and placed in a -80°C freezer for long-term storage.

#### EXAMPLE 10 EXPRESSION AND CELL LYSIS FOR PAMF-IGGS PRODUCED IN *E. COLI* WITH HIGH DENSITY FERMENTATIONS

[0188] The fermentation process began by taking a 1 mL vial of the cell bank, as described in Example 9, and inoculating a shake flask with I17-SF Shake Flask Media containing 100 µg/mL of Carbenicillin at about 3% (v/v) seeding density. Once an OD 595 nm of 2-4 was reached, the flask culture was used to inoculate a 500 mL bioreactor at a seeding density of 1.5% (v/v) in batched media, which consists of 30 µg/mL of Kanamycin, 100 µg/mL of Carbenicillin, 0.1% (v/v) P2000 antifoam, and 1.2% (v/v) 10x I17 Media in DI H<sub>2</sub>O (Tables 5, 6 and 7 describe the components of 10x I17 Media). The bioreactor temperature, dissolved oxygen, and pH setpoints were 37° C, 30%, and 7, respectively. Once the cells grew to an OD 595 nm between 2-5 in the batch phase, the fed batch phase began by feeding 10x I17 Media at an exponential rate of 0.20 h<sup>-1</sup>. After 10 hours of the fed-batch phase, the temperature of the bioreactor was decreased to 25° C, and the exponential feed rate was decreased to 0.02 h<sup>-1</sup>. An hour later, the induction phase began by adding pAMF to a target concentration of 2 mM based on the culture volume of the bioreactor before induction and L-Arabinose to a target concentration of 4 g/L based on the starting volume of the bioreactor. The induction phase took 24-48 hours before the harvest. At the end of the fermentation, the culture was collected and centrifuged at 18,592 xG and 2-8° C for 15 min in a floor centrifuge. The supernatant was discarded and the cell pellets were resuspended and washed with S30 Buffer (refer to Table 10 for more information on the components of S30 Buffer) at a concentration of 16.67% (w/w) and centrifuged again with the same conditions used in the initial harvest step. After the wash, the supernatant was discarded and the cells were resuspended with S30-5 Buffer (refer to Table 11 for information on the components of S30-5 Buffer) at a concentration of 16.67% (w/w). The cell resuspension was then passed through an Avestin Homogenizer (EmulsiFlex-C5) at 17,000 Psi to disrupt the cells and generate the crude lysate. The crude lysate was further clarified by centrifuging at 18,000-20,000 xG and 2-8° C for 30 minutes in a floor centrifuge. The supernatant (clarified lysate) was collected and aliquoted, flash frozen in liquid nitrogen and stored at -80° C.

#### EXAMPLE 11 OPTIMIZATION OF CODONS 3' TO TAG FOR EXPRESSION OF NNAA CONTAINING IGGS

[0189] To optimize the S181 codon for production of B10 Y180 pAMF containing IgGs, the HC Y180TAG gene was produced with Ser codons AGC or AGT at the Ser181 position. The

genes for the HC and LC were cloned into pJ411 with a single bicistronic operon with a T7 promoter and T7 terminator. This plasmid, pJ411-HC Y180TAG-LC B10, has a high copy pUC origin of replication and contains the gene for Kanamycin resistance. Plasmid sequences were verified by cloning. Both the HC and LC had independent ribosomal binding sites having a sequence of SEQ ID NO: 17 and SEQ ID NO: 20, respectively. pJ411-HC Y180TAG-LC B10 is identical to pJ411-HC-LC B10, except for the mutation of the Y180 codon to TAG.

[0190] Strain SBDG419 was co-transformed with the codon optimized B10 expression plasmids and pJ434 containing pAMF RS and tRNA, and pAMF containing IgGs were expressed as in Example 8.

[0191] **Figure 2** shows a comparison of titer in cell lysate for FolRa-B10 IgGs produced with 0, 2, or 4 pAMF non-natural amino acids and the indicated S181 codons. As shown in Figure 2, the HC Y180TAG gene was made with Sercodons AGC and AGT at the Ser181 position shows higher titer than the one having TCG at same position. This indicates that optimization of the 3' S181 codon can substantially improve titers.

#### EXAMPLE 12 REASSEMBLY OF IN VIVO PRODUCED IGGS

[0192] IgG reassembly was performed in cell lysate adjusted to pH 8.0 with 100 mM Tris·HCl by incubating for 16 hrs in a flowerplate (m2p-labs) at 25°C with shaking at 650 rpm. Re-assembled IgGs were purified from the cell lysate using standard affinity chromatography methods and assembly was determined by caliper bioanalyzer under non-reducing conditions.

[0193] **Figure 3** shows a comparison of assembly of FolRa-B10 IgGs produced with 0, 2, or 4 pAMF non-natural amino acids after overnight reassembly by flowerplate determined by caliper bioanalyzer under non-reducing conditions.

#### EXAMPLE 13 SCREENING OF RIBOSOMAL BINDING SITE VARIANTS FOR OPTIMAL EXPRESSION OF IGGS IN *E. COLI*

[0194] Plasmids containing ribosomal binding site (RBS) variants of both the the HC and LC for the SP7219 IgG were co-transformed in SBDG419 with pJ434 Pc0 pAMF RS-tRNA. Cells were grown overnight at 37° C in terrific broth (TB) with 50 ug/ml kanamycin and 100 ug/ml carbenicillin. In the morning, it was diluted 1:50 into fresh TB + Kan/Carb and grown at 37° C until an OD<sub>600</sub> of 1.5. At this point T7 driven transcription of HC, LC, AAtRNAs and tRNA

genes were induced by the addition of a final concentration of 0.2% arabinose and 2 mM pAMF was added to the cultures. The temperature was adjusted to 25°C and protein expression continued for 16 hours. Cells were harvested by centrifugation at 6000xg for 10 minutes. Cells were lysed by sonication in modified S30 buffer (10 mM Tris·HCl pH 8.2, 2 mM Magnesium Acetate and 60 mM Potassium Acetate) at 10 mL/g w/w, and debris was removed by centrifugation at 19,000xg for 30 min. IgGs were purified from the cell lysate using standard affinity chromatography methods and concentration was determined by UV/Vis spectroscopy.

**Table 2** – Ribosomal binding site (RBS) variants used for screening relative HC to LC expression of IgGs in *E. coli*.

| Name      | Sequence | SEQ ID NO: | Notes            |
|-----------|----------|------------|------------------|
| HC - RBS1 | AGGAGGT  | 17         | Consensus RBS    |
| HC - RBS2 | AAGAGAT  | 18         | First mutant     |
| HC - RBS3 | AAAAGAT  | 19         | Second mutant    |
| LC – RBS1 | AGGAGAT  | 20         | Canonical LC RBS |
| LC – RBS2 | AAGAGAT  | 21         | First mutant     |
| LC – RBS3 | AAAAGAT  | 22         | Double mutant    |
| LC – RBS4 | AAAATAT  | 23         | Triple mutant    |

**Table 3** – IgG titer for SP7219 comparing ribosomal binding site (RBS) variants screened.

| HC RBS | SEQ ID NO | LC RBS | SEQ ID NO | Lysate titer (g/L) | IgG Weight percentage |
|--------|-----------|--------|-----------|--------------------|-----------------------|
| RBS1   | 17        | RBS1   | 20        | 0.35               | 0.35%                 |
| RBS1   | 17        | RBS2   | 21        | 0.50               | 0.50%                 |
| RBS1   | 17        | RBS3   | 22        | 0.43               | 0.43%                 |
| RBS1   | 17        | RBS4   | 23        | 0.45               | 0.45%                 |
| RBS2   | 18        | RBS1   | 20        | 0.60               | 0.6%                  |
| RBS3   | 19        | RBS1   | 20        | 0.00               | 0%                    |

[0195] A ribosomal binding sequence for the translation of the heavy chain may be transcribed from a DNA sequence of SEQ ID NO: 28: AX<sub>1</sub>GAGX<sub>2</sub>T (wherein X<sub>1</sub> is A or G and X<sub>2</sub> is A or G), and a ribosomal binding sequence for the translation of the light chain may be transcribed from a DNA sequence of SEQ ID NO: 29: AX<sub>1</sub>X<sub>2</sub>AX<sub>3</sub>AT (wherein X<sub>1</sub> is G or A, X<sub>2</sub> is G or A, and X<sub>3</sub> is G or T)

