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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
12 November 2020 (12.11.2020)

WIPO | PCT

(10) International Publication Number
WO 2020/227206 A1

(51) International Patent Classification:
*C12N 15/52* (2006.01) *C12N 15/67* (2006.01)
*C12N 15/62* (2006.01) *C12N 15/79* (2006.01)

(21) International Application Number:
PCT/US2020/031309

(22) International Filing Date:
04 May 2020 (04.05.2020)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/844,360 07 May 2019 (07.05.2019) US

(71) Applicant: **AMGEN INC.** [US/US]; Law Department - Patent Operations, One Amgen Center, Thousand Oaks, California 91320 (US).

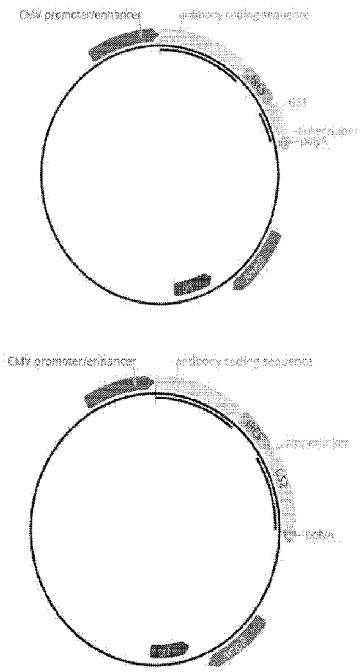
(72) Inventors: **GHATTYVENKATAKRISHNA, Pavan**; c/o Amgen, One Amgen Center Drive, Thousand Oaks, California 91320-1799 (US). **BARKHORDARIAN, Hedieh**; c/o Amgen Inc., One Amgen Center Drive, Thousand Oaks, California 91320-1799 (US). **DARIS, Kristine Marie**; c/o Amgen Inc., One Amgen Center Drive, Thousand Oaks, California 91320-1799 (US). **TEJAMO, Charilyn**; c/o Amgen Inc., One Amgen Center Drive, Thousand Oaks, California 91320-1799 (US).

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

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

(54) Title: VECTORS AND EXPRESSION SYSTEMS FOR PRODUCING RECOMBINANT PROTEINS

Figure 2




(57) Abstract: Inventions disclosed herein relates to vectors and expression systems for producing heteromeric recombinant proteins such as monoclonal antibodies. Vectors and expression systems disclosed herein are based on the finding that the selectable marker glutamine synthetase can be divided into two fragments at selected amino acid positions of the glutamine synthetase polypeptide, and the two fragments can internet and/or associate to form a monomer and then a functional multimeric glutamine synthetase protein. Z

WO 2020/227206 A1

*[Continued on next page]*

**Declarations under Rule 4.17:**

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

**Published:**

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

# VECTORS AND EXPRESSION SYSTEMS FOR PRODUCING RECOMBINANT PROTEINS

## FIELD OF THE INVENTION

**[0001]** Inventions disclosed herein generally relate to the field of vectors and expression systems for producing recombinant proteins. Specifically, inventions disclosed herein relate to vectors, expression systems and methods for producing heteromeric protein complexes comprising different polypeptides.

## BACKGROUND OF THE INVENTION

**[0002]** Therapeutic proteins such as antibodies are an important class of medicines serving patients. Typically, therapeutic proteins are produced in recombinant cells that have been adapted for long term growth in culture. Many times, multiple heterologous recombinant polypeptides (e.g., heavy chain and light chain of an antibody) are expressed in these cells that can form heteromeric complexes.

**[0003]** Production of such heteromeric complexes requires expression systems that concurrently express different polypeptides in appropriate amounts to allow proper association and assembling of the polypeptides to form heteromeric complexes. For example, in the expression of an antibody, the heavy chain and light chain of the antibody need to be expressed in roughly equal amounts for proper association of the heavy and light chains and the production of the antibody. However, a difficulty that may be encountered in antibody production is that either the heavy chain or the light chain is expressed to relatively high levels with respect to the corresponding partner, leading to improper and/or inefficient production of the antibody.

**[0004]** Different approaches have been used to address the difficulty. For instance, an expression system has been developed taking the advantage of reassociation of two fragments of the dihydrofolate reductase (DHFR) selectable maker to form an active molecule. See Bianchi A. and McGrew J., Biotechnol. Bioeng., 84(4): 439-444 (2003). In that system, the expression of each antibody chain (heavy chain and light chain) is linked to the expression of a DHFR fragment, and survival in selective media that require expression of both DHFR fragments leads to the expression of roughly equal amounts of both chains.

