DUAL SELECTION EXPRESSION VECTOR SYSTEMS FOR MULTI-CHAIN BIOLOGICS

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of priority to U.S. Provisional Patent Application No. 63/559,778, filed February 29, 2024, which is hereby incorporated by reference in its entirety.

FIELD

[0002] The present disclosure provides dual selection expression vector systems combining a metabolic selectable marker and an antibiotic resistance selectable marker, compositions and mammalian host cells, including but not limited to Chinese hamster ovary (CHO) cells, comprising the same, and methods of using the foregoing for the expression of recombinant proteins, such as multi-chain biologics (e.g., recombinant proteins comprising three or four unique polypeptide chains).

SUBMISSION OF SEQUENCE LISTING

[0003] The content of the following Sequence Listing XML is incorporated herein by reference in its entirety: file name: 10589-WO01-SEC_ST26, date created: February 27, 2025; size: 10,676 bytes.

BACKGROUND

[0004] Due to their broad applications, biologics are used worldwide in a variety of applications, including therapeutic and diagnostic applications. Mammalian cell lines are the predominant expression systems for these biologics, with Chinese hamster ovary (CHO) cells being the predominant cellular factory. (See Lalonde *et al.*, 2017, J Biotechnol 251:128-140.) Particularly with the advent of biosimilars, speed-to-market and cost efficiency in the manufacture of biologics are now more important than ever before.

[0005] The costs associated with manufacturing biologics are high due to their complexity of production, utilizing multi-step processes involving the selection of optimized cell lines, culturing production cells in large quantities, and purifying the desired biologic from the cell harvest. Manufacturing is typically even more complex for new antibody modalities, such as those having three or four unique antibody chains. While these costs are decreasing due to improvements in all facets of production, they can still be prohibitive in the widespread adoption of new antibody modalities as front-line therapies.

[0006] In order to make biological therapeutics more accessible to patients, decreasing the cost of goods for the manufacturing process is an attractive proposition. One way to achieve this objective is to increase the titers associated with production cell lines. Expression vector configurations can help to optimize expression levels of different chains in recombinant proteins, particularly for three- or four- unique polypeptide chain molecules, resulting in more balanced chain expression, reduced impurities, and higher product quality.

[0007] Establishing stable host cell lines, such as, for example, CHO cell lines, which produce multi-chain recombinant antibody modalities entails the integration of the heavy and light chains (HC and LC, respectively) from two or more expression vectors (bicistronic or monocistronic) into the genome through a selection process. The most commonly used selection method for generating CHO cell lines utilizes a single selection approach that mostly relies on a metabolic selectable marker, such as dihydrofolate reductase (DHFR) or glutamine synthetase (GS), alongside inhibitory compounds like methotrexate (MTX) or methionine sulfoximine (MSX), regardless of the number of expression vectors. However, this selection method may not result in optimized expression levels and chain ratios for complex biologics such as multi-chain recombinant proteins (e.g., therapeutic molecules with three or four unique polypeptide chains).

[0008] Accordingly, there still exists a need in the art for expression vector systems which, when transfected into host cell lines, produce recombinant proteins (e.g., three- and four- unique polypeptide chain molecules) at high titers, with minimal impact on product quality attributes and/or improvements to product quality, such as reduced product-related impurities (e.g., high molecular weight aggregates) Such expression vector systems would benefit the process development of biologics.

SUMMARY

[0009] The present disclosure relates to a dual selection strategy for mammalian host cell lines for expressing recombinant proteins, including, for example, different antibody-derived modalities with three or four unique polypeptide chains. Illustratively, provided herein are dual selection expression vector systems in which one vector employs a metabolic selectable marker and the other employs an antibiotic resistance selectable marker, compositions and mammalian host cells comprising such expression vector systems, and methods of using the foregoing in the expression of recombinant proteins.

[0010] One aspect of the present disclosure provides an expression vector system comprising a first expression vector and a second expression vector, wherein:

the first expression vector comprises a first nucleotide sequence encoding a first antibody light chain, a second nucleotide sequence encoding a first antibody heavy chain, antibody heavy chain fusion, or Fc chain fusion, and a third nucleotide sequence encoding a metabolic selectable marker; and

the second expression vector comprises a fourth nucleotide sequence encoding a second antibody light chain, a fifth nucleotide sequence encoding a second antibody heavy chain, antibody heavy chain fusion, or Fc chain fusion, and a sixth nucleotide sequence encoding an antibiotic resistance selectable marker,

wherein, if the second nucleotide sequence or the fifth nucleotide sequence encodes an Fc chain fusion, then the other nucleotide sequence encodes an antibody heavy chain or antibody heavy chain fusion.

[0011] In some embodiments, the first expression vector comprises, in 5' to 3' order, the first nucleotide sequence, the second nucleotide sequence, and the third nucleotide sequence. In some embodiments, the second expression vector comprises, in 5' to 3' order, the fourth nucleotide sequence, the fifth nucleotide sequence, and the sixth nucleotide sequence. In some embodiments, the first expression vector comprises, in 5' to 3' order, the first nucleotide sequence, the second nucleotide sequence, and the third nucleotide sequence, and the second expression vector comprises, in 5' to 3' order, the fourth nucleotide sequence, the fifth nucleotide sequence, and the sixth nucleotide sequence.

[0012] In some embodiments, the metabolic selectable marker is glutamine synthetase or dihydrofolate reductase. In some embodiments, the metabolic selectable marker is glutamine synthetase. In some embodiments, the metabolic selectable marker is dihydrofolate reductase.

[0013] In some embodiments, the antibiotic resistance selectable marker is a resistance marker to an antibiotic selected from the group consisting of puromycin, geneticin, hygromycin, blasticidin, and phleomycin D. In some embodiments, the antibiotic resistance selectable marker is a resistance marker to hygromycin.

[0014] In some embodiments, the metabolic selectable marker is glutamine synthetase, and the antibiotic resistance selectable marker is a resistance marker to hygromycin.

[0015] In some embodiments, the first expression vector does not comprise a nucleotide sequence encoding an antibiotic resistance selectable marker. In some embodiments, the second expression vector does not comprise a nucleotide sequence encoding a metabolic selectable marker. In some embodiments, the first expression vector does not comprise a nucleotide sequence encoding an antibiotic resistance selectable marker and the second expression vector does not comprise a nucleotide sequence encoding a metabolic selectable marker.

[0016] In some embodiments, a first promoter is operably linked to the first nucleotide sequence, a second promoter is operably linked to the second nucleotide sequence, and a third promoter is operably linked to the third nucleotide sequence. In some embodiments, the first expression vector comprises, in 5' to 3' order, the first promoter, the first nucleotide sequence, the second promoter, the second nucleotide sequence, the third promoter, and the third nucleotide sequence. In some embodiments, each of the first promoter and the second promoter is a CMV-derived promoter (e.g., a CMV/GAPDH, CMV/adL, or CMV/EF1a promoter). In some embodiments, each of the first promoter and the second promoter is a GAPDH promoter (e.g., a CMV/GAPDH promoter). In some embodiments, the third promoter is mPGK or SRα.

[0017] In some embodiments, a fourth promoter is operably linked to the fourth nucleotide sequence, a fifth promoter is operably linked to the fifth nucleotide sequence, and a sixth promoter is operably linked to the sixth nucleotide sequence. In some embodiments, the second expression vector comprises, in 5' to 3' order, the fourth promoter, the fourth nucleotide sequence, the fifth promoter, the fifth nucleotide sequence, the sixth promoter, and the sixth nucleotide sequence. In some embodiments, each of the fourth promoter and the fifth promoter is a CMV-derived promoter (e.g., a CMV/GAPDH, CMV/adL, or CMV/EF1a promoter). In some embodiments, each of the fourth promoter and the fifth promoter is a GAPDH promoter (e.g., a CMV/GAPDH promoter). In some embodiments, the sixth promoter is SV40.

[0018] In some embodiments, a first polyA sequence is operably linked to the first nucleotide sequence, a second polyA sequence is operably linked to the second nucleotide sequence, a third polyA sequence is operably linked to the third nucleotide sequence, a fourth polyA sequence is operably linked to the fourth nucleotide sequence, a fifth polyA sequence is operably linked to the fifth nucleotide sequence, and a sixth polyA sequence is operably linked to the sixth nucleotide sequence. In some embodiments, the first expression vector comprises, in 5' to 3' order, the first promoter, the first nucleotide sequence, the first polyA sequence, the second promoter, the second nucleotide sequence, the second polyA sequence, the third promoter, the third nucleotide sequence, and the third polyA sequence. In some embodiments, the second expression vector comprises, in 5' to 3' order, the fourth promoter, the fourth nucleotide sequence, the fourth polyA sequence, the fifth promoter, the fifth nucleotide sequence, the fifth polyA sequence, the sixth promoter, the sixth nucleotide sequence, and the sixth polyA sequence. In some embodiments, each of the first polyA sequence, the second polyA sequence, the third polyA sequence, the fourth polyA sequence, the fifth polyA sequence, and the sixth polyA sequence is independently selected from the group consisting of a rabbit beta-globin pA sequence, a thymidine kinase pA (TKpA) sequence, and a simian virus 40 (SV40) early pA sequence. In some embodiments, each polyA sequence is a SV40 early pA sequence.

[0019] In some embodiments, the expression vector system encodes a Fab-heteroFc-[scFv*], heteroIgG, asymmetric antibody-cytokine fusion, [VH-VH*Fab]-heteroFc, [VH*Fab]-heteroFc, or [Fab*]-heteroFc-[VH*VH] molecule.

[0020] Another aspect of the present disclosure provides an expression vector system comprising a first expression vector and a second expression vector, wherein:

the first expression vector comprises, in 5' to 3' order, a first promoter, a first nucleotide sequence encoding a first antibody light chain, a first polyA sequence, a second promoter, a second nucleotide sequence encoding a first antibody heavy chain, antibody heavy chain fusion, or Fc chain fusion, and a third nucleotide sequence encoding a metabolic selectable marker; and

the second expression vector comprises a fourth nucleotide sequence encoding a second antibody light chain, a fifth nucleotide sequence encoding a second antibody heavy chain, antibody

heavy chain fusion, or Fc chain fusion, and a sixth nucleotide sequence encoding an antibiotic resistance selectable marker,

wherein, if the second nucleotide sequence or the fifth nucleotide sequence encodes an Fc chain fusion, then the other nucleotide sequence encodes an antibody heavy chain or antibody heavy chain fusion.

[0021] In some embodiments, the third promoter is mPGK or SRα. In some embodiments, the sixth promoter is SV40. In some embodiments, each of the first promoter, the second promoter, the fourth promoter, and the fifth promoter is a CMV-derived promoter (e.g., a CMV/GAPDH, CMV/adL, or CMV/EF1a promoter). In some embodiments, each of the first promoter, the second promoter, the fourth promoter, and the fifth promoter is a CMV/GAPDH or CMV/adL promoter. In some embodiments, each of the first promoter, the second promoter, the fourth promoter, and the fifth promoter is a GAPDH promoter (e.g., a CMV/GAPDH promoter). In some embodiments, each of the first polyA sequence, the second polyA sequence, the third polyA sequence, the fourth polyA sequence, the fifth polyA sequence, and the sixth polyA sequence is a simian virus 40 (SV40) early pA sequence.

[0022] In some embodiments, the third promoter is mPGK or SRα; the sixth promoter is SV40; each of the first promoter, the second promoter, the fourth promoter, and the fifth promoter is a CMV/GAPDH or CMV/adL promoter; and each of the first polyA sequence, the second polyA sequence, the third polyA sequence, the fourth polyA sequence, the fifth polyA sequence, and the sixth polyA sequence is a simian virus 40 (SV40) early pA sequence.

[0023] In some embodiments, the metabolic selectable marker is glutamine synthetase, and the antibiotic resistance selectable marker is a resistance marker to hygromycin.

[0024] In some embodiments, the third promoter is mPGK or SRα; the sixth promoter is SV40; each of the first promoter, the second promoter, the fourth promoter, and the fifth promoter is a CMV/GAPDH or CMV/adL promoter; each of the first polyA sequence, the second polyA sequence, the third polyA sequence, the fourth polyA sequence, the fifth polyA sequence, and the sixth polyA sequence is a simian virus 40 (SV40) early pA sequence; the metabolic selectable marker is glutamine synthetase; and the antibiotic resistance selectable marker is a resistance marker to hygromycin.

[0025] In some embodiments, the expression vector system encodes a Fab-heteroFc-[scFv*], heteroIgG, asymmetric antibody-cytokine fusion, [VH-VH*Fab]-heteroFc, [VH*Fab]-heteroFc, or [Fab*]-heteroFc-[VH*VH] molecule.

[0026] Still another aspect of the present disclosure relates to a composition comprising an expression vector system described herein. Such compositions may be useful in methods of producing recombinant proteins, including, for example, Fab-heteroFc-[scFv*], heteroIgG, asymmetric antibody-cytokine fusion, [VH-VH*Fab]-heteroFc, [VH*Fab]-heteroFc, or [Fab*]-heteroFc-[VH*VH] molecules, as described herein.

[0027] Another aspect of the present disclosure relates to a mammalian host cell comprising an expression vector system described herein. For example, the present disclosure provides a mammalian host cell comprising a pair of expression vectors, wherein a) the first expression vector comprises nucleotide sequences encoding a first antibody light chain, a first antibody heavy chain or antibody heavy chain fusion, and a metabolic selectable marker; and b) the second expression vector comprises nucleotide sequences encoding a second antibody light chain, a second antibody heavy chain or antibody heavy chain fusion, and an antibiotic resistance selectable marker, wherein the antibiotic resistance selectable marker is expressed in said mammalian cell.

[0028] In certain embodiments, the metabolic selectable marker is selected from the group consisting of glutamine synthetase and dihydrofolate reductase. In some embodiments, the metabolic selectable marker is glutamine synthetase. In some embodiments, the metabolic selectable marker is dihydrofolate reductase.

[0029] In certain embodiments, the antibiotic resistance selectable marker is a resistance marker to an antibiotic selected from the group consisting of puromycin, geneticin, hygromycin, blasticidin, and phleomycin D. In certain embodiments, the antibiotic resistance selectable marker is a resistance marker to puromycin. In certain embodiments, the antibiotic resistance selectable marker is a resistance marker to geneticin. In certain embodiments, the antibiotic resistance selectable marker is a resistance marker to hygromycin. In certain embodiments, the antibiotic resistance selectable marker is a resistance marker to blasticidin. In certain embodiments, the antibiotic resistance selectable marker is a resistance marker to phleomycin D.

[0030] In some embodiments, the metabolic selectable marker is glutamine synthetase, and the antibiotic resistance marker is a resistance marker to hygromycin.

[0031] In some embodiments, the metabolic selectable marker is dihydrofolate reductase, and the antibiotic resistance marker is a resistance marker to puromycin.

[0032] In certain embodiments, the first antibody light chain and the second antibody light chain have the identical sequence.

[0033] In certain embodiments, the first expression vector encodes a first antibody heavy chain, the second expression vector encodes a second antibody heavy chain, and the first antibody heavy chain is different the second antibody heavy chain.

[0034] In certain embodiments, the first expression vector encodes an antibody heavy chain and the second expression vector encodes an antibody heavy chain fusion. In certain embodiments, the first expression vector encodes an antibody heavy chain fusion and the second expression vector encodes an antibody heavy chain.

[0035] In any of the embodiments described above, the antibody heavy chain fusion is selected from the group consisting of an antibody heavy chain-scFv, an antibody heavy chain-cytokine, and an antibody heavy chain-VHH.

[0036] In certain embodiments, a promoter is operably linked to each of the nucleotide sequences encoding an antibody light chain, an antibody heavy chain or antibody heavy chain fusion, and a metabolic selectable marker. In certain aspects of this embodiment, the promoter for the metabolic selectable marker is selected from the group consisting of mPGK and SRa and the promoter for the antibiotic resistance selectable marker is SV40.

[0037] In some embodiments, the metabolic selectable marker is glutamine synthetase, and the promoter for the metabolic selectable marker is mPGK. In some embodiments, the metabolic selectable marker is glutamine synthetase, and the promoter for the metabolic selectable marker is SRa.

[0038] In some embodiments, the promoter for the antibiotic resistance selectable marker is SV40, and the antibiotic resistance selectable marker is a resistance marker to hygromycin.

[0039] In some embodiments, the metabolic selectable marker is glutamine synthetase, the promoter for the metabolic selectable marker is mPGK or SRa, the promoter for the antibiotic resistance selectable marker is SV40, and the antibiotic resistance selectable marker is a resistance marker to hygromycin.

[0040] In certain embodiments, a polyA sequence is operably linked to each of the nucleotide sequences within the expression vectors. In certain aspects of this embodiment, the polyA sequences are the same or different and are selected from the group consisting of a rabbit beta-globin pA sequence, a thymidine kinase pA (TKpA) sequence, and a simian virus 40 (SV40) early pA sequence. In some embodiments, each polyA sequence is a simian virus 40 (SV40) early pA sequence.

[0041] In some embodiments, the mammalian host cell is a Chinese hamster ovary (CHO) cell.

[0042] In some embodiments, the CHO cell is dihydrofolate reductase deficient (dhfr-). In some embodiments, the CHO cell is dihydrofolate reductase deficient (dhfr-), and the metabolic selectable marker is dihydrofolate reductase.

[0043] In some embodiments, the CHO cell is a glutamine synthetase knock out (GSKO). In some embodiments, the CHO cell is a glutamine synthetase knock out (GSKO), and the metabolic selectable marker is glutamine synthetase.

[0044] In one particular aspect, the present disclosure relates to a mammalian host cell, wherein a) the first expression vector comprises a nucleotide sequence which comprises the following elements in 5' to 3' order: 1) a first promoter operably linked to a nucleotide sequence encoding an antibody light chain followed by a first polyA sequence; 2) a second promoter operably linked to a nucleotide sequence encoding an antibody heavy chain followed by a second polyA sequence; and 3) a promoter operably linked to a nucleotide sequence encoding a metabolic selectable marker followed by a third polyA sequence; and b) the second expression vector comprises a nucleotide sequence which comprises the following elements in 5' to 3' order: 1) a first promoter operably linked to a nucleotide sequence encoding an antibody light chain followed by a first polyA sequence; 2) a second promoter

operably linked to a nucleotide sequence encoding an antibody heavy chain-scFv fusion, an antibody heavy chain-VHH fusion, or heavy chain-cytokine fusion followed by a second polyA sequence; and 3) a promoter operably linked to a nucleotide sequence encoding an antibiotic resistance selectable marker followed by a third polyA sequence.

[0045] In some embodiments, the first promoter and the second promoter of the first expression vector are CMV-derived promoters (e.g., CMV/GAPDH, CMV/adL, or CMV/EF1a promoters). In some embodiments, the first promoter and the second promoter of the first expression vector are CMV/GAPDH promoters. In some embodiments, the first promoter and the second promoter of the first expression vector are CMV/adL promoters.

[0046] In some embodiments, the first promoter and the second promoter of the second expression vector are CMV-derived promoters (e.g., CMV/GAPDH, CMV/adL, or CMV/EF1a promoters). In some embodiments, the first promoter and the second promoter of the second expression vector are CMV/GAPDH promoters. In some embodiments, the first promoter and the second promoter of the second expression vector are CMV/adL promoters.

[0047] In some embodiments, the first promoter and the second promoter of the first expression vector and the first promoter and the second promoter of the second expression vector are CMV-derived promoters (e.g., CMV/GAPDH, CMV/adL, or CMV/EF1a promoters). In some embodiments, the first promoter and the second promoter of the first expression vector and the first promoter and the second promoter of the second expression vector are CMV/GAPDH promoters. In some embodiments, the first promoter and the second promoter of the first expression vector and the first promoter and the second promoter of the second expression vector are CMV/adL promoters.

[0048] In some embodiments, the mammalian host cell expresses a C1 mAb or an antibody-cytokine fusion. In one sub-aspect, the mammalian host cell encodes a C1 mAb. In one sub-aspect, the mammalian host cell encodes an antibody cytokine-fusion.

[0049] In some embodiments, the antibody light chain encoded by the first expression vector and the antibody light chain encoded by the second expression vector comprise the same amino acid sequence.

[0050] In some embodiments, the metabolic selectable marker is glutamine synthetase, the promoter operably linked to a nucleotide sequence encoding a metabolic selectable marker is mPGK or SRa, the promoter operably linked to a nucleotide sequence encoding an antibiotic resistance selectable marker is SV40, and the antibiotic resistance selectable marker is a resistance marker to hygromycin.

[0051] In some embodiments, each polyA sequence is independently selected from the group consisting of a rabbit beta-globin pA sequence, a thymidine kinase pA (TKpA) sequence, and a simian virus 40 (SV40) early pA sequence. In some embodiments, each polyA sequence is a simian virus 40 (SV40) early pA sequence.

[0052] In some embodiments, the mammalian host cell is a Chinese Hamster Ovary (CHO) cell.

[0053] In some embodiments, the CHO cell is dihydrofolate reductase deficient (dhfr-). In some embodiments, the CHO cell is dihydrofolate reductase deficient (dhfr-), and the metabolic selectable marker is dihydrofolate reductase.

[0054] In some embodiments, the CHO cell is a glutamine synthetase knock out (GSKO). In some embodiments, the CHO cell is a glutamine synthetase knock out (GSKO), and the metabolic selectable marker is glutamine synthetase.

[0055] In another particular aspect, the present disclosure relates to a mammalian host cell, wherein a) the first expression vector comprises a nucleotide sequence which comprises the following elements in 5' to 3' order: 1) a first promoter operably linked to a nucleotide sequence encoding an antibody light chain followed by a first polyA sequence; 2) a second promoter operably linked to a nucleotide sequence encoding a first antibody heavy chain followed by a second polyA sequence; and 3) a promoter operably linked to a nucleotide sequence encoding a metabolic selectable marker followed by a third polyA sequence; and b) the second expression vector comprises a nucleotide sequence which comprises the following elements in 5' to 3' order: 1) a first promoter operably linked to a nucleotide sequence encoding an antibody light chain followed by a first polyA sequence; 2) a second promoter operably linked to a nucleotide sequence encoding a second antibody heavy chain followed by a second polyA sequence; and 3) a promoter operably linked to a nucleotide sequence encoding an antibiotic resistance selectable marker followed by a third polyA sequence, wherein the first antibody heavy chain and the second antibody heavy chain are different.

[0056] In one sub-aspect, the mammalian host cell encodes a heteroIgG.

[0057] In some embodiments, the antibody light chain encoded by the first expression vector and the antibody light chain encoded by the second expression vector comprise the same amino acid sequence. In some embodiments, the antibody light chain encoded by the first expression vector and the antibody light chain encoded by the second expression vector comprise different amino acid sequences.

[0058] In some embodiments, the first promoter and the second promoter of the first expression vector are CMV-derived promoters (e.g., CMV/GAPDH, CMV/adL, or CMV/EF1a promoters). In some embodiments, the first promoter and the second promoter of the first expression vector are CMV/GAPDH promoters. In some embodiments, the first promoter and the second promoter of the first expression vector are CMV/adL promoters.

[0059] In some embodiments, the first promoter and the second promoter of the second expression vector are CMV-derived promoters (e.g., CMV/GAPDH, CMV/adL, or CMV/EF1a promoters). In some embodiments, the first promoter and the second promoter of the second expression vector are CMV/GAPDH promoters. In some embodiments, the first promoter and the second promoter of the second expression vector are CMV/adL promoters.

[0060] In some embodiments, the first promoter and the second promoter of the first expression vector and the first promoter and the second promoter of the second expression vector are CMV-derived promoters (e.g., CMV/GAPDH, CMV/adL, or CMV/EF1a promoters). In some embodiments, the first promoter and the second promoter of the first expression vector and the first promoter and the second promoter of the second expression vector are CMV/GAPDH promoters. In some embodiments, the first promoter and the second promoter of the first expression vector and the first promoter and the second promoter of the second expression vector are CMV/adL promoters.

[0061] In some embodiments, the metabolic selectable marker is glutamine synthetase, the promoter operably linked to a nucleotide sequence encoding a metabolic selectable marker is mPGK or SRa, the promoter operably linked to a nucleotide sequence encoding an antibiotic resistance selectable marker is SV40, and the antibiotic resistance selectable marker is a resistance marker to hygromycin.

[0062] In some embodiments, each polyA sequence is independently selected from the group consisting of a rabbit beta-globin pA sequence, a thymidine kinase pA (TKpA) sequence, and a simian virus 40 (SV40) early pA sequence. In some embodiments, each polyA sequence is a simian virus 40 (SV40) early pA sequence.