#### EXAMPLE 14 - SHAKE FLASK EXPRESSION OF IGGs CONTAINING PACF AND AEK NON-NATURAL AMINO ACID IN *E. COLI*

**[0196]** Strain SBDG419 was co-transformed with plasmids encoding for the B10 F404TAG IgG encoding for the along with pJ434 Pc0 pAcF RS-tRNA or pJ434 Pc0 Mm PylK RS- Mb PylK tRNA. Azidoethoxycarbonyl lysine (AEK) is a substrate for the WT pyrrolysine synthetase. These strains were grown overnight at 37° C in terrific broth (TB) with 50 ug/ml kanamycin and 100 ug/ml carbenicillin. In the morning, they were diluted 1:50 into fresh TB + Kan/Carb and grown at 37° C until an OD<sub>600</sub> of 1.5. At this point, T7 driven transcription of HC, LC, AAtRNAs and tRNA genes were induced by the addition of a final concentration of 0.2% arabinose and 2 mM pAcF or AEK was added to the cultures as appropriate. The temperature was adjusted to 25° C, and protein expression continued for 16 hours. Cells were harvested by centrifugation at 6000xg for 10 minutes. Cells were lysed by sonication in modified S30 buffer (10 mM Tris·HCl pH 8.2, 2 mM Magnesium Acetate and 60 mM Potassium Acetate ) at 10 mL/g w/w, and debris was removed by centrifugation at 19,000xg for 30 min. IgGs were purified from the cell lysate using standard affinity chromatography methods, and concentration was determined by UV/Vis spectroscopy.

**[0197]** **Figure 4** shows the expression of B10-F404pAcF and B10-F404AEK, both having a HC ribosomal binding sequence of SEQ ID NO: 17 and an LC ribosomal binding sequence of SEQ ID NO: 20 in the presence and absence of the appropriate non-natural amino acid in the culture media. The low titers in absence of NNAA indicate full-length IgG production requires NNAA for amber suppression.

#### EXAMPLE 15 SUMMARY OF IGG TITERS IN INTACT *E. COLI*

Table 4 and Table 5 summarize the results of titers of IgG produced from *E. coli* in shake flasks or bioreactors under conditions described in the Examples above. The volume of the bioreactors indicated in Table 4 was 0.5 liters.

**Table 4** – Titer of NNAA containing IgGs expressed in *E. coli*

| IgG | DA<br>R | TA<br>G<br>sites | NNAA/syst<br>em | Scale | Lysat<br>e<br>titer<br>(g/l) | Cult.<br>Titer<br>(mg/l<br>) | Wet<br>Weight<br>percenta<br>ge | Comments |
|-----|---------|------------------|-----------------|-------|------------------------------|------------------------------|---------------------------------|----------|
|-----|---------|------------------|-----------------|-------|------------------------------|------------------------------|---------------------------------|----------|

|     |   |            |         |             |       |       |        |                                                         |
|-----|---|------------|---------|-------------|-------|-------|--------|---------------------------------------------------------|
| B10 | 2 | F404       | pAMF/Mj | shake flask | 0.49  | 98    | 0.49%  | $\alpha$ FolR IgG                                       |
| B10 | 2 | Y180       | pAMF/Mj | shake flask | 0.38  | 76    | 0.38%  | $\alpha$ FolR IgG, S181 WT(TCG)                         |
| B10 | 2 | Y180       | pAMF/Mj | shake flask | 0.01  | 2     | 0.01%  | $\alpha$ FolR IgG, S181 WT(TCG), expressed without NNAA |
| B10 | 2 | Y180       | pAMF/Mj | shake flask | 0.485 | 97    | 0.485% | $\alpha$ FolR IgG, S181 AGC                             |
| B10 | 2 | Y180       | pAMF/Mj | shake flask | 0.792 | 158.4 | 0.792% | $\alpha$ FolR IgG, S181 AGT                             |
| B10 | 2 | F241       | pAMF/Mj | shake flask | 0.275 | 55    | 0.275% | $\alpha$ FolR IgG                                       |
| B10 | 4 | F404, K42  | pAMF/Mj | shake flask | 0.615 | 123   | 0.615% | $\alpha$ FolR IgG                                       |
| B10 | 4 | F404, Y180 | pAMF/Mj | shake flask | 0.095 | 19    | 0.095% | $\alpha$ FolR IgG, S181 WT(TCG)                         |
| B10 | 4 | F404, Y180 | pAMF/Mj | shake flask | 0.375 | 75    | 0.375% | $\alpha$ FolR IgG, S181 AGC                             |
| B10 | 4 | F404, Y180 | pAMF/Mj | shake flask | 0.44  | 88    | 0.44%  | $\alpha$ FolR IgG, S181 AGT                             |
| B10 | 4 | F404, Y180 | pAMF/Mj | shake flask | 0.285 | 57    | 0.285% | $\alpha$ FolR IgG, S181 WT(TCG), attenuated RF1         |
| B10 | 4 | F404, Y180 | pAMF/Mj | shake flask | 0.215 | 43    | 0.215% | $\alpha$ FolR IgG, S181 AGC attenuated RF1              |
| B10 | 4 | F404, Y180 | pAMF/Mj | shake flask | 0.31  | 62    | 0.310% | $\alpha$ FolR IgG, S181 AGT attenuated RF1              |
| B10 | 4 | F404, Y180 | pAMF/Mj | Bioreactor  | 0.38  | 383   | 0.38%  | $\alpha$ FolR IgG, S181 WT(TCG)                         |
| B10 | 4 | F404, Y180 | pAMF/Mj | Bioreactor  | 0.49  | 485   | 0.49%  | $\alpha$ FolR IgG, S181 AGC                             |

|      |   |                 |           |                                   |       |       |        |                                 |
|------|---|-----------------|-----------|-----------------------------------|-------|-------|--------|---------------------------------|
|      |   | Y180            |           |                                   |       |       |        |                                 |
| B10  | 4 | F404, Y180      | pAMF/Mj   | Bioreactor                        | 0.79  | 796   | 0.79%  | $\alpha$ FolR IgG, S181 AGT     |
| B10  | 6 | F404, Y180, K42 | pAMF/Mj   | shake flask                       | 0.15  | 30    | 0.15%  | $\alpha$ FolR IgG, S181 WT(TCG) |
| B10  | 6 | F404, Y180, K42 | pAMF/Mj   | shake flask                       | 0.29  | 58    | 0.29%  | $\alpha$ FolR IgG, S181 AGC     |
| B10  | 6 | F404, Y180, K42 | pAMF/Mj   | shake flask                       | 0.35  | 70    | 0.35%  | $\alpha$ FolR IgG, S181 AGT     |
| H01  | 4 | F404, Y180      | pAMF/Mj   | Bioreactor (a volume of 5 liters) | 1.8   | 1771  | 1.8%   | STRO2 IgG                       |
| 7219 | 2 | F404            | pAMF/Mj   | Bioreactor (a volume of 5 liters) | 1.1   | 1118  | 1.1%   | STRO1 IgG                       |
| B10  | 2 | F404            | pAcPhe/Mj | shake flask                       | 0.539 | 107.8 | 0.539% | $\alpha$ FolR IgG               |
| B10  | 2 | F404            | AEK/pylK  | shake flask                       | 0.685 | 137   | 0.685% | $\alpha$ FolR IgG               |

[0198] Purified A280 titers for NNAA IgGs expressed in Snuggle cells co-transformed with plasmids encoding listed IgG and tRNA/AA<sub>t</sub>RS for specified NNAA. All NNAAAs were used at 2 mM except for the third entry in the table which was expressed with 0 mM pAMF. The extremely low titers for this sample demonstrate that NNAA is required for amber suppression and production of full-length IgGs.

**Table 5** – Titer of *E. coli* expressed IgGs produced in a bioreactor

| Product                              | Bioreactor Volume | Culture titer(mg IgG/L media) | Wet Weight percentage |
|--------------------------------------|-------------------|-------------------------------|-----------------------|
| $\alpha$ FolRa B10 IgG               | 0.5 L             | 1575                          | 1.6%                  |
| $\alpha$ FolRa B10 Y180/F404pAMF IgG | 0.5 L             | 796                           | 0.80%                 |
| $\alpha$ FolRa H01 Y180/F404pAMF IgG | 0.5 L             | 1771                          | 1.7%                  |
| aCD74 F404pAMF IgG                   | 0.5 L             | 866                           | 0.87%                 |
| aCD74 F404pAMF IgG                   | 5 L               | 1118                          | 1.1%                  |
| aPD1 IgG                             | 0.5 L             | 5922                          | 5.9%                  |
| aTim3 IgG                            | 0.5 L             | 2813                          | 2.8%                  |

[0199] Purified A280 titers for NNAA IgGs expressed in Snuggle cells co-transformed with plasmids encoding listed IgG and tRNA/AA<sub>t</sub>RS for specified NNAA. pAMF was used at 2 mM.