**[0005]** Glutamine synthetase has been used as an amplifiable, selectable marker for high level expression of recombinant proteins. See e.g., Bebbington C. R. et al., BioTechnology, 10:169-175 (1992). Glutamine synthetase is a multimeric protein responsible for the biosynthesis of glutamine by catalyzing the condensation of ammonia and glutamate. When a gene encoding a functional glutamine synthetase is introduced in a cell lacking endogenous glutamine synthetase (e.g., a glutamine synthetase knock out mammalian cell), the cell can grow in a glutamine-free medium. There have been efforts in developing expression systems that utilize intragenic complementation of glutamine synthetase, wherein each antibody chain is linked to each of two mutant glutamine synthetases that can complement each other to form a functional glutamine synthetase such that roughly equal amounts of each chain are expressed. Intragenic complementation is a phenomenon that occurs when a multimeric protein is formed from subunits produced by different mutant alleles of a gene (e.g., mutations that mapped to the 5' end of the glutamine synthetase gene could complement those in the 3' end). See e.g., Mitchell, A. P. Genetics 111, 243-258 (1985).

**[0006]** Despite these progresses, there is a continued need for vectors, expression systems and methods that are robust and express heteromeric recombinant proteins to high levels.

## SUMMARY OF THE INVENTION

**[0007]** Disclosed herein are vectors, expression systems and methods for producing heteromeric complexes comprising different polypeptides. Vectors and expression systems disclosed herein are based on the finding that the selectable marker glutamine synthetase can be divided into two fragments at selected amino acid positions of the glutamine synthetase polypeptide, and the two fragments can interact and/or associate to form a monomer and then a functional multimeric glutamine synthetase protein. Specifically, in the vectors and expression systems disclosed herein, the expression of each polypeptide of a heteromeric complex (e.g., heavy and light chain of an antibody) is linked to the expression of a glutamine synthetase fragment, and the expression of the heteromeric complex is accomplished by growing a recombinant cell comprising the vector or the expression system under conditions that require the expression of a functional glutamine synthetase. These vectors and expression systems are robust and express proteins to high levels. In addition, they reduce the time required to select for cells expressing high levels of proteins.

**[0008]** In certain embodiments, disclosed herein is a vector comprising a) a first nucleic acid encoding a first polypeptide, b) a second nucleic acid encoding a first fragment of glutamine synthetase, wherein the transcription of the first nucleic acid is operably linked to the transcription of the second nucleic acid, c) a third nucleic acid encoding a third polypeptide, the third polypeptide is capable of associating with the first polypeptide to form a heteromeric complex, and d) a fourth nucleic acid encoding a second fragment of glutamine synthetase, wherein the transcription of the third nucleic acid is operably linked to the transcription of the fourth nucleic acid, and wherein the first fragment and the second fragment of glutamine synthetase associate to provide a selectable activity, and wherein the vector is capable of being transfected into mammalian cells and improving selection of the transfected cells.

**[0009]** In certain embodiments, disclosed herein is an expression system comprising: a) a first vector comprising a first nucleic acid encoding a first polypeptide, wherein the transcription of the first nucleic acid is operably linked to the transcription of a second nucleic acid encoding a first fragment of glutamine synthetase, and b) a second vector comprising a third nucleic acid encoding a third polypeptide, wherein the transcription of the third nucleic acid is operably linked to the transcription of a fourth nucleic acid encoding a second fragment of glutamine synthetase, wherein the first polypeptide is capable of associating with the third polypeptide to form a heteromeric complex, wherein the first and second fragments of glutamine synthetase associate to provide a selectable activity, and wherein the expression system is capable of being transfected into mammalian cells and improving selection of the transfected cells.

**[0010]** In certain embodiments, the first fragment of glutamine synthetase is an N-terminal fragment of glutamine synthetase and the second fragment of glutamine synthetase is a C-terminal fragment of glutamine synthetase, or the first fragment of glutamine synthetase is a C-terminal fragment of glutamine synthetase and the second fragment of glutamine synthetase is an N-terminal fragment of glutamine synthetase.

**[0011]** In certain embodiments of the vector or the expression system disclosed herein, the glutamine synthetase comprising the amino acid sequence of SEQ ID NO: 1, and the first and second fragments of glutamine synthetase are generated by splitting the glutamine synthetase polypeptide at an amino acid position selected from K52, E55, D92, G187, G245, R262, K291, G302 and D311 of SEQ ID NO: 1. In certain embodiments, the first and second fragments of glutamine synthetase are generated by splitting the glutamine synthetase polypeptide at an amino acid position selected from K52, E55, D92, G187, and G245 of SEQ

ID NO: 1, and in certain embodiments, the first and second fragments of glutamine synthetase are generated by splitting the glutamine synthetase polypeptide at the amino acid position D92 or G187 of SEQ ID NO: 1.