[0063] In certain embodiments, the mammalian host cell is a Chinese Hamster Ovary (CHO) cell. In certain aspects of this embodiment, the CHO cell is dihydrofolate reductase deficient (dhfr-) or a glutamine synthetase knock out (GSKO).

[0064] In some embodiments, the CHO cell is dihydrofolate reductase deficient (dhfr-). In some embodiments, the CHO cell is dihydrofolate reductase deficient (dhfr-), and the metabolic selectable marker is dihydrofolate reductase.

[0065] In some embodiments, the CHO cell is a glutamine synthetase knock out (GSKO). In some embodiments, the CHO cell is a glutamine synthetase knock out (GSKO), and the metabolic selectable marker is glutamine synthetase.

[0066] Another aspect of the present disclosure relates to a method of producing a recombinant protein comprising culturing a mammalian host cell described herein in a cell culture media adapted for metabolic and antibiotic selection and recovering the recombinant protein. Illustratively, cell culture media adaptation for metabolic and antibiotic selection may include the absence of a component necessary for cell survival that is otherwise provided by expression of the metabolic selection marker and optionally the presence of a metabolic selection agent, as well as the presence of an antibiotic, in the cell culture media.

[0067] For example, in some embodiments, the cell culture media comprises a metabolic selection agent and an antibiotic.

[0068] In some embodiments, when the mammalian host cell comprises a dual selection expression vector system in which the metabolic selectable marker is glutamine synthetase, the metabolic selection agent is methionine sulfoximine. In some embodiments, when the mammalian host cell

comprises a dual selection expression system in which the metabolic selectable marker is dihydrofolate reductase, the metabolic selection agent is methotrexate.

[0069] In some embodiments, when the mammalian host cell comprises a dual selection expression system in which the antibiotic resistance selectable marker is a resistance marker to puromycin (e.g., the pac gene), the antibiotic is puromycin. In some embodiments, when the mammalian host cell comprises a dual selection expression system in which the antibiotic resistance selectable marker is a resistance marker to neomycin (e.g., the neo gene), the antibiotic is neomycin. In some embodiments, when the mammalian host cell comprises a dual selection expression system in which the antibiotic resistance selectable marker is a resistance marker to blasticidin (e.g., the bsr gene), the antibiotic is blasticidin. In some embodiments, when the mammalian host cell comprises a dual selection expression system in which the antibiotic resistance selectable marker is a resistance marker to phleomycin (e.g., the ble gene), the antibiotic is phleomycin D. In some embodiments, when the mammalian host cell comprises a dual selection expression system in which the antibiotic resistance selectable marker is a resistance marker to hygromycin (e.g., the hph gene), the antibiotic is hygromycin B.

[0070] In some embodiments, when the mammalian host cell comprises a dual selection expression system in which the metabolic selectable marker is glutamine synthetase and the antibiotic resistance selectable marker is a resistance marker to hygromycin (e.g., the hph gene), the metabolic selection agent is methionine sulfoximine and the antibiotic is hygromycin B.

[0071] Additionally, in some embodiments, when the mammalian host cell comprises a dual selection expression system in which the metabolic selectable marker is glutamine synthetase, the cell culture media does not comprise glutamine. In some embodiments, when the mammalian host cell comprises a dual selection expression system in which the metabolic selectable marker is glutamine synthetase, the cell culture media comprises the metabolic selection agent methionine sulfoximine but does not comprise glutamine.

[0072] Additionally, in some embodiments, when the mammalian host cell comprises a dual selection expression system in which the metabolic selectable marker is glutamine synthetase and the antibiotic resistance selectable marker is a resistance marker to hygromycin (e.g., the hph gene), the cell culture media comprises the antibiotic hygromycin B but does not comprise glutamine. In some embodiments, when the mammalian host cell comprises a dual selection expression system in which the metabolic selectable marker is glutamine synthetase and the antibiotic resistance selectable marker is a resistance marker to hygromycin (e.g., the hph gene), the cell culture media comprises the metabolic selection agent methionine sulfoximine and the antibiotic hygromycin B but does not comprise glutamine.

[0073] In some embodiments, when the mammalian host cell comprises a dual selection expression system in which the metabolic selectable marker is dihydrofolate reductase, the cell culture media

does not comprise hypoxanthine or thymidine. In some embodiments, when the mammalian host cell comprises a dual selection expression system in which the metabolic selectable marker is dihydrofolate reductase, the cell culture media comprises the metabolic selection agent methotrexate but does not comprise hypoxanthine or thymidine.

[0074] Illustratively, the present disclosure provides for a method for producing a recombinant protein (e.g., an antibody modality) comprising a) culturing a mammalian host cell comprising 1) a first expression vector encoding a first light chain, a first heavy chain or a first heavy chain fusion, and a metabolic selectable marker; and 2) a second expression vector encoding a second light chain, a second heavy chain or second heavy chain fusion, and an antibiotic selectable marker; under conditions in which the antibody chains and selectable markers are expressed, wherein the culturing is in a media comprising a metabolic selection agent and an antibiotic; and b) recovering the antibody modality from the culture.

[0075] In certain embodiments, the metabolic selectable marker is glutamine synthetase and the metabolic selection agent is methionine sulfoximine or the metabolic selection marker is dihydrofolate reductase and the metabolic selection agent is methotrexate. In certain embodiments, the antibiotic resistance marker is hph gene and the antibiotic is hygromycin B, the antibiotic resistance marker is the pac gene and the antibiotic is puromycin, the antibiotic resistance marker is neo gene and the antibiotic is neomycin, the antibiotic resistance marker is bsr gene and the antibiotic is blasticidin, or the antibiotic resistance marker is the ble gene and the antibiotic is phleomycin D.

[0076] In one aspect of this method, a) the first expression vector comprises a nucleotide sequence which comprises the following elements in 5' to 3' order: 1) a first promoter operably linked to a nucleotide sequence encoding an antibody light chain followed by a first polyA sequence; 2) a second promoter operably linked to a nucleotide sequence encoding an antibody heavy chain followed by a second polyA sequence; and 3) a promoter operably linked to a nucleotide sequence encoding a metabolic selectable marker followed by a third polyA sequence; and b) the second expression vector comprises a nucleotide sequence which comprises the following elements in 5' to 3' order: 1) a first promoter operably linked to a nucleotide sequence encoding an antibody light chain followed by a first polyA sequence; 2) a second promoter operably linked to a nucleotide sequence encoding an antibody heavy chain-scFv fusion followed by a second polyA sequence; and 3) a promoter operably linked to a nucleotide sequence encoding a antibiotic resistance selectable marker followed by a third polyA sequence. In one aspect, the method encodes a C1 mAb.

[0077] In another aspect of this method, a) the first expression vector comprises a nucleotide sequence which comprises the following elements in 5' to 3' order: 1) a first promoter operably linked to a nucleotide sequence encoding a first antibody light chain followed by a first polyA sequence; 2) a second promoter operably linked to a nucleotide sequence encoding an antibody heavy chain followed by a second polyA sequence; and 3) a promoter operably linked to a nucleotide sequence encoding a

metabolic selectable marker followed by a third polyA sequence; and b) the second expression vector comprises a nucleotide sequence which comprises the following elements in 5' to 3' order: 1) a first promoter operably linked to a nucleotide sequence encoding an antibody light chain followed by a first polyA sequence; 2) a second promoter operably linked to a nucleotide sequence encoding a second antibody heavy chain followed by a second polyA sequence; and 3) a promoter operably linked to a nucleotide sequence encoding a antibiotic resistance selectable marker followed by a third polyA sequence, wherein the first antibody heavy chain and the second heavy chain are different. In one sub-aspect, the method encodes a heteroIgG.

[0078] In yet another aspect, a) the first expression vector comprises a nucleotide sequence which comprises the following elements in 5' to 3' order: 1) a first promoter operably linked to a nucleotide sequence encoding an antibody light chain followed by a first polyA sequence; 2) a second promoter operably linked to a nucleotide sequence encoding an antibody heavy chain followed by a second polyA sequence; and 3) a promoter operably linked to a nucleotide sequence encoding a metabolic selectable marker followed by a third polyA sequence; and b) the second expression vector comprises a nucleotide sequence which comprises the following elements in 5' to 3' order: 1) a first promoter operably linked to a nucleotide sequence encoding an antibody light chain followed by a first polyA sequence; 2) a second promoter operably linked to a nucleotide sequence encoding an antibody heavy chain-cytokine fusion followed by a second polyA sequence; and 3) a promoter operably linked to a nucleotide sequence encoding an antibiotic resistance selectable marker followed by a third polyA sequence, wherein the first antibody heavy chain and the second antibody heavy chain are different. In one sub-aspect, the mammalian host cell encodes an antibody cytokine-fusion.

[0079] In certain embodiments of these methods, the mammalian host cell is a Chinese Hamster Ovary (CHO) cell. In one aspect, the CHO cell is dihydrofolate reductase deficient (dhfr-) or a glutamine synthetase knock out (GSKO).

[0080] In some embodiments, the CHO cell is dihydrofolate reductase deficient (dhfr-). In some embodiments, the CHO cell is dihydrofolate reductase deficient (dhfr-), and the metabolic selectable marker is dihydrofolate reductase.

[0081] In some embodiments, the CHO cell is a glutamine synthetase knock out (GSKO). In some embodiments, the CHO cell is a glutamine synthetase knock out (GSKO), and the metabolic selectable marker is glutamine synthetase.

[0082] In certain embodiments of any of these methods, the recovered antibody modality is purified and formulated in a pharmaceutically acceptable formulation.

BRIEF DESCRIPTION OF THE DRAWINGS

[0083] FIGs. 1A-1F provide schematics of several different complex biologic modalities: a Fab-heteroFc-[scFv*] molecule (FIG. 1A); a heteroIgG molecule (FIG. 1B); an asymmetric antibody-

cytokine fusion (FIG. 1C); a [VH-VH*Fab]-heteroFc molecule (FIG. 1D); a [VH*Fab]-heteroFc molecule (FIG. 1E); and a [Fab*]-heteroFc-[VH*VH] (FIG. 1F) molecule.

[0084] FIGs. 2A-2G provide A) a schematic of vectors used to evaluate the dual selection strategy with a 3-chain C1mAb molecule; B) Recovery times; C) Bulk titer in g/L; D) Effective pool titer in g/L; E) SEC-HPLC-HMW%, F) % nrCE pre-peak impurities; and G) rCE Profile: HC-rCE-SDS %, HC-ScFv-rCE-SDS % and LC-rCE-SDS %. Shown between lines is the ideal ratio for a perfectly assembled molecule. Titer and product quality were measured on day 10 of a fed-batch production. Data is presented as mean values for two independent transfections. Error bars represent the standard deviation (SD) of two biological replicates for fed-batch production. GAPDH/CMV refers to a promoter. LC refers to a light chain. HC1 refers to a first heavy chain. HC2-scFV refers to a second heavy chain fused to a scFv sequence. mGS refers to a mouse glutamine synthetase. Hygr B refers to hygromycin B.

[0085] FIGs. 3A-3F provide A) a schematic of vectors used to evaluate the dual selection strategy with a 4-chain HeteroIgG molecule; B) Recovery times; C) Bulk titer in g/L; D) Effective pool titer in g/L; E) SEC-HPLC-HMW%, F) % nrCE pre-peak impurities. Titer and product quality were measured on day 10 of a fed-batch production. Data is presented as mean values for two independent transfections. Error bars represent the standard deviation (SD) of two biological replicates for fed-batch production. CMV/GapDH refers to a promoter. LC1 refers to a first light chain. LC2 refers to a second light chain. HC1 refers to a first heavy chain. HC2 refers to a second heavy chain. GS refers to a glutamine synthetase. Hygro refers to hygromycin B.

[0086] FIGs. 4A-4G provide A) a schematic of vectors used to evaluate the dual selection strategy with an antibody cytokine-fusion molecule; B) Recovery times; C) Bulk titer in g/L; D) Effective pool titer in g/L; E) SEC-HPLC-HMW%, F) % nrCE pre-peak impurities; and G) rCE Profile: HC-rCE-SDS %, HC-ScFv-rCE-SDS % and LC-rCE-SDS %. Shown between lines is the ideal ratio for a perfectly assembled molecule. Titer and product quality were measured on day 10 of a fed-batch production. Data is presented as mean values for two independent transfections. Error bars represent the standard deviation (SD) of two biological replicates for fed-batch production. LC refers to a light chain. HC1 refers to a first heavy chain. HC2 refers to a second heavy chain. mGS refers to a mouse glutamine synthetase. Hyg B refers to hygromycin B. Attachment of the cytokine sequence is not shown.

[0087] FIGs. 5A-5G provide A) a schematic of vectors used to evaluate the dual selection strategy with a [VH-VH*Fab]-heteroFc molecule; B) recovery times; C) normalized bulk titer (normalized to the single selection condition); D) normalized effective titer (normalized to the single selection condition); E) SEC-HPLC-HMW%, F) % nrCE pre-peak impurities; and G) rCE Profile: HC1-rCE-SDS %, HC2- rCE-SDS %, and LC-rCE-SDS %. Titer and product quality were measured on day 14 of a fed-batch production. Data are presented as mean values for two independent transfections and

two technical replicates for fed-batch production. LC refers to a light chain. HC1 refers to a first heavy chain. HC2 refers to a second heavy chain. mGS refers to a mouse glutamine synthetase. Hyg B refers to hygromycin B.

[0088] FIGs. 6A-6G provide A) a schematic of vectors used to evaluate the dual selection strategy with two [VH*Fab]-Fc molecules (molecules A and B); B) recovery times; C) normalized bulk titer (normalized to the single selection condition); D) normalized effective titer (normalized to the single selection condition); E) SEC-HPLC-HMW%, F) % nrCE pre-peak impurities; and G) rCE Profile: HC1-rCE-SDS %, HC2- rCE-SDS %, and LC-rCE-SDS %. Titer and product quality were measured on day 14 of a fed-batch production. Data are presented as mean values for two independent transfections and two technical replicates for fed-batch production. LC refers to a light chain. HC1 refers to a first heavy chain. HC2 refers to a second heavy chain. mGS refers to a mouse glutamine synthetase. Hyg B refers to hygromycin B.

[0089] FIGs. 7A-7H provide A) a schematic of vectors used to evaluate the dual selection strategy with a [Fab*]-heteroFc-[VH*VH] molecule; B) recovery times; C) normalized bulk titer (normalized to the single selection condition); D) normalized effective titer (normalized to the single selection condition); E) SEC-HPLC-HMW%, F) % nrCE pre-peak impurities; G) % nrCE post-peak impurities; and H) rCE Profile: HC1-rCE-SDS %, HC2- rCE-SDS %, and LC-rCE-SDS %. Titer and product quality were measured on day 14 of a fed-batch production. Data are presented as mean values for two independent transfections and two technical replicates for fed-batch production. LC refers to a light chain. HC1 refers to a first heavy chain. HC2 refers to a second heavy chain. mGS refers to a mouse glutamine synthetase. Hyg refers to hygromycin B.

DETAILED DESCRIPTION

[0090] The present disclosure is based on a vector system design that utilizes dual selection in a mammalian cell, where one vector utilizes metabolic selection, while the other vector utilizes antibiotic selection. Accordingly, the disclosure is directed to a mammalian host cell comprising a pair of expression vectors, wherein a) the first expression vector comprises nucleotide sequences encoding a first antibody light chain, a first antibody heavy chain, antibody heavy chain fusion, or Fc chain fusion (e.g., a first antibody heavy chain or antibody heavy chain fusion), and a metabolic selectable marker; and b) the second expression vector comprises nucleotide sequences encoding a second antibody light chain, a second antibody heavy chain, antibody heavy chain fusion, or Fc chain fusion (e.g., a second antibody heavy chain or antibody heavy chain fusion), and an antibiotic resistance selectable marker, wherein the antibiotic resistance selectable marker is expressed in the mammalian host cell. Surprisingly, this design increases specific productivity and titer and reduces impurities compared to the standard single metabolic selection method (e.g., using glutamine synthetase). This dual selection strategy ensures the integration and expression of all the antibody

chains from both vectors, therefore impacting the productivity and controlling the chain ratios that improve product quality.

[0091] In the various embodiments described herein, the first and second expression vectors are always different due to the different selectable markers, but the first and second antibody light chains can be the same (i.e., having the same sequence) or different. The first antibody heavy chain and the second antibody heavy chain can be the same or different. In certain embodiments, one expression vector encodes a heavy chain and the other expression vector encodes a heavy chain fusion, wherein the heavy chain in the heavy chain fusion is linked to, for example, a scFv, cytokine, VH/VHH, etc. In other embodiments, one expression vector encodes an antibody heavy chain and the other expression vector encodes an Fc chain fusion, wherein the heavy chain in the heavy chain fusion is linked to, for example, a scFv, cytokine, VH/VHH, etc.

[0092] This strategy is particularly useful for antibody modalities having three- or four- unique chains expressed on two different vectors. An example of a three- or four-chain antibody modality can have a heavy chain, a heavy chain-scFv fusion, and two light chains (the light chains being the same in the case of a 3-chain antibody modality or different in the case of a 4-chain modality). Another example of a four-chain antibody modality has two different heavy chains and two different light chains, where each heavy chain - light-chain pair is expressed from a different vector. Another example of a three- or four-chain antibody modality can have a heavy chain, a heavy chain-cytokine fusion, and two light chains (the light chains being the same in the case of a 3-chain antibody modality or different in the case of a 4-chain modality). Yet another example of a three-chain antibody modality comprises one light chain, one heavy chain, and one heavy chain fusion in which a VH is fused to the heavy chain at its N-terminus. Still another example of a three-chain antibody modality comprises one light chain, one heavy chain, and an Fc chain fusion in which a VH is fused to the N-terminus of the Fc. In still another example, the three-chain antibody modality comprises one light chain, one heavy chain fusion in which a VH is fused to the heavy chain at its C-terminus, and one Fc chain fusion in which a VH is fused to the Fc at its C-terminus.

[0093] Certain examples in this application show this dual selection strategy being used to express a 3-chain C1mAb (Fab-heteroFc-[scFv*]) molecule, a 4-chain heteroIgG molecule, and a 3-chain antibody cytokine-fusion. Two plasmids were used to express the antibody chains: sequences encoding a first heavy chain and a first light chain were constructed in a first plasmid under the GS selectable marker controlled by the Sra promoter, while the second plasmid contains sequences encoding a second heavy chain or a heavy chain fusion and the second light chain under the Hygromycin B antibiotic selectable marker controlled by the SV40 promoter. The use of double selection improved productivity by a factor of 1.2 to 3.5 times and reduced impurities (pre-peaks) by a factor of 3.8 to 4.5 times in 4-chain and 3-chain molecules, respectively. Additional examples illustrate that a dual selection strategy of the present disclosure increases bulk and effective titers and

reduces impurities for certain [VH-VH*Fab]-heteroFc, [VH*Fab]-heteroFc, and [Fab*]-heteroFc-[VH*VH] molecules.

[0094] By employing dual expression vector systems and mammalian host cells described herein, production of recombinant proteins can be increased while retaining or improving product quality. For example, by using the dual expression vectors described herein in mammalian host production cell lines, biopharmaceuticals can be produced in a less expensive and more consistent manner. These inventions may find particular utility in the commercial production of polypeptides (e.g., antibody modalities) having three or four unique polypeptide chains.

[0095] The dual expression vectors described herein are employed in cell lines (also referred to as "host cells"), preferably mammalian ("mammalian host cells"), grown in cell culture media to produce a recombinant protein of commercial or scientific interest. Cell lines are typically derived from a lineage arising from a primary culture that can be maintained in culture for an unlimited time. Genetically engineering the cell line involves transfecting, transforming, or transducing the cells with two expression vectors where each vector contains nucleotide sequences encoding two antibody chains so as to cause the host cell to express an antibody modality having the desired number of chains. Methods and vectors for genetically engineering cells and/or cell lines to express, for example, a protein of interest, are well known to those of skill in the art; for example, various techniques are illustrated in Current Protocols in Molecular Biology, Ausubel *et al*., eds. (Wiley & Sons, New York, 1988, and quarterly updates); Sambrook *et al*., Molecular Cloning: A Laboratory Manual (Cold Spring Laboratory Press, 1989); Kaufman, R.J., *Large Scale Mammalian Cell Culture*, 1990, pp. 15-69; and Harlow and Lane Antibodies: A Laboratory Manual Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1990).

DEFINITIONS

[0096] While the terminology used in this application is standard within the art, definitions of certain terms are provided herein to assure clarity and definiteness in the meaning of the claims. Units, prefixes, and symbols may be denoted in their SI (International System of Units) accepted form. Numeric ranges recited herein are inclusive of the numbers defining the range and include and are supportive of each integer within the defined range. The methods and techniques described herein are generally performed according to conventional methods well-known in the art and as described in various general and more specific references that are cited and discussed throughout the present specification unless otherwise indicated.

[0097] As used herein, the terms "a" and "an" mean one or more unless specifically indicated otherwise. Further, unless otherwise required by context, singular terms shall include pluralities and plural terms shall include the singular. Generally, nomenclatures used in connection with, and techniques of, cell and tissue culture, molecular biology, immunology, microbiology, genetics and

protein and nucleic acid chemistry and hybridization described herein are those well-known and commonly used in the art.

[0098] All documents, or portions of documents, cited in this application, including but not limited to patents, patent applications, articles, books, and treatises, are hereby expressly incorporated by reference. What is described in an embodiment of the invention can be combined with other embodiments of the invention.

[0099] The present disclosure provides systems and methods for expressing a "protein of interest" generally an antibody modality. A "protein of interest" includes naturally occurring proteins, recombinant proteins, and engineered proteins (e.g., proteins that do not occur in nature and which have been designed and/or created by humans). A protein of interest can, but need not be, a protein that is known or suspected to be therapeutically relevant.

[0100] As used herein, an "antibody chain" or "chain" refers to antibody light chains and antibody heavy chains. The terms "antibody heavy chain" and "antibody light chain" have the standard meaning in the art and include, for example, the various antibody heavy and light chains described elsewhere herein (e.g., heavy and light chains of IgG1, IgG2, IgG3, and IgG4 mAbs). The terms "antibody heavy chain" and "antibody light chain" include standard full-length antibody heavy chains and light chains.

[0101] As used herein, the terms "antibody heavy chain fusion" and "antibody heavy chain fusion protein" refer to a polypeptide that contains an antibody heavy chain covalently linked to one or more additional proteins or peptides. For example, an "antibody heavy chain fusion protein" can be an antibody heavy chain covalently linked to a cytokine, scFv, VH/VHH, and the like. The linkage may be direct, or via a peptide linker (e.g., a glycine-serine linker). In an antibody heavy chain fusion protein, the antibody heavy chain may be linked to additional protein(s) at the N- terminus or the C-terminus of the heavy chain (or both locations). The antibody heavy chain may also be linked to additional protein sequences, such as an scFv, at an internal amino acid residue or be between the Fab and the Fc. The terms "antibody light chain fusion protein" and "antibody light chain fusion" have the same meaning as described immediately above for "antibody heavy chain fusion protein", except for an antibody light chain. As used herein, an "antibody fusion protein" refers to an antibody as provided herein which is covalently linked to one or more additional proteins or polypeptides (e.g., via a heavy chain or light chain of the antibody). Thus, an antibody fusion protein contains at least an antibody heavy chain fusion protein or an antibody light chain fusion protein as one of the polypeptides of the antibody fusion protein. Most commonly, an antibody fusion protein is a molecule that contains two antibody light chains (which can be the same or different), one antibody heavy chain, and one antibody heavy chain fusion protein, such that the additional protein is linked to one of the heavy chains of the antibody. An "antibody chain fusion" encompasses both antibody heavy chain fusions and antibody light chain fusions.

[0102] In some embodiments, an antibody heavy chain fusion is a fusion of a heavy chain with a VH, scFv, or cytokine fused to either the N-terminus or the C-terminus of the antibody heavy chain portion of the antibody heavy chain fusion, or between the CH1 and CH2 of the antibody heavy chain portion of the antibody heavy chain fusion. The fusion may be direct or through a linker.

[0103] In some embodiments, an antibody heavy chain fusion is a fusion of an antibody heavy chain with a VH. In some embodiments, an antibody heavy chain fusion is a fusion of an antibody heavy chain with a scFv. In some embodiments, an antibody heavy chain fusion is a fusion of a antibody heavy chain with a cytokine.