#### EXAMPLE 16 MEDIA AND COMPONENTS FOR HIGH DENSITY *E. COLI* IGG FERMENTATION

**Table 6. Components of 10x I17 Media Solution**

| Component                             | Concentration | Unit |
|---------------------------------------|---------------|------|
| Potassium Phosphate (Monobasic)       | 40.0          | g/L  |
| Ammonium sulfate                      | 15.0          | g/L  |
| Potassium Chloride                    | 10.0          | g/L  |
| Sodium Citrate                        | 10.0          | g/L  |
| L-Asparagine·H <sub>2</sub> O         | 18.5          | g/L  |
| L-Methionine                          | 6.4           | g/L  |
| L-Proline                             | 22.1          | g/L  |
| MgSO <sub>4</sub> ·7H <sub>2</sub> O  | 8.4           | g/L  |
| Glucose                               | 500.0         | g/L  |
| Trace Metals (TM)                     | 3.0           | mL/L |
| Vitamins 1 (V1)                       | 10.0          | mL/L |
| DI H <sub>2</sub> O or better quality | Qs to 1 L     | mL   |

**Table 7. Components of concentrated stock solution of vitamins for 10x I17 Media Solution**

| Component               | Concentration | Units |
|-------------------------|---------------|-------|
| Choline Chloride        | 28.64         | g/L   |
| Nicotinic Acid (Niacin) | 25.12         | g/L   |

|                                            |           |      |
|--------------------------------------------|-----------|------|
| PABA (Aminobenzoic Acid {p-" or "4-<br>"}) | 25.60     | g/L  |
| Pantothenic Acid (B5)                      | 9.40      | g/L  |
| Pyridoxine (B6) (HCl)                      | 1.47      | g/L  |
| Riboflavin (B2)                            | 3.89      | g/L  |
| Thiamine (B1) (HCl)                        | 17.67     | g/L  |
| Potassium Hydroxide (6 M) *unit is<br>mL/L | 115.00    | mL/L |
| DI H <sub>2</sub> O or better quality      | Qs to 1 L | N/A  |

**Table 8. Components of concentrated stock solution of trace metals for 10x I17 Media Solution**

| <b>Component</b>                      | <b>Concentration<br/>(g/L)</b> |
|---------------------------------------|--------------------------------|
| Sodium Citrate Dihydrate              | 100.00                         |
| Ferric Chloride Hexahydrate           | 66.63                          |
| Sodium Molybdate Dihydrate            | 11.67                          |
| Cobalt Chloride Hexahydrate           | 11.32                          |
| Zinc Sulfate Heptahydrate             | 11.32                          |
| Cupric sulfate Pentahydrate           | 11.32                          |
| Manganese Sulfate Monohydrate         | 6.47                           |
| Boric Acid                            | 4.04                           |
| DI H <sub>2</sub> O or better quality | Qs to 1 L                      |

**Table 9. Components of I17-SF Shake Flask Media Solution**

| <b>Component</b>               | <b>Concentration<br/>(mL/L)</b> |
|--------------------------------|---------------------------------|
| 10x Base Salts                 | 90                              |
| 10x I17 Liquid Stock<br>Medium | 10                              |

|                                       |           |
|---------------------------------------|-----------|
| Kanamycin sulfate (30 g/L stock)      | 1         |
| pH to 6.8+/- 0.1 with 6M KOH          | As needed |
| DI H <sub>2</sub> O or better quality | Qs to 1 L |

**Table 10. Components of 10x Base Salts for I17-SF Shake Flask Media Solution**

| Component                             | Concentration (g/L) |
|---------------------------------------|---------------------|
| Potassium Phosphate Monobasic         | 75.8                |
| Ammonium sulfate                      | 15.0                |
| Potassium Chloride                    | 10.0                |
| Sodium Citrate                        | 10.0                |
| Magnesium Sulfate Heptahydrate        | 8.4                 |
| DI H <sub>2</sub> O or better quality | Qs to 1 L           |

**Table 11. Components of S30 Buffer**

| Component                             | Concentration (mL/L) |
|---------------------------------------|----------------------|
| 1 M Tris, pH 8.2 Stock                | 10                   |
| 1.4 M MgAcO Stock                     | 10                   |
| 6 M KAcO Stock                        | 10                   |
| DI H <sub>2</sub> O or better quality | 970                  |

**Table 12. Components of S30-5 Buffer**

| Component                             | Concentration (mL/L) |
|---------------------------------------|----------------------|
| 1 M Tris, pH 8.2 Stock                | 10                   |
| 1.4 M MgAcO Stock                     | 1.4                  |
| 6 M KAcO Stock                        | 83                   |
| DI H <sub>2</sub> O or better quality | 1906                 |

**EXAMPLE 17 CONJUGATION OF NNAA CONTAINING MABS.**

[0200] Click chemistry bio-conjugation reactions were carried out for B10-F404TAG IgGs expressed in vivo containing the pAMF, AEK, and pAcF non-natural amino acids. The azide nnAA containing IgGs, B10 F404pAMF, and B10 F404AEK, were conjugated to a DBCO-

maytansine linker warhead (SC236) via strain promoted alkyne-azide cycloaddition (SPAAC) “click” chemistry as follows: 1 mg/mL IgG was reacted with a 10-fold molar excess of SC236 in PBS for 16h at 22C. The B10 F404pAcF IgG was conjugated to aminooxy-PEG8-methane via oxime ligation as follows: 13 mg/mL IgG was reacted with a 40-fold molar excess of aminooxy-PEG8-methane in 100 mM sodium acetate pH 4.5 for 72h at 30C. IgGs were digested with IdeS protease, reduced with DTT, and the extent of the conjugation reaction was determined by LCMS (Agilent Technologies 6520 Accurate-Mass Q-TOF LC/MS).

[0201] **Figure 5** shows the deconvoluted LCMS spectra for Fc-fragment of B10 F404pAMF conjugated to the DBCO-maytansine drug-linker SC236. The predominant species observed has a mass of 25257.97 Da compared with the theoretical mass of 25258.04 Da for the conjugate. No species are observed near 23974.14 Da, the theoretical mass for the unconjugated Fc fragment. This conjugate has a calculated conjugation efficiency of 100%.

**Table 13.** Conjugation efficiency of in vivo produced IgGs containing the nnAAs pAMF, AEK and pAcF determined by LCMS.

| Protein          | Conjugation partner   | Conjugation efficiency |
|------------------|-----------------------|------------------------|
| B10 F404pAMF IgG | SC236                 | 100%                   |
| B10 F404AEK IgG  | SC236                 | 100%                   |
| B10 F404pAcF IgG | Aminooxy-PEG8-methane | 100%                   |

#### EXAMPLE 18. PRODUCTION OF PLASMIDS FOR EXPRESSION OF NNAA CONTAINING LC IN SNUGGLE

[0202] Nonnatural amino acid (NNAA) containing LC production requires the synthesis of Light Chain with concurrent incorporation of NNAA. To produce the plasmid, pJ411-trastuzumab LC K42TAG E161TAG, for expression of the LC with a NNAA at the K42 and E161 position, the LC gene was cloned into an operon with a T7 promoter and T7 terminator. The codons for LC residues K42 and E161 were mutated to TAG to specify the location of the NNAA. This plasmid has a high copy pUC origin of replication and contains the gene for Kanamycin resistance. Plasmid sequences were verified by cloning.

[0203] Cotranslational NNAA incorporation also requires expression of an amber suppressor tRNA orthogonal to existing aminoacyl tRNA synthetases (AAtRS) and an orthogonal AAtRS

that specifically recognizes the amber suppressor tRNA and the pAMF NNAA. The genes for the NNAA AAtRS and tRNA were cloned as a bicistronic operon into vector pJ434, downstream of the inducible T7 promoter and constitutive promoter Pc0. This vector has a p15A origin of replication and  $\beta$ -lactamase selection marker conferring resistance to carbenicillin. Both the origin and marker are compatible with pJ411. Plasmid sequences were verified by cloning.

#### EXAMPLE 19 SHAKE FLASK EXAMPLE P-AC-PHE LC EXPRESSION

[0204] The *E. coli* strain for expression of the trastuzumab LC with pAcPhe at residue K42 and E161 were made by co-transforming Snuggle strain SBDG419 with plasmid pJ411-trastuzumab LC K42TAG E161TAG and pJ434 Pc0 pAcPhe RS-tRNA. This strain was grown overnight at 37° C in terrific broth (TB) with 50 ug/ml kanamycin and 100 ug/ml carbenicillin. In the morning, it was diluted 1:50 into fresh TB + Kan/Carb and grown at 37° C until an OD<sub>600</sub> of 1.5. At this point T7 driven transcription of HC, LC, AAtRS, and tRNA genes were induced by the addition of a final concentration of 0.2% arabinose and 2 mM pAcPhe to the cultures. The temperature was adjusted to 25° C and protein expression continued for 16 hours. Cells were harvested by centrifugation at 6000xg for 10 minutes.

#### EXAMPLE 20. ANALYSIS OF P-AC-PHE LC EXPRESSION

[0205] Cells were frozen at -80 ° C overnight, thawed and then resuspended in BPer bacterial protein extraction reagent with 0.01 mg/ml lysosome and 1  $\mu$ l benzonase nuclease/ ml BPer solution to lyse the *E. coli* cells and hydrolyze nucleic acids. After incubating at room temperature for 10 minutes, cells were fully lysed and viscosity from the nucleic acid polymers was reduced. At this point, the lysate was clarified with centrifugation at 24,000 x G for 10 minutes. The LC-containing supernatant was saved for analysis. After reduction with 10 mM DTT, the samples were analyzed with a PAGE gel.

#### EXAMPLE 21. PURIFICATION AND CONJUGATION OF P-AC-PHE LC

[0206] Trastuzumab LC K42 pAcPhe E161pAcPhe was purified from lysate after shake flask expression using standard proL chromatography. 143 mg of purified LC were recovered after purification. The LCMS (**Figure 6**) of the intact LC indicates that all of the protein recovered was full length and there is no evidence of misincorporation. This means there is good fidelity with this synthetase at both TAG sites, and any truncated proteins do not have affinity for proL.

Purified K42pAcPhe E161pAcPhe LC was conjugated. LCMS of the conjugated LC (**Figure 7**) demonstrates a drug loading of 1.96 per LC or 98% conjugation efficiency.

**Table 14** - Titters of trastuzumab LC expressed with NNAA

| D<br>AR | TAG<br>sites | NNAA/s<br>ystem | scale                      | titer (mg/l<br>culture media) | comments                                 |
|---------|--------------|-----------------|----------------------------|-------------------------------|------------------------------------------|
| 1       | K42          | pAMF/M<br>j     | 0.5 liter<br>bioreactor    | 4100                          | titer estimated from gel                 |
| 1       | K42          | pAcPhe/<br>Mj   | 0.5<br>literbioreact<br>or | 4100                          | titer estimated from gel                 |
| 1       | K42          | AEK/Pyr<br>K    | 0.5 liter<br>bioreactor    | 539                           | purified protein titer                   |
| 2       | K42,<br>E161 | pAcPhe/<br>Mj   | shake flask                | 166                           | S162AGT codon,<br>purified protein titer |
| 2       | K42,<br>E161 | pAcPhe/<br>Mj   | shake flask                | 194                           | S162AGC codon,<br>purified protein titer |
| 2       | K42,<br>E170 | pAcPhe/<br>Mj   | shake flask                | 179                           | S171AGT codon,<br>purified protein titer |
| 2       | K42,<br>E170 | pAcPhe/<br>Mj   | shake flask                | 180                           | S171AGC codon,<br>purified protein titer |

#### EXAMPLE 22. QUANTIFYING PROTEIN PRODUCTION

[0207] To measure the amount of HC and LC present in cell lysate, which were reduced with 10 mM DTT, which separated the HC from the LC by treatment with 10 mM DTT and HC and LC protein concentration were determined by Sally Sue automated western blot (Protein Simple, San Jose, CA) according to the manufacturer's directions for the 12-230 kDa separation module (SM-S0001, Protein Simple). Reduced cell lysate was diluted 1:20 in the 0.1X sample buffer. Anti-Human IgG (H+L) Biotin-SP (109-065-003, Jackson ImmunoResearch) was used as a primary antibody at a 1:150 dilution in Antibody Diluant 2. A standard was prepared using a serial dilution of purified H01 antibody in 0.1X sample buffer.

#### EXAMPLE 23. PRODUCTION OF IGG UTILIZING ALTERNATIVE PROMOTER

[0208] IgG production requires the concurrent synthesis of heavy chain (HC) and light chain (LC) polypeptides. To produce IgG expression plasmid pJ401-LC-HC B10, the genes for the HC and LC were cloned into a single bicistronic operon with a T5 promoter (T5p) (SEQ ID NO: 23), which recruits the E. coli RNA polymerase for transcription. This plasmid has a high copy pUC origin of replication and contains a Kanamycin selectable marker. Plasmid sequences were verified by cloning. Both the HC and LC had independent ribosomal binding sites.

[0209] The E. coli strain for expression of B10 with this plasmid was made by transforming Snuggle, as disclosed in PCT/US2019/060345, with plasmid pJ401-LC-HC B10. This strain was grown overnight at 37° C in terrific broth (TB) with 50 ug/ml kanamycin. In the morning, it was diluted 1:50 into fresh TB + Kan and grown at 37° C until an OD600 of 1.5. At this point, expression of the HC and LC genes were induced by the addition of IPTG to a final concentration of 1 mM. The temperature was adjusted to 25° C, and protein expression continued for 16 hours. Cells were harvested by centrifugation at 6000g for 10 minutes, lysed, and then IgG expression was analyzed as previously described. The results can be seen in **Table 15** indicating that the expression of B10 IgG with the T5 promoter (“B10 WT-T5p”) is in line with other high yielding IgGs in a shake flask; conventional expression with the T7 promoter (T7p) (“B10 WT-T7p”) is shown for comparison.

**Table 15.** Titers of IgG B10 transcribed with T5 promoter

| B10 expression construct | Lysate titer (g/l) | Culture titer (shake flask) mg/l |
|--------------------------|--------------------|----------------------------------|
| B10 WT – T7p             | 0.655              | 136                              |
| B10 WT – T5p             | 0.485              | 97                               |

Based on the shake flask titers, it is expected that greater than 200 mg/l or even greater than 1g/l of the B10 IgG will be achieved when the same strain of B10 WT – T5p is produced in high density fermentation process such as in bioreactors. This is because previous studies, for example, Example 15, table 4, show that switching from shake flasks to high-density fermentation usually can increase antibody production by at least 5 fold.

#### EXAMPLE 24. PRODUCTION OF IGG UTILIZING ALTERNATIVE OPERON ORGANIZATION

[0210] IgG production requires concurrent synthesis of heavy chain (HC) and light chain (LC) polypeptides. Plasmid pJ411 B10 MO was cloned to produce IgG utilizing separate operons for the HC and LC. In this vector the genes for the HC and LC were each cloned into a monocistronic complete operon with a T7 promoter, LC or HC gene and then T7 terminator. This plasmid has a high copy pUC origin of replication and contains a Kanamycin selectable marker. Plasmid sequences were verified by cloning. Both the HC and LC had independent ribosomal binding sites. To optimize the ratio of HC and LC, mutations were made in the ribosomal binding site of the HC. Three (3) different plasmids were produced, B10-MO-SDs,

B10-MO-SDm, B10-MO-SDw, with strong, medium and weak HC ribosomal binding sites comprising SEQ ID NO: 20, 21, and 22, respectively. The LC used the same strong ribosomal binding site (SEQ ID NO: 17) for all constructs because previous studies indicate that more LC expression than HC expression is needed.

**[0211]** The E. coli strain for expression of B10 with these plasmids was made by transforming Snuggle, as disclosed in PCT/US2019/060345, with plasmids B10-MO-SDs (SEQ ID NO: 36), B10-MO-SDm (SEQ ID NO: 37), or B10-MO-SDw (SEQ ID NO: 38). This strain was grown overnight at 37° C in terrific broth (TB) with 50 ug/ml kanamycin. In the morning, it was diluted 1:50 into fresh TB + Kan and grown at 37° C until an OD600 of 1.5. At this point, expression of the HC and LC genes were induced by the addition of arabinose to a final concentration of 0.2%. The temperature was adjusted to 25° C, and protein expression continued for 16 hours. Cells were harvested by centrifugation at 6000xg for 10 minutes, lysed, and then IgG expression was analyzed as previously described. The results can be seen in **Table 16** indicating that the high expression of B10 IgG with separate operons can be achieved with proper SD tuning and HC to LC ratio. The yield for the best construct from this series was in line with other high yielding IgGs in a shake flask; conventional expression with a bicistronic operon (“B10 WT”) as described in Example 1 is shown for comparison.