[0104] In some embodiments, an antibody heavy chain fusion is a direct fusion of an antibody heavy chain with a VH. In some embodiments, an antibody heavy chain fusion is a direct fusion of an antibody heavy chain with a scFv. In some embodiments, an antibody heavy chain fusion is a direct fusion of an antibody heavy chain with a cytokine.

[0105] In some embodiments, an antibody heavy chain fusion is a fusion of an antibody heavy chain with a VH, wherein the antibody heavy chain fusion comprises a linker between the antibody heavy chain and the VH. In some embodiments, an antibody heavy chain fusion is a fusion of an antibody heavy chain with a scFv, wherein the antibody heavy chain fusion comprises a linker between the antibody heavy chain and the scFv. In some embodiments, an antibody heavy chain fusion is a fusion of an antibody heavy chain with a cytokine, wherein the antibody heavy chain fusion comprises a linker between the antibody heavy chain and the cytokine.

[0106] As used herein, an “antibody modality” refers to a protein having at least one antibody chain. An antibody modality can have two, three, or four unique chains, wherein any or all of the chains can include fusions. Reference to two, three, or four chain molecules implies that each chain is unique. Any and all of the antibody chain containing molecules described herein are considered antibody modalities.

[0107] As used herein, the terms “polypeptide” and “protein” (e.g., as used in the context of a protein of interest or a polypeptide of interest) are used interchangeably herein to refer to a polymer of amino acid residues. The terms also apply to amino acid polymers in which one or more amino acid residues is an analog or mimetic of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers. The terms can also encompass amino acid polymers that have been modified, e.g., by the addition of carbohydrate residues to form glycoproteins, or phosphorylated. Polypeptides and proteins can be produced by a naturally-occurring and non-recombinant cell, or polypeptides and proteins can be produced by a genetically-engineered or recombinant cell. Polypeptides and proteins can comprise molecules having the amino acid sequence of a native protein, or molecules having deletions from, additions to, and/or substitutions of one or more amino acids of the native sequence.

[0108] As used herein, the term "heterologous" used in connection with a nucleic acid means having a nucleic acid not naturally occurring within a host cell. This can include mutated sequences, *e.g.*,, sequences differing from the naturally occurring sequence. This can include sequences from other species. This can also include having a sequence at a different position in the genome than that naturally occurring in the host cell. This generally does not include natural mutations that may occur in a host cell. A cell already containing a heterologous nucleic acid encoding a protein of interest, for example, by stable integration of an expression cassette, would be considered to contain a heterologous nucleic acid sequence. For clarity, a CHO cell or a derivative thereof (e.g., a DHFR- or GS knockout) having a nucleic acid encoding an antigen-binding protein would be considered to have a heterologous nucleic acid.

[0109] As used herein, the term "operably linked" refers to that the nucleic acid sequences being linked are typically contiguous, or substantially contiguous, and, where necessary to join two protein coding regions, contiguous and in reading frame. However, since enhancers generally function when separated from the promoter by several kilobases and intronic sequences may be of variable lengths, some polynucleotide elements may be operably linked but not contiguous. Two or more nucleic acid sequences may be operably linked in such a manner that a nucleic acid molecule capable of directing the transcription of a given gene and/or the synthesis of a desired protein molecule is produced.

[0110] As used herein, the term "bioreactor" means any vessel useful for the growth of a cell culture. The cell culture of host cells of the instant disclosure can be performed in a bioreactor, which can be selected based on the application of a protein of interest that is produced by cells growing in the bioreactor. A bioreactor can be of any size so long as it is useful for the culturing of cells; typically, a bioreactor is sized appropriate to the volume of cell culture being grown inside of it. Typically, a bioreactor will be at least 1 liter and may be 2, 5, 10, 50, 100, 200, 250, 500, 1,000, 1500, 2000, 2,500, 5,000, 8,000, 10,000, 12,000 liters or more, or any volume in between. The internal conditions of the bioreactor, including, but not limited to pH and temperature, can be controlled during the culturing period. Those of ordinary skill in the art will be aware of, and will be able to select, suitable bioreactors for use in practicing the methods disclosed herein based on the relevant considerations. In one embodiment, 500L to 2000L bioreactors are used. In one embodiment, 1000L to 2000L bioreactors are used.

[0111] As used herein, "cell culture" or "culture" is meant the growth and propagation of cells outside of a multicellular organism or tissue. Suitable culture conditions for mammalian cells are known in the art. *See e.g.* Animal cell culture: A Practical Approach, D. Rickwood, ed., Oxford University Press, New York (1992). Mammalian cells may be cultured in suspension or while attached to a solid substrate. Fluidized bed bioreactors, hollow fiber bioreactors, roller bottles, shake flasks, or stirred tank bioreactors, with or without microcarriers, can be used.

[0112] The term "cell culture medium" (also called "culture medium," "cell culture media," "tissue culture media") refers to any nutrient solution used for growing cells, e.g., animal or mammalian cells, and which generally provides at least one or more components from the following: an energy source (usually in the form of a carbohydrate such as glucose); one or more of all essential amino acids, and generally the twenty basic amino acids, plus cysteine; vitamins and/or other organic compounds typically required at low concentrations; lipids or free fatty acids; and trace elements, e.g., inorganic compounds or naturally occurring elements that are typically required at very low concentrations, usually in the micromolar range.

[0113] The nutrient solution may optionally be supplemented with additional optional components to optimize growth of cells, such as hormones and other growth factors, e.g., transferrin, epidermal growth factor, insulin-like growth factor, insulin, serum, and the like; salts, e.g., calcium, magnesium and phosphate, and buffers, e.g., HEPES; nucleosides and bases, e.g., adenosine, thymidine, hypoxanthine; and protein and tissue hydrolysates, e.g., hydrolyzed animal or plant protein (peptone or peptone mixtures, which can be obtained from animal byproducts, purified gelatin or plant material); antibiotics, e.g., gentamycin; anti-clumping agents; cell protectants or surfactants such as Pluronic®F68 (also referred to as Lutrol® F68 and Kolliphor® P188; nonionic triblock composed of a central hydrophobic chain of polyoxypropylene (poly(propylene oxide)) flanked by two hydrophilic chains of polyoxyethylene (poly(ethylene oxide)); polyamines, e.g., putrescine, spermidine and spermine (see e.g., International Patent Application Publication No. WO 2008/154014) and pyruvate (see e.g. U.S. Pat. No. 8,053,238) depending on the requirements of the cells to be cultured and/or the desired cell culture parameters.

[0114] Cell culture media include those that are typically employed in and/or are known for use with any cell culture process, such as, but not limited to, batch, extended batch, fed-batch, and/or perfusion or continuous culturing of cells.

[0115] A "base" (or batch) cell culture medium refers to a cell culture medium that is typically used to initiate a cell culture and is sufficiently complete to support the cell culture.

[0116] A "fed-batch culture" refers to a form of suspension culture and means a method of culturing cells in which additional components are provided to the culture at a time or times subsequent to the beginning of the culture process. The provided components typically comprise nutritional supplements for the cells which have been depleted during the culturing process. Additionally or alternatively, the additional components may include supplementary components. A fed-batch culture is typically stopped at some point and the cells and/or components in the medium are harvested and optionally purified.

[0117] A "growth" cell culture medium refers to a cell culture medium that is typically used in cell cultures during a period of exponential growth, i.e., a "growth phase", and is sufficiently complete to support the cell culture during this phase. A growth cell culture medium may also contain selection

agents that confer resistance or survival to selectable markers incorporated into the host cell line. Such selection agents include, but are not limited to, geneticin (G418), neomycin, hygromycin B, puromycin, zeocin, methionine sulfoximine, methotrexate, glutamine-free cell culture medium, cell culture medium lacking glycine, hypoxanthine and thymidine, or thymidine alone.

[0118] A "perfusion" cell culture medium refers to a cell culture medium that is typically used in cell cultures that are maintained by perfusion or continuous culture methods and is sufficiently complete to support the cell culture during this process. Perfusion cell culture medium formulations may be richer or more concentrated than base cell culture medium formulations to accommodate the method used to remove the spent medium. Perfusion cell culture medium can be used during both the growth and production phases.

[0119] A "production" cell culture medium refers to a cell culture medium that is typically used in cell cultures during the transition when exponential growth is ending and protein production takes over, "transition" and/or "product" phases, and is sufficiently complete to maintain a desired cell density, viability, and/or product titer during this phase.

[0120] Concentrated cell culture medium can contain some or all of the nutrients necessary to maintain the cell culture; in particular, concentrated medium can contain nutrients identified as or known to be consumed during the course of the production phase of the cell culture. Concentrated medium may be based on just about any cell culture media formulation. Such a concentrated feed medium can contain some or all the components of the cell culture medium at, for example, about 2X, 3X, 4X, 5X, 6X, 7X, 8X, 9X, 10X, 12X, 14X, 16X, 20X, 30X, 50X, 100x, 200X, 400X, 600X, 800X, or even about 1000X of their normal amount.

[0121] The components used to prepare cell culture medium may be completely milled into a powder medium formulation; partially milled with liquid supplements added to the cell culture medium as needed; or added in a completely liquid form to the cell culture.

[0122] Cell cultures can also be supplemented with independent concentrated feeds of particular nutrients which may be difficult to formulate or are quickly depleted in cell cultures. Such nutrients may be amino acids such as tyrosine, cysteine and/or cystine (see e.g., International Patent Application Publication No. WO2012/145682). The independent feeds can begin prior to or at the start of the production phase. The independent feeds can be accomplished by fed batch to the cell culture medium on the same or different days as the concentrated feed medium. The independent feeds can also be perfused on the same or different days as the perfused medium.

[0123] "Serum-free" applies to a cell culture medium that does not contain animal sera, such as fetal bovine serum. Various tissue culture media, including defined culture media, are commercially available, for example, any one or a combination of the following cell culture media can be used: RPMI-1640 Medium, RPMI-1641 Medium, Dulbecco's Modified Eagle's Medium (DMEM), Minimum Essential Medium Eagle, F-12K Medium, Ham's F12 Medium, Iscove's Modified

Dulbecco's Medium, McCoy's 5A Medium, Leibovitz's L-15 Medium, and serum-free media such as EX-CELL™ 300 Series (JRH Biosciences, Lenexa, Kansas), MCDB 302 (Sigma Aldrich Corp., St. Louis, MO), among others. Serum-free versions of such culture media are also available. Cell culture media may be supplemented with additional or increased concentrations of components such as amino acids, salts, sugars, vitamins, hormones, growth factors, buffers, antibiotics, lipids, trace elements, and the like, depending on the requirements of the cells to be cultured and/or the desired cell culture parameters. Customized cell culture media can also be used.

[0124] "Titer" means the total amount of a polypeptide or protein of interest (which may be a naturally occurring or recombinant protein of interest) produced by a cell culture in a given amount of medium volume. Titer can be expressed in units of milligrams or micrograms of polypeptide or protein per milliliter (or other measure of volume) of medium. "Cumulative titer" is the titer produced by the cells during the course of the culture, and can be determined, for example, by measuring daily titers and using those values to calculate the cumulative titer.

[0125] As used herein, the term "host cell" is understood to include a cell that has been genetically engineered to express a polypeptide of interest. Genetically engineering a cell involves transfecting, transforming, or transducing the cell with a nucleic acid encoding a recombinant polynucleotide molecule (a "gene of interest"), and/or otherwise altering (e.g., by homologous recombination and gene activation or fusion of a recombinant cell with a non-recombinant cell) so as to cause the host cell to express a desired recombinant polypeptide. Methods and vectors for genetically engineering cells and/or cell lines to express a polypeptide of interest are well known to those of skill in the art; for example, various techniques are illustrated in Current Protocols in Molecular Biology. Ausubel *et al*., eds. (Wiley & Sons, New York, 1988, and quarterly updates); Sambrook *et al*., Molecular Cloning: A Laboratory Manual (Cold Spring Laboratory Press, 1989); Kaufman, R.J., Large Scale Mammalian Cell Culture, 1990, pp. 15-69. The term includes the progeny of the parent cell, whether or not the progeny is identical in morphology or in genetic makeup to the original parent cell, so long as the gene of interest is present. A cell culture can comprise one or more host cells. "Effective titer" refers to bulk titer normalized to percentage of correctly paired product (i.e., percent main peak SEC or percent main peak nrCE).

[0126] An "Fc" region, as the term is used herein, comprises two heavy chain fragments comprising the C_H2 and C_H3 domains of an antibody. The two heavy chain fragments are held together by two or more disulfide bonds and by hydrophobic interactions of the C_H3 domains.

[0127] As used herein, an "Fc chain fusion" refers to a heavy chain fragment comprising a C_H2 and C_H3 domain covalently linked to one or more additional proteins or peptides. For example, an "Fc chain fusion" can be a heavy chain fragment comprising a C_H2 and C_H3 domain covalently linked to a cytokine, scFv, VH/VHH, and the like. The linkage may be direct, or via a peptide linker (e.g., a glycine-serine linker). In an Fc chain fusion protein, the Fc chain may be linked to additional

protein(s) at the N- terminus or the C-terminus of the heavy chain fragment (or both locations). The heavy chain fragment may also be linked to additional protein sequences, such as an scFv, at an internal amino acid residue.

[0128] In some embodiments, an Fc chain fusion is a fusion of an Fc chain with a VH, scFv, or cytokine, wherein the VH, scFv, or cytokine is fused to the N-terminus or the C-terminus of the Fc chain. In some embodiments, an Fc chain fusion is a fusion of an Fc chain with a VH, scFv, or cytokine, wherein the VH, scFv, or cytokine is fused to the N-terminus of the Fc chain. In some embodiments, an Fc chain fusion is a fusion of an Fc chain with a VH, scFv, or cytokine, wherein the VH, scFv, or cytokine is fused to the C-terminus of the Fc chain. The fusion may be direct or through a linker.

[0129] In some embodiments, an Fc chain fusion is a fusion of an Fc chain with a VH. In some embodiments, an Fc chain fusion is a fusion of an Fc chain with a scFv. In some embodiments, an Fc chain fusion is a fusion of a Fc chain with a cytokine.

[0130] In some embodiments, an Fc chain fusion is a direct fusion of an Fc chain with a VH. In some embodiments, an Fc chain fusion is a direct fusion of an Fc chain with a scFv. In some embodiments, an Fc chain fusion is a direct fusion of an Fc chain with a cytokine.

[0131] In some embodiments, an Fc chain fusion is a fusion of an Fc chain with a VH, wherein the Fc chain fusion comprises a linker between the Fc chain and the VH. In some embodiments, an Fc chain fusion is a fusion of an Fc chain with a scFv, wherein the Fc chain fusion comprises a linker between the Fc chain and the scFv. In some embodiments, an Fc chain fusion is a fusion of an Fc chain with a cytokine, wherein the Fc chain fusion comprises a linker between the Fc chain and the cytokine.

[0132] A “hemibody” is an immunologically functional immunoglobulin construct comprising a complete heavy chain, a complete light chain and a second heavy chain Fc region paired with the Fc region of the complete heavy chain. A linker can, but need not, be employed to join the heavy chain Fc region and the second heavy chain Fc region. In particular embodiments, a hemibody is a monovalent form of an antigen-binding protein disclosed herein. In other embodiments, pairs of charged residues can be employed to associate one Fc region with the second Fc region.

[0133] It is understood that wherever embodiments are described herein with the language “comprising,” otherwise analogous embodiments described in terms of “consisting of” and/or “consisting essentially of” are also provided.

SELECTABLE MARKERS

[0134] For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable

marker (which is expressed in the mammalian cell) is generally introduced into the host cells in the same expression vector as the gene(s) of interest.

[0135] A selectable marker gene encodes a protein necessary for the survival and growth of a host cell grown in a selective culture medium. Typical selection marker genes encode proteins that (a) confer resistance to antibiotics or other toxins; (b) complement auxotrophic deficiencies of the cell; or (c) supply critical nutrients not available from complex or defined media through metabolism. The latter two can be considered metabolic selectable markers. The vector systems disclosed herein utilize one antibiotic resistance selectable marker and one metabolic selectable marker. Both of these markers are expressed in the mammalian host cell. This is distinguished from antibiotic resistance selectable markers expressed in prokaryotes. Thus, in some embodiments, the expression vector containing an antibiotic resistance selectable marker will have no metabolic selectable marker. Unless otherwise specified, reference herein to an antibiotic resistance selectable marker refers to one which is expressed in mammalian cells.

[0136] Specific antibiotic resistance selectable markers that are expressed in mammalian host cells are the kanamycin resistance gene, the ampicillin resistance gene, the tetracycline resistance gene, the hygromycin B resistance gene, the puromycin resistance gene, and the neomycin resistance gene. The antibiotic resistance selectable markers in widespread use are the NeoR (neo), BsdR (bsr), HygR (hph), PuroR (pac), and BleoR (ble) genes, which confer resistance to the selective antibiotics G418/geneticin, blasticidin, hygromycin B, puromycin, and phleomycin D, respectively. See, e.g., Guo et al., 2021, J. Biol Chem. 297:100838. Antibiotic resistance selectable markers may be described herein either by the gene name or the name of the corresponding antibiotic selection agent (i.e., hph and resistance marker to hygromycin B, pac and resistance marker to puromycin, etc. may be used interchangeably herein when referring to an antibiotic resistance selectable marker used in an expression vector). Antibiotic resistance marker and antibiotic resistance selectable marker may also be used interchangeably herein.

[0137] In some embodiments of the present disclosure, the antibiotic resistance selectable marker used in an expression vector is selected from the NeoR (neo), BsdR (bsr), HygR (hph), PuroR (pac), and BleoR (ble) genes. In some embodiments, the antibiotic resistance selectable marker used in an expression vector is the NeoR (neo) gene. In some embodiments, the antibiotic resistance selectable marker used in an expression vector is the BsdR (bsr) gene. In some embodiments, the antibiotic resistance selectable marker used in an expression vector is the HygR (hph) gene. In some embodiments, the antibiotic resistance selectable marker used in an expression vector is the PuroR (pac) gene. In some embodiments, the antibiotic resistance selectable marker used in an expression vector is the BleoR (ble) gene.

[0138] Specific metabolic selectable markers that are expressed in mammalian cells include the glutamine synthetase (GS), dihydrofolate reductase (DHFR), asparaginase (Aspg; see Ha et al.

Biotechnol Bioeng. 2023 120:1159-1166), and promoterless thymidine kinase genes. In some embodiments, the metabolic selectable marker used in an expression vector is the glutamine synthetase. In some embodiments, the metabolic selectable marker used in an expression vector is dihydrofolate reductase. In some embodiments, the metabolic selectable marker used in an expression vector is asparaginase. In some embodiments, the metabolic selectable marker used in an expression vector is promoterless thymidine kinase.

[0139] Mammalian cell transformants are placed under selection pressure wherein only the transformants are uniquely adapted to survive by virtue of the two selectable genes present in the two expression vectors. Selection pressure is imposed by culturing the transformed cells under conditions in which the concentration of the selection agents in the medium are successively increased, thereby leading to additional stringency and/or amplification of both the selectable genes and the DNAs that encodes the protein of interest. As a result, increased quantities of a polypeptide of interest are synthesized from the amplified DNA.

[0140] In certain embodiments, one selectable marker (metabolic) is glutamine synthetase or dihydrofolate reductase and the other selectable marker (antibiotic) is HygR. The selection agent for GS is methionine sulfoximine (MSX). The selection agent for DHFR is methotrexate (MTX). The selection agent for HygR is hygromycin B.

[0141] Glutamine synthetase (GS) catalyzes glutamine biosynthesis by the condensation of ammonia with glutamate. GS knockout cell lines (GSKO) provide sufficient selection stringency without MSX or with low MSX concentrations, while 25 mM MSX coupled with the GS-knockout cell line leads to higher selection efficiency compared with CHOK1SV cell lines at higher MSX concentrations. See Fan et al., 2012, Biotechnol Bioeng. 109(4):1007-1015. A previous report showed that increasing the MSX concentration in the seed train stage after clone selection increased productivity without significant impacts on cell growth, GS and target gene copy numbers and expression, and maintained product quality attributes in multiple GS knockout cell lines. See Tian et al., 2020, Engineering in Life Sciences 20(3-4): 112-125. Chain/vector expression can be influenced by increasing stringency during pool recovery/selection by adding MSX.

[0142] In certain embodiments, the promoter Sra (also referred to as SRα and Srα herein) or mPGK is operably linked to the metabolic selectable marker and the SV40 promoter is operably linked to the antibiotic resistance selectable marker. In some embodiments, the promoter Sra is operably linked to the metabolic selectable marker, and the SV40 promoter is operably linked to the antibiotic resistance selectable marker. In some embodiments, the promoter mPGK is operably linked to the metabolic selectable marker, and the SV40 promoter is operably linked to the antibiotic resistance selectable marker.

[0143] In some embodiments, the metabolic selectable marker is glutamine synthetase, the promoter mPGK is operably linked to the metabolic selectable marker, the antibiotic selectable marker is HygR, and the SV40 promoter is operably linked to the antibiotic resistance selectable marker.

[0144] In some embodiments, the metabolic selectable marker is glutamine synthetase, the promoter SRα is operably linked to the metabolic selectable marker, the antibiotic selectable marker is HygR, and the SV40 promoter is operably linked to the antibiotic resistance selectable marker.

[0145] In some embodiments, the SRa promoter comprises a nucleotide sequence that is 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to the nucleotide sequence of SEQ ID NO: 2. In some embodiments, the SRa promoter comprises the nucleotide sequence of SEQ ID NO: 2.

[0146] In some embodiments, the mPGK promoter comprises a nucleotide sequence that is 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to the nucleotide sequence of SEQ ID NO: 3. In some embodiments, the mPGK promoter comprises the nucleotide sequence of SEQ ID NO: 3.

[0147] In some embodiments, the SV40 promoter comprises a nucleotide sequence that is 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to the nucleotide sequence of SEQ ID NO: 4. In some embodiments, the SV40 promoter comprises the nucleotide sequence of SEQ ID NO: 4.

EXPRESSION VECTORS

[0148] The selectable markers disclosed herein will be incorporated into expression vectors that will typically include expression cassettes that contain one or more promoters that are recognized by the host organism and are operably linked to the nucleotide sequences encoding the antibody chains and the selectable markers. Promoters are untranscribed sequences located upstream (i.e., 5′) to the start codon of a structural gene (generally within about 100 to 1000 bp) that control transcription of the structural gene. “Promoter” and “promoter sequence” may be used interchangeably in embodiments described herein.

[0149] A typical antibody is a Y-shaped molecule having four polypeptide chains – two identical heavy chains and two identical light chains. Such an antibody is preferably expressed from a single vector. However, bispecific antibodies require the use of alternative formats and are commonly expressed on two different vectors. (See, e.g., Spiess et al., 2015, Mol. Immunol. 67:95-106; Brinkmann et al., 2017, MAbs 9:192-212; and Ma et al., 2021, Frontiers in Immunology 12:626616.)

[0150] A first expression vector for a three- or four-chain antibody modality will generally contain a first promoter driving expression of a first nucleotide sequence encoding a first antibody chain, a second promoter driving expression of a second nucleotide sequence encoding a second antibody chain, and a third promoter driving expression of a coding sequence encoding a metabolic selectable marker. A second expression vector for a three- or four-chain antibody modality will generally contain a fourth promoter driving expression of a third nucleotide sequence encoding a third antibody chain, a fifth promoter driving expression of a fifth nucleotide sequence encoding a fourth antibody

chain, and a sixth promoter driving expression of a coding sequence encoding an antibiotic selectable marker.

[0151] Promoters of particular interest for nucleotide sequences encoding antibody chains include the human cytomegalovirus IE1 gene promoter enhancer (CMV) (Boshart et al., 1985, Cell 41:521-30, GenBank Accession No. X03922) and hamster glyceraldehyde-3-phosphate dehydrogenase promoter and intron (GAPDH) (U.S. Patent No. 10,202,261). Additional sequences can also be combined with promoters to improve expression. One such example is the adenovirus tripartite leader (ADL). (See Gingeras et al., 1982, J. Biol. Chem. 257:13475-91, GenBank Accession No. J01917). All of the promoters can be different or any two, three, four, or five of the promoters can be the same.