**Table 16.** Titers of IgG B10 transcribed with separate LC and HC operons

| B10 expression construct   | Lysate titer (g/l) | Culture titer (shake flask) mg/l | Weight percentage |
|----------------------------|--------------------|----------------------------------|-------------------|
| B10 WT (SEQ ID NO: 35)     | 0.655              | 131                              | 0.655%            |
| B10-MO-SDs (SEQ ID NO: 36) | 0.153              | 31                               | 0.153%            |
| B10-MO-SDm (SEQ ID NO: 37) | 0.123              | 25                               | 0.123%            |
| B10-MO-SDw (SEQ ID NO: 38) | 0.501              | 100                              | 0.501%            |

Previous studies, for example, Example 15, table 4, show that switching from shake flasks to high density fermentation usually can increase antibody production by at least 5 fold relative to the shake flask titers. Thus, the results in this example demonstrate that greater than 200 mg/l or even greater than 1g/l of the B10 IgG can be produced when the same stain of B10-MO-SDw is cultured in high-density fermentation setting such as in bioreactors.

#### Abbreviations in this application

OD 595 nm = Optical Density measure at wavelength of 595 nm

RPM = Revolutions per minute

(v/v%) = volume of solute/volume of solution %

(w/w%) = weight of solute/weight of solution %

Qs = Quantum satis (add as much as needed to achieve desired)

SD=Shine Dalgarno sequence or ribosomal binding site or RBS;

SDs or RBSs = strong ribosomal binding site,

SDm or RBSm= medium strength ribosomal binding site,

SDw or RBSw= weak strength ribosomal binding site

## ILLUSTRATIVE SEQUENCES

### Protein sequences of IgG LC and HC expressed in *E. coli*

\* denotes site of NNAA

SEQ ID NO: 1 (WT B10 HC)

MEVQLVESGGGLVQPGGSLRLSCAASGFNTTTKSIHWVRQAPGKGLEWVGEIYPRDGIT  
DYADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCARGGWHWRSGYSYYLDYW  
GQGTILVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSG  
VHTFPAVLQSSGLYSLSSVTVTPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTC  
PPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVH  
NAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQP  
REPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSGD  
SFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK

SEQ ID NO: 2 (WT B10 LC or H01 LC or trastuzumab LC)

MDIQMTQSPSSLSASVGDRVTITCRASQDVNTAVAWYQQKPGKAPKLLIYSASFLYSGV  
PSRFGSGRSGTDFTLTITSLQPEDFATYYCQHHYTPPTFGQGTKVEIKRTVAAPSVFIFPP  
SDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSST  
LTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC

SEQ ID NO: 3 (B10 F404TAG HC)

MEVQLVESGGGLVQPGGSLRLSCAASGFNTTTKSIHWVRQAPGKGLEWVGEIYPRDGIT  
DYADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCARGGWHWRSGYSYYLDYW  
GQGTILVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSG

VHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTC  
PPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH  
NAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQP  
REPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPVLDSGD  
S\*FLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK

SEQ ID NO: 4 (B10 Y180TAG HC)

MEVQLVESGGGLVQPGGSLRLSCAASGFNTTTKSIHWVRQAPGKGLEWVGEIYPRDGIT  
DYADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCARGGWHWRSGYSYYLDYW  
GQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSG  
VHTFPAVLQSSGL\*SLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTC  
PPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH  
NAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQP  
REPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPVLDSGD  
SFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK

SEQ ID NO: 5 (B10 K42TAG LC)

MDIQMTQSPSSLSASVGDRVTITCRASQDVNTAVAWYQQKPG\*APKLLIYSASFLYSGV  
PSRFGSGRSGTDFTLTISSLQPEDFATYYCQQHYTTPPTFGQGTKVEIKRTVAAPSVFIFPP  
SDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSSTYSLSST  
LTLISKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC

SEQ ID NO: 6 (B10 Y180TAG F404TAG HC)

MEVQLVESGGGLVQPGGSLRLSCAASGFNTTTKSIHWVRQAPGKGLEWVGEIYPRDGIT  
DYADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCARGGWHWRSGYSYYLDYW  
GQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSG  
VHTFPAVLQSSGL\*SLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTC  
PPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH  
NAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQP  
REPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPVLDSGD  
S\*FLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK

SEQ ID NO: 7 (B10 241TAG HC)

MEVQLVESGGGLVQPGGSLRLSCAASGFNTTTKSIHWVRQAPGKGLEWVGEIYPRDGIT  
DYADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCARGGWHWRSGYSYYLDYW  
GQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSG  
VHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTC  
PPCPAPELLGGPSV\*LFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH  
NAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQP

REPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDG  
SFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK

SEQ ID NO: 8 (H01 Y180TAG F404TAG HC)

MEVQLVESGGGLVQPGGSLRLSCAASGFNIRTQSIHWVRQAPGKGLEWIGDIFPIDGTD  
YADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCARGSWSWPSGMDYYLDYWG  
QGTLTVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGV  
HTFPAVLQSSGL\*SLSSVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCP  
PCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHN  
AKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPR  
EPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGS  
\*FLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK

SEQ ID NO: 9 (7219 LC)

MDIQMTQSPSSVSASVGDRTITCRASQGIGSWLAWYQQKPGKAPKLLIYAADRLQSG  
VPSRFGSGSGTDFTLTISLQPEDFATYYCQQYHTYPLTFGGGKVEIKRTVAAPSVFIF  
PPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDYSLSS  
TLTSLKADYEEKHKVYACEVTHQGLSSPVTKSFNRGEC

SEQ ID NO: 10 (7219 F404TAG HC)

MQVQLVESGGGVVQPGRSLRLSCAASGFNFSDYGMHWVRQAPGKGLEWVAVIWDYD  
SISYYADSVKGRFTISRDNSTLYLQMNSLRAEDTAVYYCARGGTVEHGAVYGTVDVW  
GQGTTTVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSG  
VHTFPAVLQSSGLYSLSSVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTC  
PPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVH  
NAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPR  
REPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDG  
S\*FLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK

SEQ ID NO: 11 ( $\alpha$ PD1 HC)

MEVQLVQSGAEVKKPGASVKVSCKASGYTFDSYGISWVRQAPGQGLEWMGWISAYNG  
NTNYAQKLQGRVTMTTDTSTNTAYMELRSLRSDDTAVYYCARDVDYGTGSGYWGQG  
TLTVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTF  
PAVLQSSGLYSLSSVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCPCP  
PELLGGPSVFLFPPKPKDTLMISRTPEVTCEVVDVSHEDPEVKFNWYVDGVEVHNAKT  
KPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQ  
VYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFL  
YSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK

SEQ ID NO: 12 ( $\alpha$ PD1 LC)

MSYELTQPPSVSVSPGQTARITCSGDALPKQYAYWYQQKPGQAPVMVIYKDTERPSGIP  
ERFSGSSSGTKVTLTISGVQAEDEADYYCQSADNSITYRVFGGGTKVTVLGQPKAAPSV  
TLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTTTPSKQSNKYAA  
SSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS

SEQ ID NO: 13 ( $\alpha$ Tim3 HC)

MEVQLVESGGGLVQPGGSLRLSCAASGFNIDRYIIHWVRQAPGKGLEWVAGITPVRGY  
TEYADSVKDRFTISADTSKNTAYLQMNSLRAEDTAVYYCARGYVYRMWDSYDYWGQ  
GTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSVHT  
FPAVLQSSGLYSLSSVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCPPC  
PAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNAK  
TKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQ  
VYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFL  
YSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK

SEQ ID NO:14 ( $\alpha$ Tim3 LC)

MDIQMTQSPSSLSASVGDRTTITCRASQDVNTAVAWYQQKPGKAPKLLIYSASFLYSGV  
PSRFSGSRSGDFTLTISSLQPEDFATYYCQQHYTTPPTFGQGTKVEIKRTVAAPSVFIFPP  
SDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDYSLST  
LTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC

SEQ ID NO: 15 ( $\alpha$ LAG3 HC)

MQVQLVESGGGVVQPGRSLRLSCAASGFTFSSYGMHWVRQAPGKGLEWVAVIWYDGS  
YKYYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAREEAPENWDYALDVWG  
QGTTVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGV  
HTFPAVLQSSGLYSLSSVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCP  
PCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHN  
AKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPR  
EPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGS  
FFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK

SEQ ID NO:16 ( $\alpha$ LAG3 LC)

MEIVLTQSPGTLSLSPGERATLSCRASQSVSSSYLAWYQQKPGQAPRLLIYGASSRATGIP  
DRFSGSGSGDFTLTISRLEPEDFAVYYCQYGRSPFSFGPGTKVDIKRTVAAPSVFIFPPS  
DEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDYSLSTL  
TLISKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC

SEQ ID NO: 17 (HC-RBS1; consensus RBS))

AGGAGGT

SEQ ID NO: 18 (HC-RBS2; First mutant)

AAGAGAT

SEQ ID NO: 19 (HC-RBS3; Second mutant)

AAAAGAT

SEQ ID NO: 20 (LC-RBS1; canonical LC RBS)

AGGAGAT

SEQ ID NO: 21 (LC-RBS2; First mutant)

AAGAGAT

SEQ ID NO: 22 (LC-RBS3; Double mutant)

AAAAGAT

SEQ ID NO: 23 (LC-RBS4; Triple mutant)

AAAATAT

SEQ ID NO: 24 (H01 HC)

MEVQLVESGGGLVQPGGSLRLSCAASGFNIRTQSIHWVRQAPGKGLEWIGDIFPIDGITD  
YADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCARGSWSPSGMDYYLDYWG  
QGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGV  
HTFPAVLQSSGLYSLSSVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCP  
PCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHN  
AKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPR  
EPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGS  
FFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK

SEQ ID NO: 25 (7219 HC)

MQVQLVESGGGVVQGRSLRLSCAASGFNFSYGMHWVRQAPGKGLEWVAVIWDYDG  
SISYYADSVKGRFTISRDNSTNTLYLQMNSLRAEDTAVYYCARGGTVEHGAVYGTDVW  
GQGTTVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSG  
VHTFPAVLQSSGLYSLSSVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTC  
PPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVH  
NAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQP  
REPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDG  
SFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK

SEQ ID NO:26 (T7 promoter)

TAATACGACTCACTATAGGG

SEQ ID NO:27 (Trastuzumab HC)

MEVQLVESGGGLVQPGGSLRLSCAASGFTNIKDTYIHWVRQAPGKGLEWVARIYPTNGY  
TRYADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCSRWGGDGFYAMDYWGQG  
TLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTF  
PAVLQSSGLYSLSSVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCPPCP  
APPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKT  
KPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQ  
VYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFL  
YSKLTVDKSRWQQGNVFCFSVMHEALHNHYTQKSLSLSPGK

SEQ ID NO: 28 (Ribosomal binding sequence, e.g., for heavy chain):

AX<sub>1</sub>GAGX<sub>2</sub>T (wherein X<sub>1</sub> is A or G and X<sub>2</sub> is A or G),

SEQ ID NO: 29 (Ribosomal binding sequence, e.g., for light chain):

AX<sub>1</sub>X<sub>2</sub>AX<sub>3</sub>AT (wherein X<sub>1</sub> is G or A, X<sub>2</sub> is G or A, and X<sub>3</sub> is G or T)

SEQ ID NO: 30 (T5 promoter sequence)

TAATTGTGAGCGGATAACAATTACGAGCTTCATGCACAGTGAAATCATGAAAAATTT  
ATTTGCTTTGTGAGCGGATAACAATTATAATA

SEQ ID NO: 31 (comprising SEQ ID NO: 17 (underlined))

AAGGAGGTAAAAA

SEQ ID NO: 32 (comprising SEQ ID NO: 20 (underlined))

AGGAGATATACAT

SEQ ID NO: 33 (comprising SEQ ID NO: 21 (underlined))

AAGAGATATACAT

SEQ ID NO: 34 (comprising SEQ ID NO: 22 (underlined))

AAAAGATATACAT

SEQ ID NO: 35 (Operon sequence for B10 WT; the single-underlined (SEQ ID NO: 31) encodes the 5'UTR of the LC, and the double-underlined (SEQ ID NO:32) encodes the 5'UTR of the HC )

AAGGAGGTAAAAAATGGACATTCAAATGACCCAGTCTCCGTCGTCACTGTCCGCATC  
CGTTGGCGACCGCGTTACCATCACGTGCCGTGCGTCGCAAGATGTGAACACCGCCGT

GGCGTGGTATCAGCAAAAACCGGGCAAAGCTCCGAAGCTGCTGATCTATTCAGCCT  
 CTTTCCTGTA CTCTCGGGTGTTCCTCCCGTTTCTCAGGCTCTCGCTCGGGTACGGATTT  
 CACCCTGACTATTTCTTCACTGCAACCGGAAGATTTTGCGACGTA CTACTGTCAGCA  
 GCATTACACGACTCCGCCGACCTTTGGTCAGGGTACCAAGGTCGAGATTAAGCGTAC  
 CGTGGCTGCACCATCCGTGTTTATCTTCCCTCCGTCTGATGAGCAGCTGAAATCCGGT  
 ACGGCGTCGGTCGTCTGCTTGCTGAATAACTTCTATCCGCGTGAAGCGAAGGTGCAA  
 TGGAAGGTTGACAATGCCCTGCAGTCAGGTA ACTCCCAAGAGTCTGTTACCGAACA  
 AGATTCGAAAGACTCAACCTACTCCCTGTCTTCGACGCTGACGTTGTCCAAAGCGGA  
 CTATGAGAAACACAAGGTTTACGCATGTGAAGTGACCCACCAGGGCCTGTCATCTCC  
 GGTCACCAAATCATTTAATCGCGGTGAGTGCTAAACCCCTTTAAGAAAGGAGATATAC  
ATATGGAAGTTCAATTAGTCGAGTCTGGTGGCGGCTTAGTTCAACCTGGTGGTTCAT  
 TGCCTCTGTCTTGCGCTGCATCGGGTTTTAACACCACCAGAAATCCATTCATTGGGT  
 GCGCCAGGCCCCAGGTAAAGGCCTGGAATGGGTGCGCGAGATCTACCCGCGTGACG  
 GTATTACGATTATGCGGATTCCGTGAAGGGTCGCTTACCATCTCCGCGGATACTA  
 GCAAGAATACCGCGTATCTGCAGATGA ACTCTCTGCGTGCAGAGGACACCGCGGTC  
 TACTACTGTGCCCGTGGCGGCTGGCACTGGCGCTCGGGTTATTTCATATCTTGGAC  
 TACTGGGGTCAGGGTACGTTGGTTACCGTGTCTATCCGCGTCAACCAAGGGTCCGTCG  
 GTTTTCCGCTGGCGCGCTCGTCAAAATCTACGTCCGGTGGTACCGCCGCTCTGGGT  
 GCCTGGTTAAAGACTACTTTCCGAGCCGGTCACGGTTTCGTGGA ACTCTGGTGCCC  
 TGACTTCTGGCGTCCACACGTTCCAGCCGTTTTTGACGTCATCCGGTCTGTA CTCTGTT  
 GTCCTCTGTGGTCA CGGTGCCGTACGTCTCTGGGCACCCAAACCTATATCTGCAA  
 TGTC AACCACAAACCGTCCAATACGAAAGTTGACAAAAAAGTCGAGCCGAAATCTT  
 GCGACAAGACCCACACGTGCCCTCCGTGCCCGGCACCGGAACTGCTGGGCGGTCCG  
 TCGGTGTTCTGTTCCCGCCGAAGCCGAAAGATACTCTGATGATCTCACGTACCCCG  
 GAAGTCACGTGTGTTGTTGTTGACGTGTACACGAAGATCCAGAGGTGAAATTCAAT  
 TGGTATGTGGACGGTGTGCAAGTGATAATGCCAAAACCAAACCGCGCGAGGAACA  
 GTACAACTCCACCTACCGCGTCTGTGCGGTGTTGACCGTCCTGCATCAAGACTGGCT  
 GAACGGTAAAGAGTACAAGTGCAAGGTTTCAAATAAGGCACTGCCTGCGCCGATTG  
 AAAAGACCATCTCTAAGGCAAAGGGCCAGCCGCGTGAGCCACAGGTGTATACCTG  
 CCGCCGTCGCGTGAAGAAATGACCAAGAACCAAGTTTCACTGACGTGTCTGGTCAA  
 GGGCTTTTATCCGTCCGATATTGCGGTGGAGTGGGAGTCTAATGGCCAGCCGGA AAA  
 CAATTACAAAACGACTCCGCCGGTGTGGATTCCGACGGTTCGTTTTTCTGTATTCC  
 AAGCTGACCGTTGACAAATCACGTTGGCAGCAAGGCAACGTTTTTTCTTGTTCGGTA  
 ATGCACGAAGCGCTGCACAATCATTACACCCAGAAATCACTGTCTGTCTCCGGGC  
 AAATAA