[0152] In some embodiments, the first, second, third, fourth, fifth, and sixth promoters are the same. In some embodiments, the first, second, third, fourth, fifth, and sixth promoters are CMV-derived promoters (e.g., CMV/GAPDH, CMV/adL, or CMV/EF1a promoters). In some embodiments, the first, second, third, fourth, fifth, and sixth promoters are CMV/GAPDH promoters. In some embodiments, the first, second, third, fourth, fifth, and sixth promoters are CMV/adL promoters. In some embodiments, the first, second, third, fourth, fifth, and sixth promoters are CMV/EF1a promoters.

[0153] In some embodiments, the first, second, third, fourth, fifth, and sixth promoters are glyceraldehyde-3-phosphate dehydrogenase (GAPDH) promoters. In certain embodiments, the GAPDH promoter is operably linked to the CMV promoter enhancer; the resulting construct may be referred to as a CMV/GAPDH promoter. In some embodiments, the first, second, third, fourth, fifth, and sixth promoters are CMV/GAPDH promoters. In some embodiments, each CMV/GAPDH promoter comprises a nucleotide sequence that is 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to the nucleotide sequence of SEQ ID NO: 1. In some embodiments, each CMV/GAPDH promoter comprises the nucleotide sequence of SEQ ID NO: 1.

[0154] A polyA tail can follow each gene (i.e., the coding sequences encoding the first and second antibody chains and the selectable marker). PolyA signal sequences are known in the art and include a bovine growth hormone (BGH) polyA signal sequence (e.g., Pfarr et al., 1986, DNA, 5(2):115-22; Goodwin and Rottman, 1992, J. Biol. Chem., 267(23):16330-16334), a thymidine kinase polyA (TKpA) signal sequence (Cole and Stacy, 1985, Mol Cell Biol., 5(8):2104-13), a rabbit beta-globin polyA signal sequence (Lanoix et al., 1988; EMBO J. 7(8):2515-22; GenBank Accession No. MG356850.1), and a simian virus 40 (SV40) early polyA signal sequence (Connelly and Manley, 1988, Genes Dev., 2(4):440-52; GenBank Accession No. J02400). In some embodiments, the polyA tail can be selected from the group consisting of a rabbit beta-globin pA sequence, a thymidine kinase pA (TKpA) sequence, and a simian virus 40 (SV40) early pA sequence. Additionally, all of the polyA tail sequences can be the different or any two, three, four, or five of the polyA tail sequences can be the same.

[0155] In some embodiments, each polyA sequence is independently selected from a rabbit beta-globin pA sequence, a thymidine kinase pA (TKpA) sequence, and a simian virus 40 (SV40) early pA sequence.

[0156] In some embodiments, each polyA sequence is the same.

[0157] In some embodiments, each polyA sequence is a rabbit beta-globin pA sequence. In some embodiments, the rabbit beta-globin polyA signal sequence comprises a nucleotide sequence that is 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to the nucleotide sequence of SEQ ID NO: 5. In some embodiments, the rabbit beta-globin polyA signal sequence comprises the nucleotide sequence of SEQ ID NO: 5.

[0158] In some embodiments, each polyA sequence is a thymidine kinase pA (TKpA) sequence. In some embodiments, the thymidine kinase polyA (TKpA) signal sequence comprises a nucleotide sequence that is 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to the nucleotide sequence of SEQ ID NO: 7. In some embodiments, the thymidine kinase polyA (TKpA) signal sequence comprises the nucleotide sequence of SEQ ID NO: 7.

[0159] In some embodiments, each polyA sequence is a simian virus 40 (SV40) early pA sequence. the SV40 early polyA signal sequence comprises a nucleotide sequence that is 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to the nucleotide sequence of SEQ ID NO: 6. In some embodiments, the SV40 early polyA signal sequence comprises the nucleotide sequence of SEQ ID NO: 6.

[0160] For example, for certain three-chain modalities, one expression vector contains coding sequences for one heavy chain and one light chain and the other expression vector contains coding sequences for a heavy chain-scFv fusion or a heavy chain-cytokine fusion and a light chain (3 chains, 1 vector). In this example, either expression vector may contain the metabolic selectable marker, while the other expression vector would have the antibiotic selectable marker. The heavy chain sequences can be the same or different. The light chain sequences can be the same or different. An example of three-chain modality is a C1mAb (Fab-heteroFc-[scFv*]), asymmetric fusion. The first antibody chain is each expression cassette can be either the light chain or the heavy chain/heavy chain-scFv fusion. In one aspect, the first antibody chain is the light chain and the second antibody chain is either a heavy chain, a heavy chain-scFv fusion or a heavy chain-cytokine fusion.

[0161] A representative scheme for a C1mAb, asymmetric fusion (Fab-heteroFc-[scFv*]) is depicted in FIG. 2A.

[0162] For example, for a non-limiting example of a four-chain modality, one expression vector contains coding sequences for a first light chain and a first heavy chain and the second expression vector contains a second light chain and a second heavy chain. In this example, both the light chains are different and both the heavy chains are different. This represents a heteroIgG mAb. In this example, either expression vector may contain the metabolic selectable marker, while the other

expression vector would have the antibiotic selectable marker. The first antibody chain in each expression cassette can be either the light chain or the heavy chain.

[0163] A representative scheme for a heteroIgG mAb is depicted in FIG. 3A.

[0164] The expression vectors provided herein provide improved expression, possibly due to improved chain ratios of the expressed polypeptides. Chain ratios can be measured using techniques well-known in the art.

[0165] It has been found that using a dual selection system with one expression vector having a metabolic selection marker and the other expression vector having an antibiotic selection marker for three- and four-chain antibody modalities surprisingly results in higher production of the antibody modality.

[0166] These expression vectors are useful for transformation of a host cell and can contain additional nucleic acid sequences for plasmid maintenance and for cloning and expression of exogenous nucleotide sequences. Such sequences will typically include one or more of the following nucleotide sequences (in addition to the promoter(s), antibody chains, and selectable marker described above): one or more enhancer sequences, an origin of replication, transcriptional and translational control sequences, a transcriptional termination sequence, a complete intron sequence containing a donor and acceptor splice site, various pre- or pro-sequences to improve glycosylation or yield, a native or heterologous signal sequence (leader sequence or signal peptide) for polypeptide secretion, a ribosome binding site, a polyadenylation sequence, internal ribosome entry site (IRES) sequences, an expression augmenting sequence element (EASE), tripartite leader (TPL) and VA gene RNAs from Adenovirus 2, and a polylinker region for inserting the polynucleotide encoding the polypeptide to be expressed. Vectors may be constructed from a starting vector such as a commercially available vector, additional elements may be individually obtained and ligated into the vector. Methods used for obtaining each of the components are well known to one skilled in the art.

[0167] Vector components may be homologous (i.e., from the same species and/or strain as the host cell), heterologous (e.g.., from a species other than the host cell species or strain), hybrid (*i.e.*, a combination of flanking sequences from more than one source), synthetic or native. The sequences of components useful in the vectors may be obtained by methods well known in the art, such as those previously identified by mapping and/or by restriction endonuclease. In addition, they can be obtained by polymerase chain reaction (PCR) and/or by screening a genomic library with suitable probes.

[0168] A ribosome-binding site is usually necessary for translation initiation of mRNA and is characterized by a Shine-Dalgarno sequence (prokaryotes) or a Kozak sequence (eukaryotes). The element is typically located 3′ to the promoter and 5′ to the coding sequence of the polypeptide to be expressed.

[0169] An origin of replication aids in the amplification of the vector in a host cell. They may be included as part of commercially available prokaryotic vectors and may also be chemically

synthesized based on a known sequence and ligated into the vector. Various viral origins (e.g., SV40, polyoma, adenovirus, vesicular stomatitis virus (VSV), or papillomaviruses such as HPV or BPV) are useful for cloning vectors in mammalian cells.

[0170] Transcriptional and translational control sequences for mammalian host cell expression vectors can be excised from viral genomes. Commonly used enhancer sequences are derived from polyoma virus, adenovirus 2, simian virus 40 (SV40), and human cytomegalovirus (CMV). For example, the human CMV promoter/enhancer of immediate early gene 1 may be used. See e.g. Patterson *et al.*, 1994, Applied Microbiol. Biotechnol. 40:691-98. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early and late promoter, enhancer, splice, and polyadenylation sites can be used to provide other genetic elements for expression of a structural gene sequence in a mammalian host cell. Viral early and late promoters are particularly useful because both are easily obtained from a viral genome as a fragment, which can also contain a viral origin of replication (Fiers *et al.*, 1978, Nature 273:113; Kaufman, 1990, Meth. in Enzymol. 185:487-511). Smaller or larger SV40 fragments can also be used, provided the approximately 250 bp sequence extending from the Hind III site toward the BglI site located in the SV40 viral origin of replication site is included. For examples, an enhancer sequence may be inserted into the vector to increase transcription by higher eukaryotes. Enhancers are cis-acting elements of DNA, usually about 10-300 bp in length, which act on the promoter to increase transcription. Enhancers are relatively orientation and position independent, having been found at positions both 5′ and 3′ to the transcription unit. Several enhancer sequences available from mammalian genes are known (e.g., globin, elastase, albumin, alpha-feto-protein and insulin). Typically, however, an enhancer from a virus is used. The SV40 enhancer, the cytomegalovirus early promoter enhancer, the polyoma enhancer, and adenovirus enhancers known in the art are exemplary enhancing elements for the activation of eukaryotic promoters. While an enhancer may be positioned in the vector either 5′ or 3′ to a coding sequence, it is typically located at a site 5′ from the promoter.

[0171] A sequence encoding an appropriate native or heterologous signal sequence (leader sequence or signal peptide) can be incorporated into an expression vector, to promote extracellular secretion of the protein of interest. The choice of signal peptide or leader depends on the type of host cells in which the protein of interest to be produced, and a heterologous signal sequence can replace the native signal sequence. Examples of signal peptides that are functional in mammalian host cells include the following: the signal sequence for interleukin-7 described in U.S. Patent No. 4,965,195; the signal sequence for interleukin-2 receptor described in Cosman *et al.*, 1984, Nature 312:768; the interleukin-4 receptor signal peptide described in EP Patent No. 0367 566; the type I interleukin-1 receptor signal peptide described in U.S. Pat. No. 4,968,607; the type II interleukin-1 receptor signal peptide described in EP Patent No. 0 460 846.

[0172] Additional control sequences shown to improve expression of heterologous genes from mammalian expression vectors include such elements as the expression augmenting sequence element (EASE) derived from CHO cells (Morris *et al.*, in Animal Cell Technology, pp. 529-534 (1997); U.S. Patent Nos. 6,312,951 B1, 6,027,915, and 6,309,841 B1) and the tripartite leader (TPL) and VA gene RNAs from Adenovirus 2 (Gingeras *et al.*, 1982, J. Biol. Chem. 257:13475-13491). The internal ribosome entry site (IRES) sequences of viral origin allows bicistronic mRNAs to be translated efficiently (Oh and Sarnow, 1993, Current Opinion in Genetics and Development 3:295-300; Ramesh *et al.*, 1996, Nucleic Acids Research 24:2697-2700).

[0173] Vectors may be selected to be functional in the particular host cell employed (i.e., the vector is compatible with the host cell machinery, permitting amplification and/or expression of the gene can occur). In some embodiments, vectors are used that employ protein-fragment complementation assays using protein reporters, such as dihydrofolate reductase (see, for example, U.S. Pat. No. 6,270,964). Suitable expression vectors are known in the art and are also commercially available.

[0174] TABLE 1 provides non-limiting example synthetic nucleotide (DNA) sequences for certain expression vector components that may be used in expression vectors of the present disclosure.

TABLE 1. Non-Limiting Example Expression Vector Component Sequences

Description	SEQ ID NO:	Nucleotide Sequence
CMV/GADPH Promoter	1	GTTGACATTGATTATTGACTAGTTATTAATAGTAATCAAT TACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCG CGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACC GCCCAACGACCCCCGCCCATTGACGTCAATAATGACGT ATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGAC GTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTG GCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCT ATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTAT GCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCA GTACATCTACGTATTAGTCATCGCTATTACCATGGCCTCC AGACCACGAGAAGATCCTCAACTTTTCCACAGCCTTTG CATAAAGGGGAGAGGGTCGGCGGTGCAGCTGTGGCAC ACACGCACTTCTGCTCAACCCGCCCCCCCCCGCCCCCG TTCCTGTTCCTTCCCAGGTTCTCCCCATTTTATCGGGGC GGCAACTTTTAGGTCCCTGGGTCCTGGAAGTCCTTAGT ACACACTCTTCGTCCTTAAGTCCATAGTCTGTATTCCCT CGGTCCTATCCTGTCCCCCATCACCGGGTCACCTCCCCA GCGAAGCAATCTCAGTTCCCCTCCCCCTCTCAGCCCCG AGCCCACACGTTTGGTGCGTGCACATTTCAAAAACGAG GCGGGTCCAAAGAGAGGGGGTGGGGAGGTGCCGAGTG GCCCAGCTACTCGCGGCTTTACGGGTGCACGTAGCTCA GGCCTCAGCGCCCTTGAGCTGTGACTGGATGGATGAGC GGGGCGGGAGGCGGGGCGAGCGTCCTCGGCGCTCCCC ACCACCCCAGTTCCTATATAAGCAGAGCTCGTTTAGTGA ACCGTCAGATCCTCACTCTCTTCCGCATCGCTGTCTGCG AGGGCCAGCTGTTGGGCTCGCGGTTGAGGACAAACTC TTCGCGGTCTTTCCAGTACTCTTGGATCGGAAACCCGTC

Description	SEQ ID NO:	Nucleotide Sequence
		GGCCTCCGAACGGTACTCCGCCACCGAGGGACCTGAG CGAGTCCGCATCGACCGGATCGGAAAACCTCTCGAGAA AGGCGTCTAACCAGTCACAGTCGCAAGGTAGGCTGAG CACCGTGGCGGGCGGCAGCGGGTGGCGGTCGGGGTTG TTTCTGGCGGAGGTGCTGCTGATGATGTAATTAAAGTAG GCGCGGATGGTCGAGGTGAGGTGTGGCAGGCTTGAGA TCCAGCTGTTGGGGTGAGTACTCCCTCTCAAAAGCGGG CATTACTTCTGCGCTAAGATTGTCAGTTTCCAAAAACGA GGAGGATTTGATATTCACCTGGCCCGATCTGGCCATACA CTTGAGTGACAATGACATCCACTTTGCCTTTCTCTCCAC AGGTGTCCACTCCCAG
SRα Promoter	2	CTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCC AGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATC TCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGC TCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAA TTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCAT CCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCC CCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGA GGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAG GAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTG GTCGAGGCTCGCATCTCTCCTTCACGCGCCCGCCGCCC TACCTGAGGCCGCCATCCACGCCGGTTGAGTCGCGTTC TGCCGCCTCCCGCCTGTGGTGCCTCCTGAACTGCGTCC GCCGTCTAGGTAAGTTTAAAGCTCAGGTCGAGACCGGG CCTTTGTCCGGCGCTCCCTTGGAGCCTACCTAGACTCA GCCGGCTCTCCACGCTTTGCCTGACCCTGCTTGCTCAA CTCTACGTCTTTGTTTCGTTTTCTGTTCTGCGCCGTTACA GATCCGTCGAGGAACTGAAAAACCAGAAAGTTAACTG GTAAGTTTAGTCTTTTTGTCTTTTATTTCAGGTCCCGGAT CCGGTGGTGGTGCAAATCAAAGAACTGCTCCTCAGTGG ATGTTGCCTTTACTTCTAGGCCTGTACGGAAGTGTTACT TCTGCTCTAAAAGCTGCTGCAACAAGCTTCTA
mPGK Promoter	3	TTCTACCGGGTAGGGGAGGCGCTTTTCCCAAGGCAGTC TGGAGCATGCGCTTTAGCAGCCCCGCTGGGCACTTGGC GCTACACAAGTGGCCTCTGGCCTCGCACACATTCCACA TCCACCGGTAGGCGCCAACCGGCTCCGTTCTTTGGTGG CCCCTTCGCGCCACCTTCTACTCCTCCCCTAGTCAGGAA GTTCCCCCCCGCCCCGCAGCTCGCGTCGTGCAGGACGT GACAAATGGAAGTAGCACCTCTCACTAGTCTCGTGCAG ATGGACAGCACCGCTGAGCAATGGAAGCGGGTAGGCC TTTGGGGCAGCGGCCAATAGCAGCTTTGCTCCTTCGCT TTCTGGGCTCAGAGGCTGGGAAGGGGTGGGTCCGGGG GCGGGCTCAGGGGCGGGCTCAGGGGCGGGGCGGGCGC CCGAAGGTCCTCCGGAGGCCCGGCATTCTGCACGCTTC AAAAGCGCACGTCTGCCGCGCTGTTCTCCTCTTCCTCAT CTCCGGGCCTTTCGA
SV40 Promoter	4	GGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGC TCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAA TTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCC CAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAG TCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCG CCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCAT GGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCC

Description	SEQ ID NO:	Nucleotide Sequence
		GCCTCTGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGG CTTTTTTGGAGGCCTAGCTTTTGCAAAAAGCT
Rabbit Beta-Globin polyA	5	TCTGGCTAATAAAGGAAATTTATTTTCATTGCAATAGTGT GTTGGAATTTTTTGTGTCTCTCACTCGGAAGGACATATG GGAGGGCAAATCATTT
SV40 early polyA	6	AACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGC AATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCA CTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTA TCTTATCATGTCTG
Thymidine Kinase polyA	7	GGGGGAGGCTAACTGAAACACGGAAGGAGACAATACC GGAAGGAACCCGCGCTATGACGGCAATAAAAAGACAG AATAAAACGCACGGGTGTTGGGTCGTTTGTTCATAAAC GCGGGGTTCGGTCCCAGGGCTGGCACTCTGTCGATACC CCACCGAGACCCCATTGGGGCCAATACGCCCGCGTTTC TTCCTTTTCCCCACCCCACCCCCCAAGTTCGGGTGAAG GCCCAGGGCTCGCAGCCAACGTCGGGGCGGCAGGCCC TGCCATAGC

GENERATION OF MAMMALIAN HOST CELLS EXPRESSING A PROTEIN OF INTEREST

[0175] Expression of a protein of interest in a cell can be achieved by well-known methods, either transiently or by stable expression (Davis *et al*., Basic Methods in Molecular Biology, 2nd ed., Appleton & Lange, Norwalk, Conn., 1994; Sambrook *et al*., Molecular Cloning: A Laboratory Manual, 3rd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 2001).

[0176] Methods for stable integration are well-known in the art. Briefly, stable integration is commonly achieved by transiently introducing a heterologous polynucleotide or a vector encoding the heterologous polynucleotide into the host cell, which facilitates the stable integration of said heterologous polynucleotide into the cell genome. Typically, the heterologous polynucleotide is flanked by homology arms, i.e., sequences homologous to the region upstream and downstream to the integration site. Before their introduction into the mammalian host cell, circular vectors may be linearized to facilitate integration into the cell genome. Methods for the introduction of vectors into cells are well-known in the art and include transfection with biological methods, such as viral delivery, with chemical methods, such as using cationic polymers, calcium phosphate, cationic lipids or cationic amino acids; with physical methods, such as electroporation or microinjection; or with mixed approaches, such as protoplast fusion.

[0177] A specific method of stable integration uses recombinase mediated cassette exchange (RMCE; Bode and Baer, 2001, Curr Opin Biotechnol. 12:473-80, and Bode et al., 2000, Biol. Chem. 381:801- 813) for site-specific integration in the genome (also termed “targeted integration”). Site-specific recombinases such as Flp and Cre mediate recombination between two copies of their target sequence termed FRT and loxP, respectively. The use of two incompatible target sequences, for example FRT in combination with F3 (Schlake and Bode, 1994, Biochemistry, 33:12746-51) as well as inverted recognition target sites (Feng et al., 1999, J. Mol. Biol. 292:779-85) allows the insertion of

DNA segments into a predefined chromosomal locus carrying target sequences in a similar configuration. See also EP Patent No. EP1781796B1 and EP Patent Application Publication No. EP2789691A1.

[0178] Insertion of RMCE into a specific site in the genome can be mediated by nucleases (e.g., zinc finger protein (ZFP), transcription activator-like effector nuclease (TALEN), clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 (Cas9)) that can be engineered to create single- and double-stranded breaks (SSBs/DSBs) in the genome. There are two major and distinct pathways to repair DSBs -- homologous recombination and non-homologous end-joining (NHEJ). Homologous recombination requires the presence of a homologous sequence as a template (e.g., "donor" containing RMCE) to guide the cellular repair process and the results of the repair are error-free and predictable. In the absence of a template (or "donor") sequence for homologous recombination, the cell typically attempts to repair the DSB via the unpredictable and error-prone process of non-homologous end-joining (NHEJ).

[0179] A vector may be any molecule or entity (e.g., nucleic acid, plasmid, bacteriophage, transposon, cosmid, chromosome, virus, virus capsid, virion, naked DNA, complexed DNA and the like) suitable for use to transfer and/or transport protein encoding information into a host cell and/or to a specific location and/or compartment within a host cell. Vectors can include viral and non-viral vectors, non-episomal mammalian vectors. Vectors are often referred to as expression vectors, for example, recombinant expression vectors and cloning vectors. The vector may be introduced into a host cell to allow replication of the vector itself and thereby amplify the copies of the polynucleotide contained therein. The cloning vectors may contain sequence components generally include, without limitation, an origin of replication, promoter sequences, transcription initiation sequences, enhancer sequences, inverted terminal repeats, and selectable markers. These elements may be selected as appropriate by a person of ordinary skill in the art.

[0180] Following construction, one or more vectors may be inserted into a suitable cell for amplification and/or polypeptide expression. The transformation of an expression vector into a selected cell may be accomplished by well-known methods including transfection, infection, calcium phosphate co-precipitation, electroporation, nucleofection, microinjection, DEAE-dextran mediated transfection, cationic lipids mediated delivery, liposome mediated transfection, microprojectile bombardment, receptor-mediated gene delivery, delivery mediated by polylysine, histone, chitosan, and peptides. The method selected will in part be a function of the type of host cell to be used. These methods and other suitable methods are well known to the skilled artisan and are set forth in manuals and other technical publications, for example, in Sambrook *et al*., Molecular Cloning: A Laboratory Manual, 3rd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2001).

[0181] The term “transformation” refers to a change in a cell's genetic characteristics, and a cell has been transformed when it has been modified to contain new DNA or RNA. For example, a cell is

transformed where it is genetically modified from its native state by introducing new genetic material via transfection, transduction, or other techniques. Following transfection or transduction, the transforming DNA can recombine with that of the cell by physically integrating into a chromosome of the cell or can be maintained transiently as an episomal element without being replicated, or can replicate independently as a plasmid. A cell is considered to have been "stably transformed" when the transforming DNA is replicated with the division of the cell.

[0182] The term "transfection" refers to the uptake of foreign or exogenous DNA by a cell. A number of transfection techniques are well known in the art and are disclosed herein. See, *e.g.*, Graham *et al.*, 1973, Virology 52:456; Sambrook *et al.*, 2001, Molecular Cloning: A Laboratory Manual, supra; Davis *et al.*, 1986, Basic Methods in Molecular Biology, Elsevier; Chu *et al.*, 1981, Gene 13:197.

[0183] The term "transduction" refers to the process whereby foreign DNA is introduced into a cell via viral vector. See Jones *et al.*, (1998). Genetics: principles and analysis. Boston: Jones & Bartlett Publ.

[0184] Following the methods described above to transfect a mammalian host cell with a pair of expression vectors as described herein, a suitable mammalian host cell can be obtained. Accordingly, the disclosure is directed to a mammalian host cell comprising a pair of expression vectors, wherein a) the first expression vector comprises nucleotide sequences encoding a first antibody light chain, a first antibody heavy chain or antibody heavy chain fusion, and a metabolic selectable marker; and b) the second expression vector comprises nucleotide sequences encoding a second antibody light chain, a second antibody heavy chain or antibody heavy chain fusion, and an antibiotic resistance selectable marker, wherein the second expression vector does not have a metabolic selectable marker.

CELL LINES

[0185] In the methods disclosed herein, any mammalian cell line can be used. A wide variety of mammalian cell lines suitable for growth in culture are available from the American Type Culture Collection (Manassas, Va.) and commercial vendors. Examples of cell lines commonly used in the industry include monkey kidney CVl line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture, (Graham *et al*, 1977, J. Gen Virol. 36:59); baby hamster kidney cells (BHK, ATCC CCL 10); mouse Sertoli cells (TM4, Mather, 1980, Biol. Reprod. 23:243-251); monkey kidney cells (CVl ATCC CCL 70); African green monkey kidney cells (VERO-76, ATCC CRL-1587); human cervical carcinoma cells (HELA, ATCC CCL 2); canine kidney cells (MDCK, ATCC CCL 34); buffalo rat liver cells (BRL 3A, ATCC CRL 1442); human lung cells (W138, ATCC CCL 75); human hepatoma cells (Hep G2, HB 8065); mouse mammary tumor (MMT 060562, ATCC CCL51); TRI cells (Mather *et al.*, 1982, Annals N.Y

Acad. Sci. 383:44-68); MRC 5 cells or FS4 cells; mammalian myeloma cells, and a number of other cell lines and Chinese hamster ovary (CHO) cells.

[0186] Large-scale production of proteins for commercial applications is typically carried out in suspension culture. Therefore, mammalian host cells used to generate the recombinant mammalian cells described herein can, but need not be, adapted to growth in suspension culture. A variety of host cells adapted to growth in suspension culture are known, including mouse myeloma NS0 cells and CHO cells from CHO-S, DG44, and DXB11 cell lines. Other suitable cell lines include mouse myeloma SP2/0 cells, baby hamster kidney BHK-21 cells, human PER.C6® cells, human embryonic kidney HEK-293 cells, and cell lines derived or engineered from any of the cell lines disclosed herein.

[0187] CHO cells are widely used to produce complex recombinant proteins, including CHOK1 cells (ATCC CCL61). The dihydrofolate reductase (DHFR)-deficient mutant cell lines (Urlaub *et al.*, 1980, *Proc Natl Acad Sci USA* 77: 4216-4220), DXB11 and DG-44, are desirable CHO host cell lines because the efficient DHFR selectable and amplifiable gene expression system allows high level recombinant protein expression in these cells (Kaufman R. J., 1990, *Meth Enzymol* 185:537-566). Also included are the glutamine synthase (GS)-knockout CHOK1SV cell lines, making use of glutamine synthetase (GS)-based methionine sulfoximine (MSX) selection. Other suitable CHO host cells could include, but are not limited to the following (ECACC accession numbers in brackets): CHO (85050302), CHO (PROTEIN FREE) (00102307), CHO-K1 (85051005), CHO-K1/SF (93061607), CHO/DHFR-(94060607), CHO/DHFR-AC-free (05011002), RR-CHOKI (92052129).

[0188] In some embodiments of the present disclosure, the mammalian host cell comprising a pair of selection vectors is a CHO cell. In some embodiments, the mammalian host cell is a GS KO CHO cell. In some embodiments, the mammalian host cell is a DHFR- CHO cell.

CELL CULTURE PROCESSES

[0189] The present disclosure also provides methods for producing a recombinant protein (e.g., an antibody modality) comprising a) culturing a mammalian host cell comprising 1) a first expression vector encoding a first light chain, a first heavy chain or a first heavy chain fusion, and a metabolic selectable marker; and 2) a second expression vector encoding a second light chain, a second heavy chain or a second heavy chain fusion, and an antibiotic selectable marker; under conditions in which the antibody chains and selectable markers are expressed, wherein the culturing is in a media comprising a metabolic selection agent and an antibiotic; and b) recovering the recombinant protein (e.g., the antibody modality) from the culture.

[0190] The pair of expression vectors used in such methods (i.e., the first expression vector and the second expression vector) may be any expression vector system described herein. Host cells transfected with the vector systems described herein can be used with adherent culture or suspension cultures grown in stirred tank reactors (including traditional batch and fed-batch cell cultures, which

may but need not comprise a spin filter), perfusion systems (including alternating tangential flow ("ATF") cultures, acoustic perfusion systems, depth filter perfusion systems, and other systems), hollow fiber bioreactors (HFB, which in some cases may be employed in perfusion processes), as well as various other cell culture methods (see, e.g., Tao *et al.*, 2003, *Biotechnol. Bioeng.* 82:751-65; Kuystermans & Al-Rubeai, (2011) "Bioreactor Systems for Producing Antibody from Mammalian Cells" in Antibody Expression and Production, Cell Engineering 7:25-52, Al-Rubeai (ed) Springer; Catapano *et al.*, (2009) "Bioreactor Design and Scale-Up" in Cell and Tissue Reaction Engineering: Principles and Practice, Eibl *et al.* (eds) Springer-Verlag, incorporated herein by reference in their entireties).

[0191] During recombinant protein production, it is desirable to have a controlled system where cells are grown to a desired density and then the physiological state of the cells is switched to a growth-arrested, high productivity state where the cells use energy and substrates to produce the recombinant protein of interest instead of making more cells. Various methods for accomplishing this goal exist, and include temperature shifts and amino acid starvation, as well as use of a cell-cycle inhibitor or other molecule that can arrest cell growth without causing cell death.

[0192] The production of a recombinant protein begins with establishing a mammalian cell production culture of cells that express the protein, in a culture plate, flask, tube, bioreactor, or other suitable vessel. Smaller production bioreactors are typically used, for example, in one embodiment, the bioreactors are 500L to 2000L. In another embodiment, 1000L to 2000L bioreactors are used. The seed cell density used to inoculate the bioreactor can have a positive impact on the level of recombinant protein produced. In some embodiments, the bioreactor can be inoculated with at least 0.5×10^6, such as, e.g., up to and beyond 3.0×10^6 viable cells/mL, in a serum-free culture medium. In another embodiment, the inoculation is 1.0×10^6 viable cells/mL.

[0193] The mammalian cells then undergo an exponential growth phase. The cell culture can be maintained without supplemental feeding until a desired cell density is achieved. The cell culture can be maintained for up to three days with or without supplemental feeding. The culture can be inoculated at a desired cell density to begin the production phase without a brief growth phase. The switch from the growth phase to production phase can be initiated by any of the methods known in the art.

[0194] Three methods are typically used in commercial processes for the production of recombinant proteins by mammalian cell culture: batch culture, fed-batch culture, and perfusion culture. Batch culture is a discontinuous method where cells are grown in a fixed volume of culture media for a short period of time followed by a full harvest. Cultures grown using the batch method experience an increase in cell density until a maximum cell density is reached, followed by a decline in viable cell density as the media components are consumed and levels of metabolic by-products (such as lactate and ammonia) accumulate. Harvest typically occurs at the point when the maximum cell density is

achieved (e.g., $5x10^6$ cells/mL or greater, depending on media formulation, cell line, etc.). The batch process is the simplest culture method; however, viable cell density is limited by the nutrient availability and once the cells are at maximum density, the culture declines and production decreases. There is no ability to extend a production phase because the accumulation of waste products and nutrient depletion rapidly lead to culture decline, typically around 3 to 7 days.

[0195] Fed-batch culture improves on the batch process by providing bolus or continuous media feeds to replenish those media components that have been consumed. Since fed-batch cultures receive additional nutrients throughout the run, they have the potential to achieve higher cell densities (>10 to 30 x 10^6 cells/mL, depending on media formulation, cell line, etc.) and increased product titers, when compared to the batch method. Unlike the batch process, a biphasic culture can be created and sustained by manipulating feeding strategies and media formulations to distinguish the period of cell proliferation to achieve a desired cell density (the growth phase) from the period of suspended or slow cell growth (the production phase). As such, fed batch cultures have the potential to achieve higher product titers compared to batch cultures. Typically, a batch method is used during the growth phase and a fed-batch method is used during the production phase, but a fed-batch feeding strategy can be used throughout the entire process. However, unlike the batch process, bioreactor volume is a limiting factor which limits the amount of feed. Also, as with the batch method, metabolic by-product accumulation will lead to culture decline, which limits the duration of the production phase, for example, to about 10 to 21 days. Fed-batch cultures are discontinuous, and harvest typically occurs when metabolic by-product levels or culture viability reach predetermined levels. When compared to a batch culture, in which no feeding occurs, a fed batch culture can produce greater amounts of recombinant protein. *See e.g.* U.S. Patent No. 5,672,502.

[0196] Perfusion culture is one in which the cell culture receives fresh perfusion feed medium while simultaneously removing spent medium. Perfusion can be continuous, stepwise, intermittent, or a combination of any or all of any of these. Perfusion rates can be less than a working volume to many working volumes per day. The cells are retained in the culture and the spent medium that is removed is substantially free of cells or has significantly fewer cells than the culture. Recombinant proteins expressed by the cell culture can also be retained in the culture. Perfusion can be accomplished by a number of means including centrifugation, sedimentation, or filtration. *See e.g.* Voisard *et al.*, 2003, Biotechnology and Bioengineering 82:751-65. An example of a filtration method is alternating tangential flow filtration. Alternating tangential flow is maintained by pumping medium through hollow-fiber filter modules. *See e.g.* US Patent No. 6,544,424; Furey, 2002, Gen. Eng. News. 22 (7):62-63.

[0197] "Perfusion flow rate" is the amount of media that is passed through (added and removed) from a bioreactor, typically expressed as some portion or multiple of the working volume, in a given time. "Working volume" refers to the amount of bioreactor volume used for cell culture. In one

embodiment, the perfusion flow rate is one working volume or less per day. Perfusion feed medium can be formulated to maximize perfusion nutrient concentration to minimize perfusion rate.

[0198] Cell cultures can be supplemented with concentrated feed medium containing components, such as nutrients and amino acids, which are consumed during the course of the production phase of the cell culture.

[0199] Concentrated feed medium may be based on just about any cell culture media formulation. Such a concentrated feed medium can contain most of the components of the cell culture medium at, for example, about 5X, 6X, 7X, 8X, 9X, 10X, 12X, 14X, 16X, 20X, 30X, 50X, 100X, 200X, 400X, 600X, 800X, or even about 1000X of their normal amount. Concentrated feed media are often used in fed batch culture processes.

[0200] Samples from the cell culture can be monitored and evaluated using any of the analytical techniques known in the art. A variety of parameters including recombinant protein and medium quality and characteristics can be monitored for the duration of the culture. Samples can be taken and monitored intermittently at a desirable frequency, including continuous monitoring, real time or near real time.

[0201] Typically, the cell cultures that precede the final production culture (N-x to N-1) are used to generate the seed cells that will be used to inoculate the production bioreactor, the N-1 culture. The seed cell density can have a positive impact on the level of recombinant protein produced. Product levels tend to increase with increasing seed density. Improvement in titer is tied not only to higher seed density, but is likely to be influenced by the metabolic and cell cycle state of the cells that are placed into production.

[0202] Seed cells can be produced by any culture method. One such method is a perfusion culture using alternating tangential flow filtration. An N-1 bioreactor can be run using alternating tangential flow filtration to provide cells at high density to inoculate a production bioreactor. The N-1 stage may be used to grow cells to densities of >90 x 10^6 cells/mL. The N-1 bioreactor can be used to generate bolus seed cultures or can be used as a rolling seed stock culture that could be maintained to seed multiple production bioreactors at high seed cell density. The duration of the growth stage of production can range from 7 to 14 days and can be designed so as to maintain cells in exponential growth prior to inoculation of the production bioreactor. Perfusion rates, medium formulation and timing are optimized to grow cells and deliver them to the production bioreactor in a state that is most conducive to optimizing their production. Seed cell densities of >15 x 10^6 cells/mL can be achieved for seeding production bioreactors. Higher seed cell densities at inoculation can decrease or even eliminate the time needed to reach a desired production density.

[0203] In certain embodiments, the mammalian host cells can be used to generate a high yield of a protein of interest. High yield, or high volumetric productivity, corresponds to the ability of cells to produce high levels of a protein of interest. The particular yield will depend on the protein of interest

and can be at least 0.05 g/L, at least 0.1 g/L, at least 0.15 g/L, at least 0.2 g/L, at least 0.25 g/L, at least 0.3 g/L, at least 0.35 g/L, at least 0.4 g/L, at least 0.45 g/L, at least 0.5 g/L, at least 0.6 g/L, at least 0.7 g/L, at least 0.8 g/L, at least 0.9 g/L, at least 1 g/L, at least 1.5 g/L, at least 2 g/L, or more, in a 10-day culture grown in fed batch or perfusion conditions, using a feed medium suitable for the mammalian host cell and containing amino acids, vitamins, or trace elements. In specific embodiments, the host cells and methods of the present disclosure express a protein of interest and are capable of producing at least 0.5 g/L, at least 0.6 g/L, at least 0.7 g/L, at least 0.8 g/L, at least 0.9 g/L, at least 1 g/L, at least 1.5 g/L, at least 2 g/L, or more, preferably up to about 3 g/L, 4 g/L, 5 g/L or 10 g/L when grown under the culture conditions described above.

[0204] Yield can also be measured in terms of the specific productivity of a cell line, determined based on the amount of protein produced per cell per day (expressed as pg/cell/day). Mammalian host cells of the present disclosure are capable of producing at least 1 pg/cell/day, at least 2 pg/cell/day, at least 3 pg/cell/day, at least 4 pg/cell/ day, at least 5 pg/cell/day, at least 6 pg/cell/day, at least 7 pg/cell/day, at least 8 pg/cell/day, at least 9 pg/cell/day, at least 10 pg/cell/day, at least 11 pg/cell/day, at least 12 pg/cell/day, at least 13 pg/cell/day, at least 14 pg/cell/day, at least 15 pg/cell/day, at least 20 pg/cell/day, at least 25 pg/cell/day, or more, preferably up to 50 pg/cell/day in a 10-day culture grown in fed batch or perfusion conditions, using a feed medium suitable for the mammalian host cell and containing amino acids, vitamins, or trace elements. In specific embodiments, mammalian host cells of the present disclosure express an protein of interest and have a specific productivity of at least 10 pg/cell/day, at least 11 pg/cell/day, at least 12 pg/cell/day, at least 13 pg/cell/day, at least 14 pg/cell/day, at least 15 pg/cell/day, at least 20 pg/cell/day, at least 25 pg/cell/day, or more, preferably up to 50 pg/cell/day under the culture conditions described above.

[0205] The mammalian host cells described herein can be used to express a protein of interest. The expressed protein may be secreted into the culture medium from which they can be recovered and/or collected. In addition, the proteins can be purified, or partially purified, from such culture or component (e.g., from culture medium) using known processes and products available from commercial vendors. The purified proteins can then be “formulated”, meaning buffer exchanged, sterilized, bulk-packaged, and/or packaged for a final user. Suitable formulations for pharmaceutical compositions include those described in Remington’s Pharmaceutical Sciences, 18thed. 1995, Mack Publishing Company, Easton, Pa.

[0206] In certain embodiments, a CHO DHFR- cell or a CHO GSKO cell can be cultured under conditions to express the antibody chains under methotrexate stringency in the case of a CHO DHFR- cell or methionine sulfoximine stringency in the case of a CHO GSKO cell to favor expression of the difficult-to-express chain paired with the stronger GS promoter.

[0207] In certain embodiments, a CHO DHFR- cell or a CHO GSKO cell can be cultured under conditions to express the antibody chains under methotrexate stringency in the case of a CHO DHFR-

cell or methionine sulfoximine stringency in the case of a CHO GSKO cell to favor expression of the difficult-to-express chain paired with the weaker GS promoter.

[0208] A variety of known techniques can be utilized in making the polynucleotides, polypeptides, vectors, host cells, immune cells, compositions, and the like according to the present disclosure.

PROTEINS OF INTEREST

[0209] Polypeptides and proteins of interest, some of which may be produced using expression vector systems, host cells, and methods described herein, can be of scientific or commercial interest, including protein-based therapeutics. Proteins of interest include, among other things, secreted proteins, non-secreted proteins, intracellular proteins, or membrane-bound proteins. Polypeptides and proteins of interest can be produced by recombinant animal cell lines using cell culture methods and may be referred to as "recombinant proteins". The expressed protein(s) may be produced intracellularly or secreted into the culture medium from which it can be recovered and/or collected. The term "isolated protein" or "isolated recombinant protein" refers to a polypeptide or protein of interest, which is purified away from proteins or polypeptides or other contaminants that would interfere with its therapeutic, diagnostic, prophylactic, research or other use. Proteins of interest include proteins that exert a therapeutic effect by binding a target, particularly a target among those listed below, including targets derived therefrom, targets related thereto, and modifications thereof.

[0210] Proteins of interest include "antigen-binding proteins," including, for example, "antibody modalities." Antigen-binding protein refers to proteins or polypeptides that comprise an antigen-binding region or antigen-binding portion that has affinity for another molecule to which it binds (antigen). Antigen-binding proteins encompass antibodies, peptibodies, antibody fragments, antibody derivatives, antibody analogs, fusion proteins (including single-chain variable fragments (scFvs), double-chain (divalent) scFvs, and IgGscFv (see, e.g., Orcutt *et al.*, 2010, Protein Eng Des2 Sel 23:221-228), hetero-IgG (see, e.g., Liu *et al.*, 2015, J Biol Chem 290:7535-7562), muteins, and XmAb® (Xencor, Inc., Monrovia, CA). Examples of antigen-binding proteins include, but are not limited to, a human antibody, a humanized antibody; a chimeric antibody; a recombinant antibody; a single chain antibody; a diabody; a triabody; a tetrabody; a Fab fragment; a $F(ab')_2$ fragment; an IgD antibody; an IgE antibody; an IgM antibody; an IgG1 antibody; an IgG2 antibody; an IgG3 antibody; or an IgG4 antibody, and fragments thereof. Also included are bispecific T cell engagers (BiTE®) molecules, as well as bispecific T cell engager molecules having extensions, such as half-life extensions, for example HLE BiTE molecules, HeteroIg BITE molecules, and the like.

[0211] As used herein, the term "antigen-binding protein" is used in its broadest sense and means a protein comprising a portion that binds to an antigen or target and, optionally, a scaffold or framework portion that allows the antigen binding portion to adopt a conformation that promotes binding of the antigen-binding protein to the antigen. The antigen-binding protein can comprise, for example, an

alternative protein scaffold or artificial scaffold with grafted CDRs or CDR derivatives. Such scaffolds include, but are not limited to, antibody-derived scaffolds comprising mutations introduced to, for example, stabilize the three-dimensional structure of the antigen-binding protein as well as wholly synthetic scaffolds comprising, for example, a biocompatible polymer. *See*, *e.g.*, Korndorfer *et al.*, 2003, *Proteins: Structure, Function, and Bioinformatics*, 53(1):121-129; Roque *et al.*, 2004, *Biotechnol. Prog.* 20:639-654. In addition, peptide antibody mimetics ("PAMs") can be used, as well as scaffolds based on antibody mimetics utilizing fibronectin components as a scaffold.

[0212] An antigen-binding protein can have, for example, the structure of a naturally occurring immunoglobulin. An "immunoglobulin" is a tetrameric molecule. In a naturally occurring immunoglobulin, each tetramer is composed of two identical pairs of polypeptide chains, each pair having one "light" chain (about 25 kDa) and one "heavy" chain (about 50-70 kDa). The amino-terminal portion of each chain includes a variable region of about 100 to 110 or more amino acids primarily responsible for antigen recognition. The carboxy-terminal portion of each chain defines a constant region primarily responsible for effector function. Human light chains are classified as kappa and lambda light chains. Heavy chains are classified as mu, delta, gamma, alpha, or epsilon, and define the antibody's isotype as IgM, IgD, IgG, IgA, and IgE, respectively.

[0213] Naturally occurring immunoglobulin chains exhibit the same general structure of relatively conserved framework regions (FR) joined by three hypervariable regions, also called complementarity determining regions or CDRs. From N-terminus to C-terminus, both light and heavy chains comprise the domains FR1, CDR1, FR2, CDR2, FR3, CDR3 and FR4. The assignment of amino acids to each domain can be done in accordance with the definitions of Kabat *et al.* in Sequences of Proteins of Immunological Interest, 5th Ed., US Dept. of Health and Human Services, PHS, NIH, NIH Publication no. 91-3242, (1991). As desired, the CDRs can also be redefined according to an alternative nomenclature scheme, such as that of Chothia (see Chothia and Lesk, 1987, *J. Mol. Biol.* 196:901-917; Chothia *et al.*, 1989, *Nature* 342:878-883 or Honegger and Pluckthun, 2001, *J. Mol. Biol.* 309:657-670).

[0214] In the context of the instant disclosure, an antigen-binding protein is said to "specifically bind" or "selectively bind" its target antigen when the dissociation constant (K_D) is $\leq 10^{-8}$ M. The antibody specifically binds antigen with "high affinity" when the K_D is $\leq 5x\ 10^{-9}$ M, and with "very high affinity" when the K_D is $\leq 5x\ 10^{-10}$ M.

[0215] The term "antibody" includes reference to both glycosylated and non-glycosylated immunoglobulins of any isotype or subclass or to an antigen-binding region thereof that competes with the intact antibody for specific binding, unless otherwise specified. Unless otherwise specified, antibodies include human, humanized, chimeric, multi-specific, monoclonal, polyclonal, heteroIgG, bispecific, and oligomers. Antibodies include the lgG1-, lgG2- lgG3- or lgG4-type.

[0216] An antigen-binding protein can have one or more binding sites. If there is more than one binding site, the binding sites can be identical to one another or can be different. For example, a naturally occurring human immunoglobulin typically has two identical binding sites, while a "bispecific" or "bifunctional" antibody has two different binding sites. One standard nomenclature for multi-specific antibody modalities is VERITAS. See Biswas et al., 2023, mAbs 15:1-9.(

[0217] An antigen binding fragment or region include Fab, Fab', F(ab')2, Fv, diabodies, Fd, dAb, maxibodies, single chain antibody molecules, single domain V_HH, complementarity determining region (CDR) fragments, scFv, diabodies, triabodies, tetrabodies and polypeptides that contain at least a portion of an immunoglobulin that is sufficient to confer specific antigen binding to a target polypeptide.

[0218] A Fab fragment is a monovalent fragment having the V_L, V_H, C_L and C_H1 domains; a $F(ab')_2$ fragment is a bivalent fragment having two Fab fragments linked by a disulfide bridge at the hinge region; a Fd fragment has the V_H and C_H1 domains; an Fv fragment has the V_L and V_H domains of a single arm of an antibody; and a dAb fragment has a V_H domain, a V_L domain, or an antigen-binding fragment of a V_H or V_L domain (U.S. Pat. Nos. 6,846,634, 6,696,245, U.S. Patent Application Publication Nos. 2005/0202512, 2004/0202995, 2004/0038291, 2004/0009507, 2003/0039958, Ward *et al.*, 1989, *Nature* 341:544-546).

[0219] A single-chain antibody (scFv) is an antibody in which a V_L and a V_H region are joined via a linker (*e.g.*, a synthetic sequence of amino acid residues) to form a continuous protein chain wherein the linker is long enough to allow the protein chain to fold back on itself and form a monovalent antigen binding site (*see*, *e.g.*, Bird *et al.*, 1988, *Science* 242:423-26 and Huston *et al.*, 1988, *Proc. Natl. Acad. Sci. USA* 85:5879-83), U.S. Patent Nos. 7,741,465, and 6,319,494 as well as Eshhar *et al.*, 1997, Cancer Immunol Immunotherapy 45:131-136. An scFv retains the parent antibody's ability to specifically interact with target antigen.

[0220] Diabodies are bivalent antibodies comprising two polypeptide chains, wherein each polypeptide chain comprises V_H and V_L domains joined by a linker that is too short to allow for pairing between two domains on the same chain, thus allowing each domain to pair with a complementary domain on another polypeptide chain (*see*, *e.g.*, Holliger *et al.*, 1993, *Proc. Natl. Acad. Sci. USA* 90:6444-48; and Poljak *et al.*, 1994, *Structure* 2:1121-23). If the two polypeptide chains of a diabody are identical, then a diabody resulting from their pairing will have two identical antigen binding sites. Polypeptide chains having different sequences can be used to make a diabody with two different antigen binding sites. Similarly, tribodies and tetrabodies are antibodies comprising three and four polypeptide chains, respectively, and forming three and four antigen binding sites, respectively, which can be the same or different.

[0221] For purposes of clarity, and as described herein, it is noted that an antigen-binding protein can, but need not, be of human origin (*e.g.*, a human antibody), and in some cases will comprise a

non-human protein, for example a rat or murine protein, and in other cases an antigen-binding protein can comprise a hybrid of human and non-human proteins (*e.g.*, a humanized antibody).