SEQ ID NO: 36 (Operon sequence for B10-MO-SDs (the single-underlined (SEQ ID NO: 31) encodes the 5'UTR of the LC and the double-underlined (SEQ ID NO: 32) encodes the 5'UTR of the HC)

AAGGAGGTAAAAAATGGACATTCAAATGACCCAGTCTCCGTCTGCTCACTGTCCGCATC  
 CGTTGGCGACCGCGTTACCATCACGTGCCGTGCGTCGCAAGATGTGAACACCGCCGT  
 GGCGTGGTATCAGCAAAAACCGGGCAAAGCTCCGAAGCTGCTGATCTATTCAGCCT  
 CTTTCCTGTA CTCTCGGGTGTTCCTCCCGTTTCTCAGGCTCTCGCTCGGGTACGGATTT  
 CACCCTGACTATTTCTTCACTGCAACCGGAAGATTTTGCGACGTA CTACTGTCAGCA  
 GCATTACACGACTCCGCCGACCTTTGGTCAGGGTACCAAGGTCGAGATTAAGCGTAC

CGTGGCTGCACCATCCGTGTTTATCTTCCCTCCGTCTGATGAGCAGCTGAAATCCGGT  
 ACGGCGTCGGTCGTCTGCTTGAATAACTTCTATCCGCGTGAAGCGAAGGTGCAA  
 TGGAAGGTTGACAATGCCCTGCAGTCAGGTAACCCCAAGAGTCTGTTACCGAACA  
 AGATTCGAAAGACTCAACCTACTCCCTGTCTTCGACGCTGACGTTGTCCAAAGCGGA  
 CTATGAGAAACACAAGGTTTACGCATGTGAAGTGACCCACCAGGGCCTGTCATCTCC  
 GGTCACCAAATCATTTAATCGCGGTGAGTGCTAACCCCTAGCATAACCCCTTGGGG  
 CCTCTAAACGGGTCTTGAGGGGTTTTTTGCCCTGAGACGCGTCAATCGAGTTTCGTA  
 CCTAAGGGCGACACCCCTAATTAGCCCGGATTTTTTTTGCAGAAATTAATACGACTC  
 ACTATAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAG  
GAGATATACATATGGAAGTTCAATTAGTCGAGTCTGGTGGCGGCTTAGTTCAACCTG  
GTGGTTCATTGCGTCTGTCTTGCGCTGCATCGGGTTTTAACACCACCACGAAATCCAT  
 TCATTGGGTGCGCCAGGCCCCAGGTAAAGGCCTGGAATGGGTGCGCGAGATCTACC  
 CGCGTGACGGTATTACGGATTATGCGGATTCGGTGAAGGGTCGCTTCACCATCTCCG  
 CGGATACTAGCAAGAATACCGCGTATCTGCAGATGAACTCTCTGCGTGCAGAGGAC  
 ACCGCGGTCTACTACTGTGCCCGTGGCGGCTGGCACTGGCGCTCGGGTTATTCTATC  
 TATCTGGACTACTGGGGTCAGGGTACGTTGGTTACCGTGTCTATCCGCGTCAACCAAG  
 GGTCCGTCGGTTTTTCCGCTGGCGCGCTGCTCAAAATCTACGTCCGGTGGTACCGCC  
 GCTCTGGGTGCTGTTAAAGACTACTTTCCGAGCCGGTCACGGTTTCGTGGGAAC  
 TCTGGTGCCCTGACTTCTGGCGTCCACACGTTCCAGCCGTTTTGCAGTCATCCGGTC  
 TGACTCGTTGTCTCTGTGGTCACGGTGCCGTCATCGTCTCTGGGCACCCAAACCTA  
 TATCTGCAATGTCAACCACAAACCGTCCAATACGAAAGTTGACAAAAAAGTCGAGC  
 CGAAATCTTGCGACAAGACCCACACGTGCCCTCCGTGCCCGGCACCGGAACGTGTG  
 GGCGGTCCGTCGGTGTTCCTGTTCCCGCCGAAGCCGAAAGATACTCTGATGATCTCA  
 CGTACCCCGGAAGTCACGTGTGTTGTTGTTGACGTGTCACACGAAGATCCAGAGGTG  
 AAATTCAATTGGTATGTGGACGGTGTGCAAGTGCATAATGCCAAAACCAACCGCG  
 CGAGGAACAGTACAACCTCCACCTACCGCGTCGTGTCGGTGTGACCGTCCTGCATCA  
 AGACTGGCTGAACGGTAAAGAGTACAAGTGAAGGTTTCAAATAAGGCACTGCCTG  
 CGCCGATTGAAAAGACCATCTCTAAGGCAAAGGGCCAGCCGCGTGAGCCACAGGTG  
 TATACCCTGCCGCCGTGCGGTGAAGAAATGACCAAGAACCAAGTTTCACTGACGTGT  
 CTGGTCAAGGGCTTTTATCCGTCCGATATTGCGGTGGAGTGGGAGTCTAATGGCCAG  
 CCGGAAAACAATTACAAAACGACTCCGCCGGTGTGGATTCCGACGGTTCGTTTTTC  
 CTGTATTCCAAGCTGACCGTTGACAAATCACGTTGGCAGCAAGGCAACGTTTTTCT  
 TGTTCCGTAAATGCACGAAGCGCTGCACAATCATTACCCAGAAATCACTGTCGTTG  
 TCTCCGGGCAAATAATGAGTCGACCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAG  
 TTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGG  
 GTCTTGAGGGGTTTTT

SEQ ID NO: 37 (Operon sequence for B10-MO-SDm; the single-underlined (SEQ ID NO: 31) encodes the 5'UTR of the LC, and the double-underlined (SEQ ID NO: 33) encodes the 5' UTR of the HC)

AAGGAGGTAAAAAATGGACATTCAAATGACCCAGTCTCCGTGCTCACTGTCCGCATC  
 CGTTGGCGACCGCGTTACCATCACGTGCCGTGCGTCGCAAGATGTGAACACCGCCGT  
 GGCGTGGTATCAGCAAAAACCGGGCAAAGCTCCGAAGCTGCTGATCTATTACGCCT  
 CTTTCCTGTACTCGGGTGTTCGTCCCGTTTCTCAGGCTCTCGCTCGGGTACGGATT  
 CACCCTGACTATTTCTTCACTGCAACCGGAAGATTTTGCAGCTACTACTGTCAGCA  
 GCATTACACGACTCCGCCGACCTTTGGTCAGGGTACCAAGGTCGAGATTAAGCGTAC