[0222] A protein of interest can comprise a human antibody. The term "human antibody" includes all antibodies that have one or more variable and constant regions derived from human immunoglobulin sequences. In one embodiment, all of the variable and constant domains are derived from human immunoglobulin sequences (a fully human antibody). Such antibodies can be prepared in a variety of ways, including through the immunization with an antigen of interest of a mouse that is genetically modified to express antibodies derived from human heavy and/or light chain-encoding genes, such as a mouse derived from a Xenomouse®, UltiMab™, or Velocimmune® system, or a rat derived from UniRat®. Phage-based approaches can also be employed.

[0223] Alternatively, a protein of interest can comprise a humanized antibody. A "humanized antibody" has a sequence that differs from the sequence of an antibody derived from a non-human species by one or more amino acid substitutions, deletions, and/or additions, such that the humanized antibody is less likely to induce an immune response, and/or induces a less severe immune response, as compared to the non-human species antibody, when it is administered to a human subject. In one embodiment, certain amino acids in the framework and constant domains of the heavy and/or light chains of the non-human species antibody are mutated to produce the humanized antibody. In another embodiment, the constant domain(s) from a human antibody are fused to the variable domain(s) of a non-human species. Examples of how to make humanized antibodies can be found in U.S. Pat. Nos. 6,054,297, 5,886,152 and 5,877,293.

[0224] Also included are modified proteins, such as are proteins modified chemically by a non-covalent bond, covalent bond, or both a covalent and non-covalent bond. Also included are proteins further comprising one or more post-translational modifications, which may be made by cellular modification systems or modifications introduced *ex vivo* by enzymatic and/or chemical methods or introduced in other ways.

[0225] Proteins of interest may also include recombinant fusion proteins comprising, for example, a multimerization domain, such as a leucine zipper, a coiled coil, an Fc portion of an immunoglobulin, and the like. Also included are proteins comprising all or part of the amino acid sequences of differentiation antigens (referred to as CD proteins) or their ligands or proteins substantially similar to either of these.

[0226] In some embodiments, proteins of interest may include proteins that bind specifically to one or more CD proteins, HER receptor family proteins, cell adhesion molecules, growth factors, nerve growth factors, fibroblast growth factors, transforming growth factors (TGF), insulin-like growth factors, osteoinductive factors, insulin and insulin-related proteins, coagulation and coagulation-related proteins, colony stimulating factors (CSFs), other blood and serum proteins blood group antigens; receptors, receptor-associated proteins, growth hormones, growth hormone receptors, T-cell

receptors; neurotrophic factors, neurotrophins, relaxins, interferons, interleukins, viral antigens, lipoproteins, integrins, rheumatoid factors, immunotoxins, surface membrane proteins, transport proteins, homing receptors, addressins, regulatory proteins, and immunoadhesins.

[0227] In some embodiments, proteins of interest bind to one of more of the following, alone or in any combination: CD proteins including but not limited to CD3, CD4, CD5, CD7, CD8, CD19, CD20, CD22, CD25, CD30, CD33, CD34, CD38, CD40, CD70, CD123, CD133, CD138, CD171, and CD174, HER receptor family proteins, including, for instance, HER2, HER3, HER4, and the EGF receptor, EGFRvIII, cell adhesion molecules, for example, LFA-1, Mol, p150,95, VLA-4, ICAM-1, VCAM, and alpha v/beta 3 integrin, growth factors, including but not limited to, for example, vascular endothelial growth factor ("VEGF"); VEGFR2, growth hormone, thyroid stimulating hormone, follicle stimulating hormone, luteinizing hormone, growth hormone releasing factor, parathyroid hormone, mullerian-inhibiting substance, human macrophage inflammatory protein (MIP-1-alpha), erythropoietin (EPO), nerve growth factor, such as NGF-beta, platelet-derived growth factor (PDGF), fibroblast growth factors, including, for instance, aFGF and bFGF, epidermal growth factor (EGF), Cripto, transforming growth factors (TGF), including, among others, TGF-α and TGF-β, including TGF-β1, TGF-β2, TGF-β3, TGF-β4, or TGF-β5, insulin-like growth factors-I and -II (IGF-I and IGF-II), des(1-3)-IGF-I (brain IGF-I), and osteoinductive factors, insulins and insulin-related proteins, including but not limited to insulin, insulin A-chain, insulin B-chain, proinsulin, and insulin-like growth factor binding proteins; (coagulation and coagulation-related proteins, such as, among others, factor VIII, tissue factor, von Willebrand factor, protein C, alpha-1-antitrypsin, plasminogen activators, such as urokinase and tissue plasminogen activator ("t-PA"), bombazine, thrombin, thrombopoietin, and thrombopoietin receptor, colony stimulating factors (CSFs), including the following, among others, M-CSF, GM-CSF, and G-CSF, other blood and serum proteins, including but not limited to albumin, IgE, and blood group antigens, receptors and receptor-associated proteins, including, for example, flk2/flt3 receptor, obesity (OB) receptor, growth hormone receptors, and T-cell receptors; neurotrophic factors, including but not limited to, bone-derived neurotrophic factor (BDNF) and neurotrophin-3, -4, -5, or -6 (NT-3, NT-4, NT-5, or NT-6); relaxin A-chain, relaxin B-chain, and prorelaxin, interferons, including for example, interferon-alpha, -beta, and -gamma, interleukins (ILs), e.g., IL-1 to IL-10, IL-12, IL-15, IL-17, IL-23, IL-12/IL-23, IL-2Ra, IL1-R1, IL-6 receptor, IL-4 receptor and/or IL-13 to the receptor, IL-13RA2, or IL-17 receptor, IL-1RAP; viral antigens, including but not limited to, an AIDS envelope viral antigen, lipoproteins, calcitonin, glucagon, atrial natriuretic factor, lung surfactant, tumor necrosis factor-alpha and -beta, enkephalinase, BCMA, IgKappa, ROR-1, ERBB2, mesothelin, RANTES (regulated on activation normally T-cell expressed and secreted), mouse gonadotropin-associated peptide, DNase, FR-alpha, inhibin, and activin, integrin, protein A or D, rheumatoid factors, immunotoxins, bone morphogenetic protein (BMP), superoxide dismutase, surface membrane proteins, decay accelerating factor (DAF),

AIDS envelope, transport proteins, homing receptors, MIC (MIC-a, MIC-B), ULBP 1-6, EPCAM, addressins, regulatory proteins, immunoadhesins, antigen-binding proteins, somatropin, CTGF, CTLA4, eotaxin-1, MUC1, CEA, c-MET, Claudin-18, GPC-3, EPHA2, FPA, LMP1, MG7, NY-ESO-1, PSCA, ganglioside GD2, ganglioside GM2, BAFF, OPGL (RANKL), myostatin, Dickkopf-1 (DKK-1), Ang2, NGF, IGF-1 receptor, hepatocyte growth factor (HGF), TRAIL-R2, c-Kit, B7RP-1, PSMA, NKG2D-1, programmed cell death protein 1 and ligand, PD1 and PDL1, mannose receptor/hCGβ, hepatitis-C virus, mesothelin dsFv[PE38] conjugate, Legionella pneumophila (lly), IFN gamma, interferon gamma induced protein 10 (IP10), IFNAR, TALL-1, thymic stromal lymphopoietin (TSLP), proprotein convertase subtilisin/Kexin Type 9 (PCSK9), stem cell factors, Flt-3, calcitonin gene-related peptide (CGRP), OX40L, α4β7, platelet specific (platelet glycoprotein IIb/IIIb (PAC-1), transforming growth factor beta (TFGβ), Zona pellucida sperm-binding protein 3 (ZP-3), TWEAK, platelet derived growth factor receptor alpha (PDGFRα), sclerostin, and biologically active fragments or variants of any of the foregoing.

[0228] Examples of therapeutic proteins include, but are not limited to, abciximab, adalimumab, adecatumumab, aflibercept, alemtuzumab, alirocumab, anakinra, atacicept, basiliximab, belimumab, bevacizumab, biosozumab, blinatumomab, brentuximab vedotin, brodalumab, cantuzumab mertansine, canakinumab, cetuximab, certolizumab pegol, conatumumab, daclizumab, denosumab, eculizumab, edrecolomab, efalizumab, epratuzumab, etanercept, evolocumab, galiximab, ganitumab, gemtuzumab, golimumab, ibritumomab tiuxetan, infliximab, ipilimumab, lerdelimumab, lumiliximab, lxdkizumab, mapatumumab, motesanib diphosphate, muromonab-CD3, natalizumab, nesiritide, nimotuzumab, nivolumab, ocrelizumab, ofatumumab, omalizumab, oprelvekin, palivizumab, panitumumab, pembrolizumab, pertuzumab, pexelizumab, ranibizumab, rilotumumab, rituximab, romiplostim, romosozumab, sargamostim, tocilizumab, tositumomab, trastuzumab, ustekinumab, vedolizumab, visilizumab, volociximab, zanolimumab, zalutumumab, and biosimilars of any of the foregoing.

[0229] Proteins of interest further include antibodies comprising 1, 2, 3, 4, 5, or 6 of the complementarity determining regions (CDRs) of any of the aforementioned antibodies. One or more CDRs can be incorporated into a molecule either covalently or noncovalently to make it an antigen-binding protein. An antigen-binding protein can incorporate the CDR(s) as part of a larger polypeptide chain, can covalently link the CDR(s) to another polypeptide chain, or can incorporate the CDR(s) noncovalently. The CDRs permit the antigen-binding protein to specifically bind to a particular antigen of interest. Also included are variants that comprise a region that is 70% or more, especially 80% or more, more especially 90% or more, yet more especially 95% or more, particularly 97% or more, more particularly 98% or more, yet more particularly 99% or more identical in amino acid sequence to a reference amino acid sequence of a protein of interest. Identity in this regard can be determined using a variety of well-known and readily available amino acid sequence analysis

software. Preferred software includes those that implement the Smith-Waterman algorithms, considered a satisfactory solution to the problem of searching and aligning sequences. Other algorithms also may be employed, particularly where speed is an important consideration. Commonly employed programs for alignment and homology matching of DNAs, RNAs, and polypeptides that can be used in this regard include FASTA, TFASTA, BLASTN, BLASTP, BLASTX, TBLASTN, PROSRCH, BLAZE, and MPSRCH, the latter being an implementation of the Smith-Waterman algorithm for execution on massively parallel processors made by MasPar.

[0230] Proteins of interest comprising an Fc region, including antigen-binding proteins and Fc chain fusion proteins, form another aspect of the instant disclosure. Additionally, a hemibody can be a protein of interest in the context of the instant disclosure.

ADDITIONAL NON-LIMITING EXAMPLE EMBODIMENTS

[0231] Non-limiting example embodiments of the present disclosure also include:

E1. A mammalian host cell comprising a pair of expression vectors, wherein

a) the first expression vector comprises nucleotide sequences encoding a first antibody light chain, a first antibody heavy chain or antibody heavy chain fusion, and a metabolic selectable marker; and

b) the second expression vector comprises nucleotide sequences encoding a second antibody light chain, a second antibody heavy chain or antibody heavy chain fusion, and an antibiotic resistance selectable marker which is expressed in said mammalian host cell.

E2. The mammalian host cell of E1, wherein the metabolic selectable marker is selected from the group consisting of glutamine synthetase and dihydrofolate reductase.

E3. The mammalian host cell of E1 or E2, wherein the antibiotic resistance marker is a resistance marker to an antibiotic selected from the group consisting of puromycin, geneticin, hygromycin, blasticidin, and phleomycin D.

E4. The mammalian host cell of any one of E1-E3, wherein the first antibody light chain and second antibody light chain are the same.

E5. The mammalian host cell of any one of E1-E4, wherein the first expression vector encodes a first antibody heavy chain, the second expression vector encodes a second antibody heavy chain, and the first antibody heavy chain and second antibody heavy chain are different.

E6. The mammalian host cell of any one of E1-E4, wherein the first expression vector encodes an antibody heavy chain fusion and the second expression vector encodes an antibody heavy chain.

E7. The mammalian host cell of any one of E1-E4, wherein the first expression vector encodes an antibody heavy chain and the second expression vector encodes an antibody heavy chain fusion.

E8. The mammalian host cell of any one of E1-E7, wherein the antibody heavy chain fusion is selected from the group consisting of an antibody heavy chain-scFv, an antibody heavy chain-cytokine, and an antibody heavy chain-VHH.

E9. The mammalian host cell of any one of E1-E8, wherein a promoter is operably linked to each of the nucleotide sequences encoding an antibody light chain, an antibody heavy chain or antibody heavy chain fusion, and a metabolic selectable marker.

E10. The mammalian host cell of any one of E1-E9, wherein the promoter for the metabolic selectable marker is selected from the group consisting of mPGK and Sra and the promoter for the antibiotic resistance selectable marker is SV40.

E11. The mammalian host cell of any one of E1-E9, wherein a polyA sequence is operably linked to each of the nucleotide sequences.

E12. The mammalian host cell of E11, wherein the polyA sequences are the same or different and are selected from the group consisting of a rabbit beta-globin pA sequence, a thymidine kinase pA (TKpA) sequence and a simian virus 40 (SV40) early pA sequence.

E13. The mammalian host cell of E1, wherein

a) the first expression vector comprises a nucleotide sequence which comprises the following elements in 5' to 3' order:
    1) a first promoter operably linked to a nucleotide sequence encoding an antibody light chain followed by a first polyA sequence;
    2) a second promoter operably linked to a nucleotide sequence encoding an antibody heavy chain followed by a second polyA sequence; and
    3) a promoter operably linked to a nucleotide sequence encoding a metabolic selectable marker followed by a third polyA sequence; and

b) the second expression vector comprises a nucleotide sequence which comprises the following elements in 5' to 3' order:

1) a first promoter operably linked to a nucleotide sequence encoding an antibody light chain followed by a first polyA sequence;
2) a second promoter operably linked to a nucleotide sequence encoding an antibody heavy chain-scFv fusion followed by a second polyA sequence; and
3) a promoter operably linked to a nucleotide sequence encoding an antibiotic resistance selectable marker followed by a third polyA sequence.

E14. The mammalian host cell of E13, which encodes a C1 mAb.

E15. The mammalian host cell of E1, wherein

a) the first expression vector comprises a nucleotide sequence which comprises the following elements in 5' to 3' order:

1) a first promoter operably linked to a nucleotide sequence encoding an antibody light chain followed by a first polyA sequence;
2) a second promoter operably linked to a nucleotide sequence encoding a first antibody heavy chain followed by a second polyA sequence; and
3) a promoter operably linked to a nucleotide sequence encoding a metabolic selectable marker followed by a third polyA sequence; and

b) the second expression vector comprises a nucleotide sequence which comprises the following elements in 5' to 3' order:

1) a first promoter operably linked to a nucleotide sequence encoding an antibody light chain followed by a first polyA sequence;
2) a second promoter operably linked to a nucleotide sequence encoding a second antibody heavy chain followed by a second polyA sequence; and
3) a promoter operably linked to a nucleotide sequence encoding an antibiotic resistance selectable marker followed by a third polyA sequence,

wherein the first antibody heavy chain and the second antibody heavy chain are different.

E16. The mammalian host cell of E15, which encodes a heteroIgG.

E17. The mammalian host cell of E1, wherein

a) the first expression vector comprises a nucleotide sequence which comprises the following elements in 5' to 3' order:

1) a first promoter operably linked to a nucleotide sequence encoding an antibody light chain followed by a first polyA sequence;
2) a second promoter operably linked to a nucleotide sequence encoding an antibody heavy chain followed by a second polyA sequence; and
3) a promoter operably linked to a nucleotide sequence encoding a metabolic selectable marker followed by a third polyA sequence; and

b) the second expression vector comprises a nucleotide sequence which comprises the following elements in 5' to 3' order:
4) a first promoter operably linked to a nucleotide sequence encoding an antibody light chain followed by a first polyA sequence;
5) a second promoter operably linked to a nucleotide sequence encoding an antibody heavy chain-cytokine fusion followed by a second polyA sequence; and
6) a promoter operably linked to a nucleotide sequence encoding a antibiotic resistance selectable marker followed by a third polyA sequence.

E18. The mammalian host cell of E17, which encodes an antibody cytokine-fusion.

E19. The mammalian host cell of any one of E1-E18, wherein the cell is a Chinese Hamster Ovary (CHO) cell.

E20. The mammalian host cell of E19, wherein the CHO cell is dihydrofolate reductase deficient (dhfr-) or a glutamine synthetase knock out (GSKO).

E21. A method for producing an antibody modality comprising

a) culturing a mammalian host cell comprising
1) a first expression vector encoding a first light chain, a first heavy chain or a first heavy chain fusion and a metabolic selectable marker; and
2) a second expression vector encoding a second light chain, a second heavy chain or a second heavy chain fusion and an antibiotic selectable marker;

under conditions in which the antibody chains and selectable markers are expressed, wherein the culturing is in a media comprising a metabolic selection agent and an antibiotic; and

b) recovering the antibody modality from the culture.

E22. The method of E21, wherein the metabolic selectable marker is glutamine synthetase and the metabolic selection agent is methionine sulfoximine or the metabolic selection marker is dihydrofolate reductase and the metabolic selection agent is methotrexate.

E23. The method of E22, wherein the antibiotic resistance marker is hph gene and the antibiotic is hygromycin B, the antibiotic resistance marker is the pac gene and the antibiotic is puromycin, the antibiotic resistance marker is neo gene and the antibiotic is neomycin, the antibiotic resistance marker is bsr gene and the antibiotic is blasticidin, or the antibiotic resistance marker is the ble gene and the antibiotic is phleomycin D.

E24. The method of E21, wherein

a) the first expression vector comprises a nucleotide sequence which comprises the following elements in 5' to 3' order:
    1) a first promoter operably linked to a nucleotide sequence encoding an antibody light chain followed by a first polyA sequence;
    2) a second promoter operably linked to a nucleotide sequence encoding an antibody heavy chain followed by a second polyA sequence; and
    3) a promoter operably linked to a nucleotide sequence encoding a metabolic selectable marker followed by a third polyA sequence; and
b) the second expression vector comprises a nucleotide sequence which comprises the following elements in 5' to 3' order:
    1) a first promoter operably linked to a nucleotide sequence encoding an antibody light chain followed by a first polyA sequence;
    2) a second promoter operably linked to a nucleotide sequence encoding an antibody heavy chain-scFv fusion followed by a second polyA sequence; and
    3) a promoter operably linked to a nucleotide sequence encoding a antibiotic resistance selectable marker followed by a third polyA sequence.

E25. The method of E24, which encodes a C1 mAb.

E26. The method of E21, wherein

a) the first expression vector comprises a nucleotide sequence which comprises the following elements in 5' to 3' order:
    1) a first promoter operably linked to a nucleotide sequence encoding an antibody light chain followed by a first polyA sequence;

2) a second promoter operably linked to a nucleotide sequence encoding a first antibody heavy chain followed by a second polyA sequence; and
3) a promoter operably linked to a nucleotide sequence encoding a metabolic selectable marker followed by a third polyA sequence; and

b) the second expression vector comprises a nucleotide sequence which comprises the following elements in 5' to 3' order:
1) a first promoter operably linked to a nucleotide sequence encoding an antibody light chain followed by a first polyA sequence;
2) a second promoter operably linked to a nucleotide sequence encoding a second antibody heavy chain followed by a second polyA sequence; and
3) a promoter operably linked to a nucleotide sequence encoding a antibiotic resistance selectable marker followed by a third polyA sequence,

wherein the first antibody heavy chain and second antibody heavy chain are different.

E27. The method of E26, which encodes a heteroIgG.

E28. The method of E21, wherein

a) the first expression vector comprises a nucleotide sequence which comprises the following elements in 5' to 3' order:
1) a first promoter operably linked to a nucleotide sequence encoding an antibody light chain followed by a first polyA sequence;
2) a second promoter operably linked to a nucleotide sequence encoding an antibody heavy chain followed by a second polyA sequence; and
3) a promoter operably linked to a nucleotide sequence encoding a metabolic selectable marker followed by a third polyA sequence; and

b) the second expression vector comprises a nucleotide sequence which comprises the following elements in 5' to 3' order:
1) a first promoter operably linked to a nucleotide sequence encoding an antibody light chain followed by a first polyA sequence;
2) a second promoter operably linked to a nucleotide sequence encoding an antibody heavy chain-cytokine fusion followed by a second polyA sequence; and
3) a promoter operably linked to a nucleotide sequence encoding a antibiotic resistance selectable marker followed by a third polyA sequence.

E29. The method of E28, which encodes an antibody cytokine-fusion.

E30. The method of any one of E21-E29, wherein the mammalian host cell is a Chinese Hamster Ovary (CHO) cell.

E31. The method of E30, wherein the CHO cell is dihydrofolate reductase deficient (dhfr-) or a glutamine synthetase knock out (GSKO).

E32. The method of any one of E21-E31, wherein the recovered antibody modality is purified and formulated in a pharmaceutically acceptable formulation.

[0232] Further non-limiting example embodiments/features of the present disclosure include:

F1. An expression vector system comprising a first expression vector and a second expression vector, wherein:

the first expression vector comprises a first nucleotide sequence encoding a first antibody light chain, a second nucleotide sequence encoding a first antibody heavy chain, antibody heavy chain fusion, or Fc chain fusion, and a third nucleotide sequence encoding a metabolic selectable marker; and

the second expression vector comprises a fourth nucleotide sequence encoding a second antibody light chain, a fifth nucleotide sequence encoding a second antibody heavy chain, antibody heavy chain fusion, or Fc chain fusion, and a sixth nucleotide sequence encoding an antibiotic resistance selectable marker,

wherein, if the second nucleotide sequence or the fifth nucleotide sequence encodes an Fc chain fusion, then the other nucleotide sequence encodes an antibody heavy chain or antibody heavy chain fusion.

F2. The expression vector system of F1, wherein the first expression vector comprises, in 5' to 3' order, the first nucleotide sequence, the second nucleotide sequence, and the third nucleotide sequence.

F3. The expression vector system of F1 or F2, wherein the second expression vector comprises, in 5' to 3' order, the fourth nucleotide sequence, the fifth nucleotide sequence, and the sixth nucleotide sequence.

F4. The expression vector system of any one of F1-F3, wherein:

the first expression vector comprises, in 5' to 3' order, the first nucleotide sequence, the second nucleotide sequence, and the third nucleotide sequence; and

the second expression vector comprises, in 5' to 3' order, the fourth nucleotide sequence, the fifth nucleotide sequence, and the sixth nucleotide sequence.

F5. The expression vector system of any one of F1-F4, wherein the first expression vector does not comprise a nucleotide sequence encoding an antibiotic resistance selectable marker.

F6. The expression vector system of any one of F1-F5, wherein the second expression vector does not comprise a nucleotide sequence encoding a metabolic selectable marker.

F7. The expression vector system of any one of F1-F6, wherein the first expression vector does not comprise a nucleotide sequence encoding an antibiotic resistance selectable marker, and the second expression vector does not comprise a nucleotide sequence encoding a metabolic selectable marker.

F8. The expression vector system of any one of F1-F7, wherein the metabolic selectable marker is glutamine synthetase or dihydrofolate reductase.

F9. The expression vector system of any one of F1-F8, wherein the metabolic selectable marker is glutamine synthetase.

F10. The expression vector system of any one of F1-F8, wherein the metabolic selectable marker is dihydrofolate reductase.

F11. The expression vector system of any one of F1-F10, wherein the antibiotic resistance selectable marker is a resistance marker to an antibiotic selected from the group consisting of puromycin, geneticin, hygromycin, blasticidin, and phleomycin D.

F12. The expression vector system of any one of F1-F11, wherein the antibiotic resistance selectable marker is a resistance marker to puromycin.

F13. The expression vector system of any one of F1-F11, wherein the antibiotic resistance selectable marker is a resistance marker to geneticin.

F14. The expression vector system of any one of F1-F11, wherein the antibiotic resistance selectable marker is a resistance marker to hygromycin.

F15. The expression vector system of any one of F1-F11, wherein the antibiotic resistance selectable marker is a resistance marker to blasticidin.

F16. The expression vector system of any one of F1-F11, wherein the antibiotic resistance selectable marker is a resistance marker to phleomycin D.

F17. The expression vector system of any one of F1-F11, wherein the antibiotic resistance selectable marker is the pac gene, the hph gene, the bsr gene, or the ble gene.

F18. The expression vector system of any one of F1-F11, F12, or F17, wherein the antibiotic resistance selectable marker is the pac gene.

F19. The expression vector system of any one of F1-F11, F14 or F17, wherein the antibiotic resistance selectable marker is the hph gene.

F20. The expression vector system of any one of F1-F11, F15, or F17, wherein the antibiotic resistance selectable marker is the bsr gene.

F21. The expression vector system of any one of F1-F11, F16, or F17, wherein the antibiotic resistance selectable marker is the ble gene.

F22. The expression vector system of any one of F1-F9, F11, or F14, wherein the metabolic selectable marker is glutamine synthetase, and the antibiotic resistance selectable marker is a resistance marker to hygromycin.

F23. The expression vector system of any one of F1-F9, F11, F14, F17, or F19, wherein the metabolic selectable marker is glutamine synthetase, and the antibiotic resistance selectable marker is the hph gene.

F24. The expression vector system of any one of F1-F23, wherein a first promoter is operably linked to the first nucleotide sequence, a second promoter is operably linked to the second nucleotide sequence, and a third promoter is operably linked to the third nucleotide sequence.

F25. The expression vector system of F24, wherein the first expression vector comprises, in 5’ to 3’ order, the first promoter, the first nucleotide sequence, the second promoter, the second nucleotide sequence, the third promoter, and the third nucleotide sequence.

F26. The expression vector system of any one of F1-F25, wherein a fourth promoter is operably linked to the fourth nucleotide sequence, a fifth promoter is operably linked to the fifth nucleotide sequence, and a sixth promoter is operably linked to the sixth nucleotide sequence.

F27. The expression vector system of F26, wherein the second expression vector comprises, in 5' to 3' order, the fourth promoter, the fourth nucleotide sequence, the fifth promoter, the fifth nucleotide sequence, the sixth promoter, and the sixth nucleotide sequence.

F28. The expression vector system of any one of F1-F23, wherein:

a first promoter is operably linked to the first nucleotide sequence, a second promoter is operably linked to the second nucleotide sequence, a third promoter is operably linked to the third nucleotide sequence, a fourth promoter is operably linked to the fourth nucleotide sequence, a fifth promoter is operably linked to the fifth nucleotide sequence, and a sixth promoter is operably linked to the sixth nucleotide sequence;

the first expression vector comprises, in 5' to 3' order, the first promoter, the first nucleotide sequence, the second promoter, the second nucleotide sequence, the third promoter, and the third nucleotide sequence; and

the second expression vector comprises, in 5' to 3' order, the fourth promoter, the fourth nucleotide sequence, the fifth promoter, the fifth nucleotide sequence, the sixth promoter, and the sixth nucleotide sequence.

F29. The expression vector system of any one of F24, F25, or F28, wherein the third promoter is mPGK or SRα.

F30. The expression vector system of any one of F24, F25, F28, or F29, wherein the third promoter is mPGK.

F31. The expression vector system of any one of F24, F25, F28, or F29, wherein the third promoter is SRα.

F32. The expression vector system of any one of F26-F31, wherein the sixth promoter is SV40.

F33. The expression vector system of F26-F29 or F31, wherein the third promoter is mPGK or SRα and the sixth promoter is SV40.

F34. The expression vector system of any one of F1-F23, wherein:

a first promoter is operably linked to the first nucleotide sequence, a second promoter is operably linked to the second nucleotide sequence, a third promoter is operably linked to the third nucleotide sequence, a fourth promoter is operably linked to the fourth nucleotide sequence, a fifth promoter is operably linked to the fifth nucleotide sequence, and a sixth promoter is operably linked to the sixth nucleotide sequence;

the first expression vector comprises, in 5' to 3' order, the first promoter, the first nucleotide sequence, the second promoter, the second nucleotide sequence, the third promoter, and the third nucleotide sequence;

the second expression vector comprises, in 5' to 3' order, the fourth promoter, the fourth nucleotide sequence, the fifth promoter, the fifth nucleotide sequence, the sixth promoter, and the sixth nucleotide sequence;

the third promoter is mPGK or SRα; and

the sixth promoter is SV40.

F35. The expression vector system of any one of F26-F34, wherein each of the first promoter, the second promoter, the fourth promoter, and the fifth promoter is a CMV-derived promoter (e.g., a CMV/GAPDH, CMV/adL, or CMV/EF1a promoter).

F36. The expression vector system of any one of F26-F35, wherein each of the first promoter, the second promoter, the fourth promoter, and the fifth promoter is a CMV/GAPDH or CMV/adL promoter.

F37. The expression vector system of any one of F26-F35, wherein:

a first promoter is operably linked to the first nucleotide sequence, a second promoter is operably linked to the second nucleotide sequence, a third promoter is operably linked to the third nucleotide sequence, a fourth promoter is operably linked to the fourth nucleotide sequence, a fifth promoter is operably linked to the fifth nucleotide sequence, and a sixth promoter is operably linked to the sixth nucleotide sequence;

the first expression vector comprises, in 5' to 3' order, the first promoter, the first nucleotide sequence, the second promoter, the second nucleotide sequence, the third promoter, and the third nucleotide sequence;

the second expression vector comprises, in 5' to 3' order, the fourth promoter, the fourth nucleotide sequence, the fifth promoter, the fifth nucleotide sequence, the sixth promoter, and the sixth nucleotide sequence;

the third promoter is mPGK or SRα;

the sixth promoter is SV40; and

each of the first promoter, the second promoter, the fourth promoter, and the fifth promoter is a CMV-derived promoter (e.g., a CMV/GAPDH, CMV/adL, or CMV/EF1a promoter).

F38. The expression vector system of any one of F26-F37, wherein:

a first promoter is operably linked to the first nucleotide sequence, a second promoter is operably linked to the second nucleotide sequence, a third promoter is operably linked to the third nucleotide sequence, a fourth promoter is operably linked to the fourth nucleotide sequence, a fifth promoter is operably linked to the fifth nucleotide sequence, and a sixth promoter is operably linked to the sixth nucleotide sequence;

the first expression vector comprises, in 5' to 3' order, the first promoter, the first nucleotide sequence, the second promoter, the second nucleotide sequence, the third promoter, and the third nucleotide sequence;

the second expression vector comprises, in 5' to 3' order, the fourth promoter, the fourth nucleotide sequence, the fifth promoter, the fifth nucleotide sequence, the sixth promoter, and the sixth nucleotide sequence;

the third promoter is mPGK or SRα;

the sixth promoter is SV40; and

each of the first promoter, the second promoter, the fourth promoter, and the fifth promoter is a CMV/GAPDH or CMV/adL promoter.

F39. The expression vector system of any one of F1-F38, wherein a first polyA sequence is operably linked to the first nucleotide sequence, a second polyA sequence is operably linked to the second nucleotide sequence, a third polyA sequence is operably linked to the third nucleotide sequence, a fourth polyA sequence is operably linked to the fourth nucleotide sequence, a fifth polyA sequence is operably linked to the fifth nucleotide sequence, and a sixth polyA sequence is operably linked to the sixth nucleotide sequence.

F40. The expression vector system of F39, wherein the first expression vector comprises, in 5' to 3' order, the first promoter, the first nucleotide sequence, the first polyA sequence, the second promoter, the second nucleotide sequence, the second polyA sequence, the third promoter, the third nucleotide sequence, and the third polyA sequence.

F41. The expression vector system of F39 or F40, wherein the second expression vector comprises, in 5' to 3' order, the fourth promoter, the fourth nucleotide sequence, the fourth polyA sequence, the

fifth promoter, the fifth nucleotide sequence, the fifth polyA sequence, the sixth promoter, the sixth nucleotide sequence, and the sixth polyA sequence.

F42. The expression vector system of any one of F39-F41, wherein:

the first expression vector comprises, in 5' to 3' order, the first promoter, the first nucleotide sequence, the first polyA sequence, the second promoter, the second nucleotide sequence, the second polyA sequence, the third promoter, the third nucleotide sequence, and the third polyA sequence; and

the second expression vector comprises, in 5' to 3' order, the fourth promoter, the fourth nucleotide sequence, the fourth polyA sequence, the fifth promoter, the fifth nucleotide sequence, the fifth polyA sequence, the sixth promoter, the sixth nucleotide sequence, and the sixth polyA sequence.

F43. The expression vector system of any one of F39-F42, wherein each of the first polyA sequence, the second polyA sequence, the third polyA sequence, the fourth polyA sequence, the fifth polyA sequence, and the sixth polyA sequence is independently selected from the group consisting of a rabbit beta-globin pA, a thymidine kinase pA (TKpA) sequence, and a simian virus 40 (SV40) early pA sequence.

F44. The expression vector system of any one of F39-F43, wherein each of the first polyA sequence, the second polyA sequence, the third polyA sequence, the fourth polyA sequence, the fifth polyA sequence, and the sixth polyA sequence is a rabbit beta-globin pA sequence.

F45. The expression vector system of any one of F39-F43, wherein each of the first polyA sequence, the second polyA sequence, the third polyA sequence, the fourth polyA sequence, the fifth polyA sequence, and the sixth polyA sequence is a thymidine kinase pA (TKpA) sequence.

F46. The expression vector system of any one of F39-F43, wherein each of the first polyA sequence, the second polyA sequence, the third polyA sequence, the fourth polyA sequence, the fifth polyA sequence, and the sixth polyA sequence is a simian virus 40 (SV40) early pA sequence.

F47. The expression vector system of any one of F1-F23, wherein:

a first promoter and a first polyA sequence are operably linked to the first nucleotide sequence, a second promoter and a second polyA sequence are operably linked to the second nucleotide sequence, a third promoter and a third polyA sequence are operably linked to the third nucleotide sequence, a fourth promoter and a fourth polyA sequence are operably linked to the fourth nucleotide sequence, a fifth promoter and a fifth polyA sequence are operably linked to the fifth

nucleotide sequence, and a sixth promoter and a sixth polyA sequence are operably linked to the sixth nucleotide sequence;

the first expression vector comprises, in 5' to 3' order, the first promoter, the first nucleotide sequence, the first polyA sequence, the second promoter, the second nucleotide sequence, the second polyA sequence, the third promoter, the third nucleotide sequence, and the third polyA sequence;

the second expression vector comprises, in 5' to 3' order, the fourth promoter, the fourth nucleotide sequence, the fourth polyA sequence, the fifth promoter, the fifth nucleotide sequence, the fifth polyA sequence, the sixth promoter, the sixth nucleotide sequence, and the sixth polyA sequence;

the third promoter is mPGK or SRα;

the sixth promoter is SV40;

each of the first promoter, the second promoter, the fourth promoter, and the fifth promoter is a CMV/GAPDH or CMV/adL promoter; and

each of the first polyA sequence, the second polyA sequence, the third polyA sequence, the fourth polyA sequence, the fifth polyA sequence, and the sixth polyA sequence is a simian virus 40 (SV40) early pA sequence.

F48. The expression vector system of any one of F34, F37, F38, or F47, wherein the metabolic selectable marker is glutamine synthetase, and the antibiotic resistance selectable marker is a resistance marker to hygromycin.

F49. The expression vector system of any one of F34, F37, F38, F47, or F48, wherein the metabolic selectable marker is glutamine synthetase, and the antibiotic resistance selectable marker is a resistance marker to hygromycin.

F50. The expression vector system of any one of F1-F49, wherein the first antibody light chain and the second antibody light chain comprise the same amino acid sequence.

F51. The expression vector system of any one of F1-F50, wherein the second nucleotide sequence encodes a first antibody heavy chain or antibody heavy chain fusion, and the fifth nucleotide sequence encodes a second antibody heavy chain or antibody heavy chain fusion.

F52. The expression vector system of any one of F1-F51, wherein the first nucleotide sequence encodes a first antibody heavy chain, the fifth nucleotide sequence encodes a second antibody heavy chain, the first antibody heavy chain and the second antibody heavy chain comprise different amino acid sequences, and the first antibody light chain and the second antibody light chain comprise the same amino acid sequence.

F53. The expression vector system of any one of F1-F51, wherein:

the first nucleotide sequence encodes a first antibody heavy chain fusion and the fifth nucleotide sequence encodes a second antibody heavy chain; or

the first nucleotide sequence encodes a first antibody heavy chain and the fifth nucleotide sequence encodes a second antibody heavy chain fusion.

F54. The expression vector system of any one of F1-F51 or F53, wherein:

(a) (i) the first nucleotide sequence encodes a first antibody heavy chain fusion and the fifth nucleotide sequence encodes a second antibody heavy chain; or (ii) the first nucleotide sequence encodes a first antibody heavy chain and the fifth nucleotide sequence encodes a second antibody heavy chain fusion; and

(b) the first antibody light chain and the second antibody light chain comprise the same amino acid sequence.

F55. The expression vector system of any one of F1-F50, wherein:

the first nucleotide sequence encodes a first antibody heavy chain or antibody heavy chain fusion and the fifth nucleotide sequence encodes a second Fc chain fusion; or

the first nucleotide sequence encodes a first Fc chain fusion and the fifth nucleotide sequence encodes a second antibody heavy chain or antibody heavy chain fusion.

F56. The expression vector system of any one of F1-F50 or F55, wherein:

the first nucleotide sequence encodes a first antibody heavy chain fusion and the fifth nucleotide sequence encodes a second Fc chain fusion; or

the first nucleotide sequence encodes a first Fc chain fusion and the fifth nucleotide sequence encodes a second antibody heavy chain fusion.

F57. The expression vector system of any one of F1-F51 or F53-F56, wherein the first antibody heavy chain fusion and/or the second antibody heavy chain fusion is selected from the group consisting of an antibody heavy chain-scFv, an antibody heavy chain-cytokine, and an antibody heavy chain-VH.

F58. The expression vector system of any one of F1-F50, F55, or F56, wherein the first Fc chain fusion or the second Fc chain fusion is selected from the group consisting of an Fc chain-scFv, an Fc chain-cytokine, and an Fc chain-VH.

F59. A composition comprising an expression vector system of any one of F1-F58.

F60. A mammalian host cell comprising an expression vector system of any one of F1-F58.

F61. The mammalian host cell of F60, wherein the mammalian host cell is a Chinese hamster ovary (CHO) cell.

F62. The mammalian host cell of F60 or F61, wherein the mammalian host cell is a GS KO CHO cell.

F63. A method for producing a recombinant protein comprising:

culturing a mammalian host cell of any one of F60-F62 in a cell culture media adapted for metabolic and antibiotic selection; and

recovering the recombinant protein.

F64. The method of F63, wherein the cell culture media comprises an antibiotic but does not comprise a component necessary for cell survival that is otherwise provided by expression of the metabolic selection marker.

F65. The method of F63 or F64, wherein the cell culture media comprises an antibiotic and a metabolic selection agent.

F66. The method of any one of F63-F65, wherein the cell culture media comprises an antibiotic and a metabolic selection agent but does not comprise a component necessary for cell survival that is otherwise provided by expression of the metabolic selection marker.

F67. The method of F65 or F66, wherein the mammalian host cell comprises an expression vector system in which the metabolic selectable marker is glutamine synthetase and the metabolic selection agent is methionine sulfoximine.

F68. The method of F65 or F66, wherein the mammalian host cell comprises an expression vector system in which the metabolic selectable marker is dihydrofolate reductase and the metabolic selection agent is methotrexate.

F69. The method of F65 or F66, wherein the mammalian host cell comprises an expression vector system in which the metabolic selectable marker is glutamine synthetase and the antibiotic resistance selectable marker is a resistance marker to hygromycin (e.g., the hph gene), the metabolic selection agent is methionine sulfoximine, and the antibiotic is hygromycin B.

F70. The method of any one of F63-F66, wherein the mammalian host cell comprises an expression vector system in which the metabolic selectable marker is glutamine synthetase and the cell culture media does not comprise glutamine.

j

F71. The method of F70, wherein the cell culture media comprises methionine sulfoximine.

F72. The method of F65 or F66, wherein the mammalian host cell comprises an expression vector system in which the metabolic selectable marker is glutamine synthetase and the antibiotic resistance selectable marker is a resistance marker to hygromycin (e.g., the hph gene), the metabolic selection agent is methionine sulfoximine, the antibiotic is hygromycin B, and the cell culture media does not comprise glutamine.

[0233] The present invention is not to be limited in scope by the specific embodiments described herein that are intended as single illustrations of individual aspects of the invention, and functionally equivalent methods and components are within the scope of the invention. Indeed, various modifications of the invention, in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and accompanying drawings. Such modifications are intended to fall within the scope of the appended claims.

EXAMPLES

EXAMPLE 1. Dual Selection Strategy to Improve Productivity for a Three-Chain C1mAb Antibody Modality

Materials and Methods

[0234] *Plasmid Generation.* The coding sequences of the light and heavy chains were inserted into the pPBGS plasmid backbone using Golden Gate cloning to generate two bi-cistronic vectors. Briefly, CMV derived promoters were used to control the LC and HC with SV40-Poly A, followed by the SRα promoter driving the expression of mGS-polyA and/or SV40 driving the expression of Hygromycin B-polyA in that order. All the fragments were unidirectionally assembled using combinations of overhang sequences to facilitate Golden Gate cloning. The various vector configurations are shown in FIG. 2A.

[0235] *Transfection of Plasmids into GS KO Host.* A proprietary GS KO clonal cell host, derived from the CHO-K1 parental host, was used for generating stable pools expressing the 3-chain antibody modality. Host cells were passaged at a seeding density of 0.3-0.4 × 10^6 cells/mL every 3-4 days in a proprietary DMEM-F12-based media in shake flasks at 120 rpm, 36°C and 5% CO_2. Twenty-four

hours before transfection, the host cells were seeded at 1×10^6 cells/mL to ensure the cells would be in exponential growth phase at transfection.

[0236] Stable pools expressing recombinant antibodies were generated using a Gene Pulser XCell (BioRad Laboratories; Hercules, CA) following the manufacturer's protocol. Duplicate transfections were performed for each of the vector configurations. Briefly, 10 mg of each plasmid in combination with 5 mg of a proprietary piggyBac transposase were electroporated into 20×10^6 host cells. The transfected cells were recovered in 20 mL of growth media in 50 mL spin tubes at 225 rpm, 36°C and 5% CO_2.

[0237] *Selection and Recovery.* Seventy-two hours post transfection, the cells were spun down and transferred into selection media without glutamine and with 25 μM of MSX and 400 mg/ml of Hygromycin B. The cells were passaged at seeding densities around $1\text{-}2 \times 10^6$ cells/mL every 3-4 days until viability reached over 90%, when the seeding density was reduced to $0.3\text{-}0.4 \times 10^6$ cells/mL.

[0238] *Fed-Batch Production.* Fully recovered cells were inoculated for fed-batch production at 1×10^6 cells/mL in a proprietary basal media. The cultures were supplemented with proprietary feeds on days 3, 6, and 8 and harvested on day 10. Cell count and viability were determined using a Vi-Cell BLU cell viability analyzer (Beckman Coulter, Brea, CA). Supernatants were analyzed for 1) titer (protein A-HPLC) and 2) product quality attributes including aggregates, clips, and isoforms using size-exclusion chromatography (SEC-UHPLC) (Waters UPLC H-class series), reduced capillary electrophoresis (rCE-SDS) (Sciex PA 800 Plus Pharmaceutical Analysis System), and analytical hydrophobic interaction chromatography (HIC-HPLC) (Agilent HPLC 1100/1200 series), respectively.

Results

[0239] Results in FIG. 2D show a 2.45-fold increase in normalized effective titer using the dual selection approach (black) compared to the optimized vector (grey) configuration that used the standard single metabolic GS selection. In addition, dual selection reduces the impurities measured as nr-CE pre-peaks by 2 times compared to the optimized vector configuration (FIG. 2F). This reduction is due to a much proper chain ratio to achieve perfectly assembled antibodies (FIG. 2G).

EXAMPLE 2. Dual Selection Strategy to Improve Productivity for a heteroIgG Antibody

Materials and Methods

[0240] *heteroIgG mAb Plasmid Generation.* The coding sequences of the light and heavy chains were inserted into the pPBGS plasmid backbone using Golden Gate cloning to generate two bi-cistronic vectors. Briefly, CMV derived promoters were used to control the LC and HC with SV40-Poly A, followed by the SRα or mPGK promoters driving the expression of mGS-polyA and/or SV40 driving the expression of Hygromycin B-polyA in that order. All the fragments were unidirectionally

assembled using combinations of overhang sequences to facilitate Golden Gate cloning. The different vector configurations are shown in FIG. 3A.

[0241] *Transfection of Plasmids into GS KO Host.* A proprietary GS KO clonal cell host, derived from the CHO-K1 parental host, was used for generating stable pools expressing a 3-chain antibody modality. Host cells were passaged at a seeding density of 0.3-0.4 × 10^6 cells/mL every 3-4 days in a proprietary DMEM-F12-based media in shake flasks at 120 rpm, 36°C and 5% CO_2. Twenty-four hours before transfection, the host cells were seeded at 1 × 10^6 cells/mL to ensure the cells would be in exponential growth phase at transfection.

[0242] Stable pools expressing recombinant antibodies were generated using a Gene Pulser XCell (BioRad Laboratories; Hercules, CA) following the manufacturer's protocol. Duplicate transfections were performed for each of the vector configurations. Briefly, 10 mg of each plasmid in combination with 5 mg of a proprietary piggyBac transposase were electroporated into 20 × 10^6 host cells. The transfected cells were recovered in 20 mL of growth media in 50 mL spin tubes at 225 rpm, 36°C and 5% CO_2.

[0243] *Selection and Recovery.* Seventy-two hours post transfection, the cells were spun down and transferred into selection media without glutamine and with 25 μM of MSX and 400 μg/ml of Hygromycin B. The cells were passaged at seeding densities around 1-2 × 10^6 cells/mL every 3-4 days until viability reached over 90%, when the seeding density was reduced to 0.3-0.4 × 10^6 cells/mL.

[0244] *Fed-Batch Production.* Fully recovered cells were inoculated fed-batch production at 1 × 10^6 cells/mL in a proprietary basal media. The cultures were supplemented with proprietary feeds on days 3, 6, and 8 and harvested on day 10. Cell count and viability were determined using a Vi-Cell BLU cell viability analyzer (Beckman Coulter, Brea, CA). Supernatants were analyzed for 1) titer (protein A-HPLC) and 2) product quality attributes including aggregates, clips, and isoforms using size-exclusion chromatography (SEC-UHPLC) (Waters UPLC H-class series), reduced capillary electrophoresis (rCE-SDS) (Sciex PA 800 Plus Pharmaceutical Analysis System), and analytical hydrophobic interaction chromatography (HIC-HPLC) (Agilent HPLC 1100/1200 series), respectively.

Results

[0245] Results in FIGs. 3C and 3D show a 1.2 to 1.4-fold increase in bulk and higher effective titer using the dual selection approach (black) compared to using the standard single GS selection (gray). In addition, dual selection results in comparable %HMW species by SEC and reduces the impurities measured as nr-CE pre-peaks by 3.8 times compared to single selection, respectively (FIGs. 3E and 3F). Bulk and effective titers are normalized to single GS selection samples.

EXAMPLE 3. Dual Selection Strategy to Improve Productivity for an Antibody-Cytokine Fusion

Materials and Methods

[0246] *Antibody-Cytokine Fusion Plasmid Generation.* The coding sequences of the light and heavy chains were inserted into the pPBGS plasmid backbone using Golden Gate cloning to generate two bi-cistronic vectors. Briefly, CMV-derived promoters were used to control the LC and HC with SV40-Poly A, followed by the SRα or mPGK promoters driving the expression of mGS-polyA and/or SV40 driving the expression of Hygromycin B-polyA in that order. All the fragments were unidirectionally assembled using combinations of overhang sequences to facilitate Golden Gate cloning. The different vector configurations are shown in FIG. 4A.

[0247] *Transfection of Plasmids into GS KO Host.* A proprietary GS KO clonal cell host, derived from the CHO-K1 parental host, was used for generating stable pools expressing this 3-chain antibody modality. Host cells were passaged at a seeding density of 0.3-0.4 × 10^6 cells/mL every 3-4 days in a proprietary DMEM-F12-based media in shake flasks at 120 rpm, 36°C and 5% CO_2. Twenty-four hours before transfection, the host cells were seeded at 1 × 10^6 cells/mL to ensure the cells would be in exponential growth phase at transfection.

[0248] Stable pools expressing recombinant antibodies were generated using a Gene Pulser XCell (BioRad Laboratories; Hercules, CA) following the manufacturer's protocol. Duplicate transfections were performed for each of the vector configurations. Briefly, 10 mg of each plasmid in combination with 5 mg of a proprietary piggyBac transposase were electroporated into 20 × 10^6 host cells. The transfected cells were recovered in 20 mL of growth media in 50 mL spin tubes at 225 rpm, 36°C and 5% CO_2.