CGTGGCTGCACCATCCGTGTTTATCTTCCCTCCGTCTGATGAGCAGCTGAAATCCGGT  
 ACGGCGTCGGTCGTCTGCTTGAATAACTTCTATCCGCGTGAAGCGAAGGTGCAA  
 TGGAAGGTTGACAATGCCCTGCAGTCAGGTAACCTCCCAAGAGTCTGTTACCGAACA  
 AGATTCGAAAGACTCAACCTACTCCCTGTCTTCGACGCTGACGTTGTCCAAAGCGGA  
 CTATGAGAAACACAAGGTTTACGCATGTGAAGTGACCCACCAGGGCCTGTCATCTCC  
 GGTCACCAAATCATTTAATCGCGGTGAGTGCTAACCCCTAGCATAACCCCTTGGGG  
 CCTCTAAACGGGTCTTGAGGGGTTTTTTGCCCTGAGACGCGTCAATCGAGTTTCGTA  
 CCTAAGGGCGACACCCCTAATTAGCCCGGATTTTTTTTGCAGAAATTAATACGACTC  
 ACTATAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAA  
GAGATATACATATGGAAGTTCAATTAGTCGAGTCTGGTGGCGGCTTAGTTCAACCTG  
GTGGTTCATTGCGTCTGTCTTGCGCTGCATCGGGTTTTAACACCACCACGAAATCCAT  
 TCATTGGGTGCGCCAGGCCCCAGGTAAAGGCCTGGAATGGGTGCGCGAGATCTACC  
 CGCGTGACGGTATTACGGATTATGCGGATTCGGTGAAGGGTCGCTTCACCATCTCCG  
 CGGATACTAGCAAGAATACCGCGTATCTGCAGATGAACTCTCTGCGTGCAGAGGAC  
 ACCGCGGTCTACTACTGTGCCCGTGGCGGCTGGCACTGGCGCTCGGGTTATTCTATC  
 TATCTGGACTACTGGGGTCAGGGTACGTTGGTTACCGTGTCTATCCGCGTCAACCAAG  
 GGTCCGTCGGTTTTTCCGCTGGCGCGCTGCTCAAAATCTACGTCCGGTGGTACCGCC  
 GCTCTGGGTGCTGTTAAAGACTACTTTCCGAGCCGGTCACGGTTTCGTGGAAC  
 TCTGGTGCCCTGACTTCTGGCGTCCACACGTTCCAGCCGTTTTGCAGTCATCCGGTC  
 TGACTCGTTGTCTCTGTGGTCACGGTGCCGTCATCGTCTCTGGGCACCCAAACCTA  
 TATCTGCAATGTCAACCACAAACCGTCCAATACGAAAGTTGACAAAAAAGTCGAGC  
 CGAAATCTTGCGACAAGACCCACACGTGCCCTCCGTGCCCGGCACCGGAACGTGTG  
 GGCGGTCCGTCGGTGTTCCTGTTCCCGCCGAAGCCGAAAGATACTCTGATGATCTCA  
 CGTACCCCGGAAGTCACGTGTGTTGTTGTTGACGTGTCACACGAAGATCCAGAGGTG  
 AAATTCAATTGGTATGTGGACGGTGTGCAAGTGCATAATGCCAAAACCAACCGCG  
 CGAGGAACAGTACAACCTCCACCTACCGCGTCGTGTCGGTGTGACCGTCCTGCATCA  
 AGACTGGCTGAACGGTAAAGAGTACAAGTGCAAGGTTTCAAATAAGGCACTGCCTG  
 CGCCGATTGAAAAGACCATCTCTAAGGCAAAGGGCCAGCCGCGTGAGCCACAGGTG  
 TATACCCTGCCGCCGTGCGGTGAAGAAATGACCAAGAACCAAGTTTCACTGACGTGT  
 CTGGTCAAGGGCTTTTATCCGTCCGATATTGCGGTGGAGTGGGAGTCTAATGGCCAG  
 CCGGAAAACAATTACAAAACGACTCCGCCGGTGTGGATTCCGACGGTTCGTTTTTC  
 CTGTATTCCAAGCTGACCGTTGACAAATCACGTTGGCAGCAAGGCAACGTTTTTCT  
 TGTTCCGTAAATGCACGAAGCGCTGCACAATCATTACCCAGAAATCACTGTCGTTG  
 TCTCCGGGCAAATAATGAGTCGACCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAG  
 TTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGG  
 GTCTTGAGGGGTTTTT

SEQ ID NO: 38 (Operon sequence for B10-MO-SDw; the single-underlined (SEQ ID NO: 31) encodes the 5'UTR of the LC, and the double-underlined (SEQ ID NO: 34) encodes the 5'UTR of the HC)

AAGGAGGTAAAAAATGGACATTCAAATGACCCAGTCTCCGTGCTCACTGTCCGCATC  
 CGTTGGCGACCGCGTTACCATCACGTGCCGTGCGTCGCAAGATGTGAACACCGCCGT  
 GGCGTGGTATCAGCAAAAACCGGGCAAAGCTCCGAAGCTGCTGATCTATTACGCCT  
 CTTTCCTGTACTCGGGTGTTCGTCCCGTTTCTCAGGCTCTCGCTCGGGTACGGATT  
 CACCCTGACTATTTCTTCACTGCAACCGGAAGATTTTGCAGCTACTACTGTCAGCA  
 GCATTACACGACTCCGCCGACCTTTGGTCAGGGTACCAAGGTCGAGATTAAGCGTAC

CGTGGCTGCACCATCCGTGTTTATCTTCCCTCCGTCTGATGAGCAGCTGAAATCCGGT  
ACGGCGTCGGTCGTCTGCTTGCTGAATAACTTCTATCCGCGTGAAGCGAAGGTGCAA  
TGGAAGGTTGACAATGCCCTGCAGTCAGGTAACTCCCAAGAGTCTGTTACCGAACA  
AGATTCGAAAGACTCAACCTACTCCCTGTCTTCGACGCTGACGTTGTCCAAAGCGGA  
CTATGAGAAACACAAGGTTTACGCATGTGAAGTGACCCACCAGGGCCTGTCATCTCC  
GGTCACCAAATCATTTAATCGCGGTGAGTGCTAACCCCTAGCATAACCCCTTGGGG  
CCTCTAAACGGGTCTTGAGGGGTTTTTTGCCCCTGAGACGCGTCAATCGAGTTTCGTA  
CCTAAGGGCGACACCCCTAATTAGCCCGGATTTTTTTTGCGAAATTAATACGACTC  
ACTATAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAA  
AAGATATACATATGGAAGTTCAATTAGTCGAGTCTGGTGGCGGCTTAGTTCAACCTG  
GTGGTTCATTGCGTCTGTCTTGCGCTGCATCGGGTTTTTAACACCACCACGAAATCCAT  
TCATTGGGTGCGCCAGGCCCCAGGTAAAGGCCTGGAATGGGTGCGCGAGATCTACC  
CGCGTGACGGTATTACGGATTATGCGGATTCGGTGAAGGGTCGCTTCACCATCTCCG  
CGGATACTAGCAAGAATACCGCGTATCTGCAGATGAACTCTCTGCGTGCAGAGGAC  
ACCGCGGTCTACTACTGTGCCCCTGGCGGCTGGCACTGGCGCTCGGGTTATTCTATC  
TATCTGGACTACTGGGGTCAGGGTACGGTTGGTTACCGTGTCTCCGCTCAACCAAG  
GGTCCGCTCGGTTTTTCCGCTGGCGCCGCTCGTCAAAATCTACGTCCGGTGGTACCGC  
GCTCTGGGTGCTGTTAAAGACTACTTTCCGGAGCCGGTCACGGTTTCGTGGAAC  
TCTGGTGCCCTGACTTCTGGCGTCCACACGTTCCAGCCGTTTTGCAGTCATCCGGTC  
TGTA CTGTTGTCTCTGTGGTCACGGTGCCGTCATCGTCTCTGGGCACCCAAACCTA  
TATCTGCAATGTCAACCACAAACCGTCCAATACGAAAGTTGACAAAAAAGTCGAGC  
CGAAATCTTGCGACAAGACCCACACGTGCCCTCCGTGCCCGGCACCGGAACGTCTG  
GGCGGTCCGTCGGTGTTCTGTGTTCCCGCCGAAGCCGAAAGATACTCTGATGATCTCA  
CGTACCCCGGAAGTCACGTGTGTTGTTGTTGACGTGTCACACGAAGATCCAGAGGTG  
AAATTCAATTGGTATGTGGACGGTGTGCAAGTGCATAATGCCAAAACCAAACCGCG  
CGAGGAACAGTACAACCTCCACCTACCGCGTCGTGTCGGTGTTGACCGTCCTGCATCA  
AGACTGGCTGAACGGTAAAGAGTACAAGTGCAAGGTTTCAAATAAGGCACTGCCTG  
CGCCGATTGAAAAGACCATCTCTAAGGCAAAGGGCCAGCCGCGTGAGCCACAGGTG  
TATACCCTGCCGCCGTGCGGTGAAGAAATGACCAAGAACCAAGTTTCACTGACGTGT  
CTGGTCAAGGGCTTTTATCCGTCCGATATTGCGGTGGAGTGGGAGTCTAATGGCCAG  
CCGGAACAATTACAAAACGACTCCGCCGGTGTGGATTCCGACGGTTCGTTTTTCT  
CTGTATTCCAAGCTGACCGTTGACAAATCACGTTGGCAGCAAGGCAACGTTTTTCT  
TGTTCCGTAAATGCACGAAGCGCTGCACAATCATTACCCAGAAATCACTGTCGTTG  
TCTCCGGGCAAATAATGAGTCGACCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAG  
TTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGG  
GTCTTGAGGGGTTTTT