[0249] *Selection and Recovery.* Seventy-two hours post-transfection, the cells were spun down and transferred into selection media without glutamine and with 25 μM of MSX and 400 μg/ml of Hygromycin B. The cells were passaged at seeding densities around 1-2 × 10^6 cells/mL every 3-4 days until viability reached over 90%, when the seeding density was reduced to 0.3-0.4 × 10^6 cells/mL.

[0250] *Fed-batch Production.* Fully recovered cells were inoculated for fed-batch production at 1 × 10^6 cells/mL in a proprietary basal media. The cultures were supplemented with proprietary feeds on days 3, 6, and 8 and harvested on day 10. Cell count and viability were determined using a Vi-Cell BLU cell viability analyzer (Beckman Coulter, Brea, CA). Supernatants were analyzed for 1) titer (protein A-HPLC) and 2) product quality attributes including aggregates, clips, and isoforms using size-exclusion chromatography (SEC-UHPLC) (Waters UPLC H-class series), reduced capillary electrophoresis (rCE-SDS) (Sciex PA 800 Plus Pharmaceutical Analysis System), and analytical hydrophobic interaction chromatography (HIC-HPLC) (Agilent HPLC 1100/1200 series), respectively.

Results

[0251] Results in FIG. 4D show a 2.2-fold increase in higher normalized effective titer using the dual selection approach (black) compared to the optimized vector (gray) configuration that uses the standard single GS selection. In addition, dual selection reduces the impurities measured as nr-CE pre-peaks by 1.5 times compared to the optimized vector configuration (FIGs. 4E and 4G).

EXAMPLE 4. Dual Selection Strategy to Improve Productivity for a [VH-VH*Fab]-heteroFc Molecule

Materials and Methods

[0252] *Plasmid Generation.* The coding sequences of the light and heavy chains were inserted into a pPBGS plasmid backbone using Golden Gate cloning to generate bicistronic vectors. Briefly, CMV-derived promoters were used to control the light chains and heavy chains with SV40-Poly A, followed by the SRα or mPGK promoters driving the expression of mGS-polyA and/or SV40 driving the expression of Hygromycin B-polyA in that order. All the fragments were unidirectionally assembled using combinations of overhang sequences to facilitate Golden Gate cloning. The different vector configurations are shown in FIG. 5A.

[0253] *Transfection of Plasmids into GS KO Host.* A proprietary GS KO clonal cell host, derived from the CHO-K1 parental host, was used for generating stable pools expressing this [VH-VH*Fab]-heteroFc molecule. Host cells were passaged at a seeding density of 0.5-0.6 × 10^6 cells/mL every 3-4 days in a proprietary DMEM-F12-based media in shake flasks at 120 rpm, 36°C, and 5% CO_2. Twenty-four hours before transfection, the host cells were seeded at 1 × 10^6 cells/mL to ensure the cells would be in exponential growth phase at transfection.

[0254] Stable pools expressing recombinant proteins were generated using a Gene Pulser XCell (BioRad Laboratories; Hercules, CA) following the manufacturer's protocol. Duplicate transfections were performed for each of the vector configurations. Briefly, 10 mg of each plasmid in combination with 5 mg of a proprietary piggyBac transposase were electroporated into 20 × 10^6 host cells. The transfected cells were recovered in 20 mL of growth media in 50 mL spin tubes at 225 rpm, 36°C, and 5% CO_2.

[0255] *Selection and Recovery.* Seventy-two hours post-transfection, the cells were spun down and transferred into selection media without glutamine and with 400 µg/mL of Hygromycin B. The cells were passaged at seeding densities around 1-2 × 10^6 cells/mL every 3-4 days until viability reached over 90%, when the seeding density was reduced to 0.5-0.6 × 10^6 cells/mL.

[0256] *Fed-Batch Production.* Fully recovered cells were inoculated for fed-batch production at 1.3 × 10^6 cells/mL in a proprietary basal media. The cultures were supplemented with proprietary feeds on days 3, 6, 8, 10, and 13 and harvested on day 15. Cell count and viability were determined using a Vi-Cell BLU cell viability analyzer (Beckman Coulter, Brea, CA). Supernatants were analyzed for 1)

titer (protein A-HPLC) and 2) product quality attributes including aggregates, clips, and isoforms using size-exclusion chromatography (SEC-UHPLC) (Waters UPLC H-class series), reduced capillary electrophoresis (rCE-SDS), and non-reduced capillary electrophoresis (nrCE-SDS) (Sciex PA 800 Plus Pharmaceutical Analysis System).

Results

[0257] As shown in FIG. 5B, recovery times for both vector configurations were substantially the same. However, as shown in FIGs. 5C and 5D, use of the dual selection vector increased the bulk and effective titer of Molecule A approximately 1.7-fold and 1.9-fold, respectively, compared to the single selection configuration. Additionally, the dual selection vector configuration reduced the number of impurities, measured as nr-CE pre-peaks, by about 2.5-fold compared to the single selection configuration (FIG. 5F), while maintaining similar aggregates, measured as SEC-HMW (FIG. 5E). rCE-SDS profiles resulting from the two vector configurations are depicted in FIG. 5G.

EXAMPLE 5. Dual Selection Strategy to Improve Productivity for Two [VH*Fab]-heteroFc Molecules

Materials and Methods

[0258] *Plasmid Generation.* The coding sequences of the light and heavy chains for molecules A and B were inserted into a pPBGS plasmid backbone using Golden Gate cloning to generate bicistronic vectors. Briefly, CMV-derived promoters were used to control the light chains and heavy chains with SV40-Poly A, followed by the SRα or mPGK promoters driving the expression of mGS-polyA and/or SV40 driving the expression of Hygromycin B-polyA in that order. All the fragments were unidirectionally assembled using combinations of overhang sequences to facilitate Golden Gate cloning. The different vector configurations are shown in FIG. 6A.

[0259] *Transfection of Plasmids into GS KO Host.* A proprietary GS KO clonal cell host, derived from the CHO-K1 parental host, was used for generating stable pools expressing the two [VH*Fab]-heteroFc molecules (molecules A and B). Host cells were passaged at a seeding density of 0.5-0.6 × 10^6 cells/mL every 3-4 days in a proprietary DMEM-F12-based media in shake flasks at 120 rpm, 36°C, and 5% CO_2. Twenty-four hours before transfection, the host cells were seeded at 1 × 10^6 cells/mL to ensure the cells would be in exponential growth phase at transfection.

[0260] Stable pools expressing recombinant proteins were generated using a Gene Pulser XCell (BioRad Laboratories; Hercules, CA) following the manufacturer's protocol. Duplicate transfections were performed for each of the vector configurations. Briefly, 10 mg of each plasmid in combination with 5 mg of a proprietary piggyBac transposase were electroporated into 20 × 10^6 host cells. The transfected cells were recovered in 20 mL of growth media in 50 mL spin tubes at 225 rpm, 36°C, and 5% CO_2.

[0261] *Selection and Recovery.* Seventy-two hours post-transfection, the cells were spun down and transferred into selection media without glutamine and with 400 μg/mL of Hygromycin B. The cells were passaged at seeding densities around 1-2 × 10^6 cells/mL every 3-4 days until viability reached over 90%, when the seeding density was reduced to 0.5-0.6 × 10^6 cells/mL.

[0262] *Fed-Batch Production.* Fully recovered cells were inoculated for fed-batch production at 1.3 × 10^6 cells/mL in a proprietary basal media. The cultures were supplemented with proprietary feeds on days 3, 6, 8, 10, and 13 and harvested on day 15. Cell count and viability were determined using a Vi-Cell BLU cell viability analyzer (Beckman Coulter, Brea, CA). Supernatants were analyzed for 1) titer (protein A-HPLC) and 2) product quality attributes including aggregates, clips, and isoforms using size-exclusion chromatography (SEC-UHPLC) (Waters UPLC H-class series), reduced capillary electrophoresis (rCE-SDS), and non-reduced capillary electrophoresis (nrCE-SDS) (Sciex PA 800 Plus Pharmaceutical Analysis System).

Results

[0263] As shown in FIG. 6B, the recovery time for both vector configurations was similar for molecule A and slightly longer (by about four days) for molecule B.

[0264] The dual selection vector configuration increased the bulk and effective titer of molecules A and B by approximately 1.5-fold and 1.3-fold, and 1.7-fold and 3.4-fold, respectively, compared to the single selection vector configuration, as shown in FIGs. 6C and 6D. Additionally, the dual selection vector configuration reduced impurities by about 1.4-fold or 1.8-fold, measured by aggregates as SEC-HMW (FIG. 6E) and nr-CE pre-peaks (FIG. 6F), respectively, for molecule B, compared to the single selection vector configuration. Similar impurity levels were observed for molecule A in the dual selection vector and single selection vector configurations (FIGs. 6E and 6F). rCE-SDS profiles resulting from the two vector configurations for molecules A and B are depicted in FIG. 6G.

EXAMPLE 6. Dual Selection Strategy to Improve Productivity for a [Fab*]-heteroFc-[VH*VH] Molecule

Materials and Methods

[0265] *Plasmid Generation.* The coding sequences of the light and heavy chains were inserted into a pPBGS plasmid backbone using Golden Gate cloning to generate bicistronic vectors. Briefly, CMV-derived promoters were used to control the light chains and heavy chains with SV40-Poly A, followed by the SRα or mPGK promoters driving the expression of mGS-polyA and/or SV40 driving the expression of Hygromycin B-polyA in that order. All the fragments were unidirectionally assembled using combinations of overhang sequences to facilitate Golden Gate cloning. The different vector configurations are shown in FIG. 7A.

[0266] *Transfection of Plasmids into GS KO Host.* A proprietary GS KO clonal cell host, derived from the CHO-K1 parental host, was used for generating stable pools expressing the [Fab*]-heteroFc-

[VH*VH] molecule. Host cells were passaged at a seeding density of 0.5-0.6 × 10^6 cells/mL every 3-4 days in a proprietary DMEM-F12-based media in shake flasks at 120 rpm, 36°C, and 5% CO_2. Twenty-four hours before transfection, the host cells were seeded at 1 × 10^6 cells/mL to ensure the cells would be in exponential growth phase at transfection.

[0267] Stable pools expressing recombinant proteins were generated using a Gene Pulser XCell (BioRad Laboratories; Hercules, CA) following the manufacturer's protocol. Duplicate transfections were performed for each of the vector configurations. Briefly, 10 mg of each plasmid in combination with 5 mg of a proprietary piggyBac transposase were electroporated into 20 × 10^6 host cells. The transfected cells were recovered in 20 mL of growth media in 50 mL spin tubes at 225 rpm, 36°C, and 5% CO_2.

[0268] *Selection and Recovery.* Seventy-two hours post-transfection, the cells were spun down and transferred into selection media without glutamine and with 400 µg/mL of Hygromycin B. The cells were passaged at seeding densities around 1-2 × 10^6 cells/mL every 3-4 days until viability reached over 90%, when the seeding density was reduced to 0.5-0.6 × 10^6 cells/mL.

[0269] *Fed-Batch Production.* Fully recovered cells were inoculated for fed-batch production at 1.3 × 10^6 cells/mL in a proprietary basal media. The cultures were supplemented with proprietary feeds on days 3, 6, 8, 10, and 13 and harvested on day 15. Cell count and viability were determined using a Vi-Cell BLU cell viability analyzer (Beckman Coulter, Brea, CA). Supernatants were analyzed for 1) titer (protein A-HPLC) and 2) product quality attributes including aggregates, clips, and isoforms using size-exclusion chromatography (SEC-UHPLC) (Waters UPLC H-class series), reduced capillary electrophoresis (rCE-SDS), and non-reduced capillary electrophoresis (nrCE-SDS) (Sciex PA 800 Plus Pharmaceutical Analysis System).

Results

[0270] As shown in FIG. 7B, the recovery times for the dual selection and single selection vector configurations were similar. Additionally, the double selection vector configuration had a similar bulk and effective titer compared to the single selection vector configuration, as shown in FIGs. 7C and 7D. However, the dual selection vector configuration reduced the number of impurities, measured as SEC-HMW aggregates by about 1.7-fold and nr-CE post-peaks by about 2.2-fold compared to the single selection vector configuration (FIGs. 7E and 7G, respectively), while maintaining similar nr-CE pre-peaks (FIG. 7F). rCE-SDS profiles resulting from the two vector configurations are depicted in FIG. 7H.

What is claimed is:

1. An expression vector system comprising a first expression vector and a second expression vector, wherein:

the first expression vector comprises a first nucleotide sequence encoding a first antibody light chain, a second nucleotide sequence encoding a first antibody heavy chain, antibody heavy chain fusion, or Fc chain fusion, and a third nucleotide sequence encoding a metabolic selectable marker; and

the second expression vector comprises a fourth nucleotide sequence encoding a second antibody light chain, a fifth nucleotide sequence encoding a second antibody heavy chain, antibody heavy chain fusion, or Fc chain fusion, and a sixth nucleotide sequence encoding an antibiotic resistance selectable marker,

wherein, if the second nucleotide sequence or the fifth nucleotide sequence encodes an Fc chain fusion, then the other nucleotide sequence encodes an antibody heavy chain or antibody heavy chain fusion.

2. The expression vector system of claim 1, wherein the first expression vector comprises, in 5' to 3' order, the first nucleotide sequence, the second nucleotide sequence, and the third nucleotide sequence.

3. The expression vector system of claim 1 or claim 2, wherein the second expression vector comprises, in 5' to 3' order, the fourth nucleotide sequence, the fifth nucleotide sequence, and the sixth nucleotide sequence.

4. The expression vector system of any one of claims 1-3, wherein the first expression vector does not comprise a nucleotide sequence encoding an antibiotic resistance selectable marker, and the second expression vector does not comprise a nucleotide sequence encoding a metabolic selectable marker.

5. The expression vector system of any one of claims 1-4, wherein the metabolic selectable marker is glutamine synthetase or dihydrofolate reductase.

6. The expression vector system of any one of claims 1-5, wherein the metabolic selectable marker is glutamine synthetase.

7. The expression vector system of any one of claims 1-6, wherein the antibiotic resistance selectable marker is a resistance marker to an antibiotic selected from the group consisting of puromycin, geneticin, hygromycin, blasticidin, and phleomycin D.

8. The expression vector system of any one of claims 1-7, wherein the antibiotic resistance selectable marker is a resistance marker to hygromycin.

9. The expression vector system of any one of claims 1-8, wherein the metabolic selectable marker is glutamine synthetase, and the antibiotic resistance selectable marker is a resistance marker to hygromycin.

10. The expression vector system of any one of claims 1-9, wherein a first promoter is operably linked to the first nucleotide sequence, a second promoter is operably linked to the second nucleotide sequence, and a third promoter is operably linked to the third nucleotide sequence.

11. The expression vector system of claim 10, wherein the first expression vector comprises, in 5' to 3' order, the first promoter, the first nucleotide sequence, the second promoter, the second nucleotide sequence, the third promoter, and the third nucleotide sequence.

12. The expression vector system of claim 10 or claim 11, wherein the third promoter is mPGK or SRα.

13. The expression vector system of any one of claims 10-12, wherein the third promoter is mPGK.

14. The expression vector system of any one of claims 10-13, wherein the third promoter is SRα.

15. The expression vector system of any one of claims 1-14, wherein a fourth promoter is operably linked to the fourth nucleotide sequence, a fifth promoter is operably linked to the fifth nucleotide sequence, and a sixth promoter is operably linked to the sixth nucleotide sequence.

16. The expression vector system of claim 15, wherein the second expression vector comprises, in 5' to 3' order, the fourth promoter, the fourth nucleotide sequence, the fifth promoter, the fifth nucleotide sequence, the sixth promoter, and the sixth nucleotide sequence.

17. The expression vector system of claim 15 or claim 16, wherein the sixth promoter is SV40.

18. The expression vector system of any one of claims 1-15, wherein:

a first promoter is operably linked to the first nucleotide sequence, a second promoter is operably linked to the second nucleotide sequence, a third promoter is operably linked to the third nucleotide sequence, a fourth promoter is operably linked to the fourth nucleotide sequence, a fifth promoter is operably linked to the fifth nucleotide sequence, and a sixth promoter is operably linked to the sixth nucleotide sequence;

the first expression vector comprises, in 5' to 3' order, the first promoter, the first nucleotide sequence, the second promoter, the second nucleotide sequence, the third promoter, and the third nucleotide sequence;

the second expression vector comprises, in 5' to 3' order, the fourth promoter, the fourth nucleotide sequence, the fifth promoter, the fifth nucleotide sequence, the sixth promoter, and the sixth nucleotide sequence;

the third promoter is mPGK or SRα; and

the sixth promoter is SV40.

19. The expression vector system of any one of claims 1-18, wherein each of the first promoter, the second promoter, the fourth promoter, and the fifth promoter is a CMV-derived promoter.

20. The expression vector system of any one of claims 1-19, wherein each of the first promoter, the second promoter, the fourth promoter, and the fifth promoter is a CMV/GAPDH or CMV/adL promoter.

21. The expression vector system of any one of claims 1-20, wherein:

a first promoter is operably linked to the first nucleotide sequence, a second promoter is operably linked to the second nucleotide sequence, a third promoter is operably linked to the third nucleotide sequence, a fourth promoter is operably linked to the fourth nucleotide sequence, a fifth promoter is operably linked to the fifth nucleotide sequence, and a sixth promoter is operably linked to the sixth nucleotide sequence;

the first expression vector comprises, in 5' to 3' order, the first promoter, the first nucleotide sequence, the second promoter, the second nucleotide sequence, the third promoter, and the third nucleotide sequence;

the second expression vector comprises, in 5' to 3' order, the fourth promoter, the fourth nucleotide sequence, the fifth promoter, the fifth nucleotide sequence, the sixth promoter, and the sixth nucleotide sequence;

the third promoter is mPGK or SRα;

the sixth promoter is SV40; and

each of the first promoter, the second promoter, the fourth promoter, and the fifth promoter is a CMV/GAPDH or CMV/adL promoter.

22. The expression vector system of any one of claims 1-21, wherein a first polyA sequence is operably linked to the first nucleotide sequence, a second polyA sequence is operably linked to the second nucleotide sequence, a third polyA sequence is operably linked to the third nucleotide sequence, a fourth polyA sequence is operably linked to the fourth nucleotide sequence, a fifth polyA sequence is operably linked to the fifth nucleotide sequence, and a sixth polyA sequence is operably linked to the sixth nucleotide sequence.

23. The expression vector system of claim 22, wherein:

the first expression vector comprises, in 5' to 3' order, the first promoter, the first nucleotide sequence, the first polyA sequence, the second promoter, the second nucleotide sequence, the second polyA sequence, the third promoter, the third nucleotide sequence, and the third polyA sequence;

the second expression vector comprises, in 5' to 3' order, the fourth promoter, the fourth nucleotide sequence, the fourth polyA sequence, the fifth promoter, the fifth nucleotide sequence, the fifth polyA sequence, the sixth promoter, the sixth nucleotide sequence, and the sixth polyA sequence.

24. The expression vector system of claim 22 or claim 23, wherein each of the first polyA sequence, the second polyA sequence, the third polyA sequence, the fourth polyA sequence, the fifth polyA sequence, and the sixth polyA sequence is independently selected from the group consisting of a rabbit beta-globin pA sequence, a thymidine kinase pA (TKpA) sequence, and a simian virus 40 (SV40) early pA sequence.

25. The expression vector system of any one of claims 22-24, wherein each of the first polyA sequence, the second polyA sequence, the third polyA sequence, the fourth polyA sequence, the fifth polyA sequence, and the sixth polyA sequence is a simian virus 40 (SV40) early pA sequence.

26. The expression vector system of any one of claims 1-9, wherein:

a first promoter and a first polyA sequence are operably linked to the first nucleotide sequence, a second promoter and a second polyA sequence are operably linked to the second nucleotide sequence, a third promoter and a third polyA sequence are operably linked to the third nucleotide sequence, a fourth promoter and a fourth polyA sequence are operably linked to the fourth nucleotide sequence, a fifth promoter and a fifth polyA sequence are operably linked to the fifth

nucleotide sequence, and a sixth promoter and a sixth polyA sequence are operably linked to the sixth nucleotide sequence;

the first expression vector comprises, in 5' to 3' order, the first promoter, the first nucleotide sequence, the first polyA sequence, the second promoter, the second nucleotide sequence, the second polyA sequence, the third promoter, the third nucleotide sequence, and the third polyA sequence;

the second expression vector comprises, in 5' to 3' order, the fourth promoter, the fourth nucleotide sequence, the fourth polyA sequence, the fifth promoter, the fifth nucleotide sequence, the fifth polyA sequence, the sixth promoter, the sixth nucleotide sequence, and the sixth polyA sequence;

the third promoter is mPGK or SRα;

the sixth promoter is SV40;

each of the first promoter, the second promoter, the fourth promoter, and the fifth promoter is a CMV/GAPDH or CMV/adL promoter; and

each of the first polyA sequence, the second polyA sequence, the third polyA sequence, the fourth polyA sequence, the fifth polyA sequence, and the sixth polyA sequence is a simian virus 40 (SV40) early pA sequence.

27. The expression vector system of any one of claims 18, 21, or 26, wherein the metabolic selectable marker is glutamine synthetase, and the antibiotic resistance selectable marker is a resistance marker to hygromycin.

28. The expression vector system of any one of claims 1-27, wherein the first antibody light chain and the second antibody light chain comprise the same amino acid sequence.

29. The expression vector system of any one of claim 1-28, wherein the second nucleotide sequence encodes a first antibody heavy chain or antibody heavy chain fusion, and the fifth nucleotide sequence encodes a second antibody heavy chain or antibody heavy chain fusion.

30. The expression vector system of any one of claims 1-29, wherein the first nucleotide sequence encodes a first antibody heavy chain, the fifth nucleotide sequence encodes a second antibody heavy chain, the first antibody heavy chain and the second antibody heavy chain comprise different amino acid sequences, and the first antibody light chain and the second antibody light chain comprise the same amino acid sequence.

31. The expression vector system of any one of claims 1-29, wherein:

the first nucleotide sequence encodes a first antibody heavy chain fusion and the fifth nucleotide sequence encodes a second antibody heavy chain; or

the first nucleotide sequence encodes a first antibody heavy chain and the fifth nucleotide sequence encodes a second antibody heavy chain fusion.

32. The expression vector system of any one of claims 1-29, wherein:

(a) (i) the first nucleotide sequence encodes a first antibody heavy chain fusion and the fifth nucleotide sequence encodes a second antibody heavy chain; or (ii) the first nucleotide sequence encodes a first antibody heavy chain and the fifth nucleotide sequence encodes a second antibody heavy chain fusion; and

(b) the first antibody light chain and the second antibody light chain comprise the same amino acid sequence.

33. The expression vector system of any one of claims 1-28, wherein:

the first nucleotide sequence encodes a first antibody heavy chain or antibody heavy chain fusion and the fifth nucleotide sequence encodes a second Fc chain fusion; or

the first nucleotide sequence encodes a first Fc chain fusion and the fifth nucleotide sequence encodes a second antibody heavy chain or antibody heavy chain fusion.

34. The expression vector system of any one of claims 1-8 or 13, wherein:

the first nucleotide sequence encodes a first antibody heavy chain fusion and the fifth nucleotide sequence encodes a second Fc chain fusion; or

the first nucleotide sequence encodes a first Fc chain fusion and the fifth nucleotide sequence encodes a second antibody heavy chain fusion.

35. The expression vector system of any one of claims 1-29 or 31-34, wherein the first antibody heavy chain fusion and/or the second antibody heavy chain fusion is selected from the group consisting of an antibody heavy chain-scFv, an antibody heavy chain-cytokine, and an antibody heavy chain-VH.

36. The expression vector system of any one of claims 1-28 or 33-35, wherein the first Fc chain fusion or the second Fc chain fusion is selected from the group consisting of an Fc chain-scFv, an Fc chain-cytokine, and an Fc chain-VH.

37. A composition comprising an expression vector system of any one of claims 1-36.

38. A mammalian host cell comprising an expression vector system of any one of claims 1-36.

39. The mammalian host cell of claim 38, wherein the mammalian host cell is a Chinese hamster ovary (CHO) cell.

40. A method for producing a recombinant protein comprising:

culturing a mammalian host cell of claim 38 or claim 39 in a cell culture media adapted for metabolic and antibiotic selection; and

recovering the recombinant protein.