**[0012]** In certain embodiments of the vector or the expression system disclosed herein, a) the first glutamine synthetase fragment comprises amino acid residues 1 to 51 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 52 to 373 of SEQ ID NO: 1; or b) the first glutamine synthetase fragment comprises amino acid residues 1 to 52 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 53 to 373 of SEQ ID NO: 1; or c) the first glutamine synthetase fragment comprises amino acid residues 1 to 54 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 55 to 373 of SEQ ID NO: 1; or d) the first glutamine synthetase fragment comprises amino acid residues 1 to 55 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 56 to 373 of SEQ ID NO: 1; or e) the first glutamine synthetase fragment comprises amino acid residues 1 to 91 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 92 to 373 of SEQ ID NO: 1; or f) the first glutamine synthetase fragment comprises amino acid residues 1 to 92 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 93 to 373 of SEQ ID NO: 1; or g) the first glutamine synthetase fragment comprises amino acid residues 1 to 186 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 187 to 373 of SEQ ID NO: 1; or h) the first glutamine synthetase fragment comprises amino acid residues 1 to 187 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 188 to 373 of SEQ ID NO: 1; or i) the first glutamine synthetase fragment comprises amino acid residues 1 to 244 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 245 to 373 of SEQ ID NO: 1; or j) the first glutamine synthetase fragment comprises amino acid residues 1 to 245 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 246 to 373 of SEQ ID NO: 1; or k) the first glutamine synthetase fragment comprises amino acid residues 1 to 261 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 262 to 373 of SEQ ID NO: 1; or l) the first glutamine synthetase fragment comprises amino acid residues 1 to 262 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 263 to 373 of SEQ ID NO: 1; or m) the first glutamine synthetase fragment comprises amino acid residues 1 to 301 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 302 to 373 of SEQ ID NO: 1; or

n) the first glutamine synthetase fragment comprises amino acid residues 1 to 302 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 303 to 373 of SEQ ID NO: 1; or o) the first glutamine synthetase fragment comprises amino acid residues 1 to 310 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 311 to 373 of SEQ ID NO: 1; or p) the first glutamine synthetase fragment comprises amino acid residues 1 to 311 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 312 to 373 of SEQ ID NO: 1.

**[0013]** In certain embodiments, the glutamine synthetase is a mammalian glutamine synthetase having an amino acid sequence different from SEQ ID NO:1, and the first and second fragments of glutamine synthetase are generated by splitting the mammalian glutamine synthetase polypeptide at an amino acid position equivalent to an amino acid position selected from K52, E55, D92, G187, G245, R262, K291, G302 and D311 of SEQ ID NO: 1 according to sequence alignment; in certain embodiments, the first and second fragments of glutamine synthetase are generated by splitting the mammalian glutamine synthetase polypeptide at an amino acid position equivalent to an amino acid position selected from K52, E55, D92, G187, and G245 of SEQ ID NO:1 according to sequence alignment; in certain embodiments, the first and second fragments of glutamine synthetase are generated by splitting the mammalian glutamine synthetase polypeptide at an amino acid position equivalent to the amino acid position D92 or G187 of SEQ ID NO:1 according to sequence alignment.

**[0014]** In certain embodiments, the heteromeric complex is an antibody or an antigen binding molecule. In certain embodiments, a) the first polypeptide is a heavy chain of an antibody or a fragment thereof and the third polypeptide is a light chain of an antibody or a fragment thereof; or b) the first polypeptide is a heavy chain of an antibody or a fragment thereof and the third polypeptide is a light chain of an antibody or a fragment thereof. In certain embodiments, a) the first nucleic acid encodes an antibody heavy chain or a fragment thereof and the third nucleic acid encodes an antibody light chain or a fragment thereof; or b) the first nucleic acid encodes an antibody light chain or a fragment thereof and the third nucleic acid encodes an antibody heavy chain or a fragment thereof.

**[0015]** In certain embodiments, the vector disclosed herein, or one or both vectors of the expression system disclosed herein further comprises an internal ribosomal entry site (IRES) and/or an expression augmenting sequence element (EASE). In certain embodiments, the IRES occurs at a site selected from: a) a site between the first nucleic acid and the second nucleic acid; b) a site between the third nucleic acid and the fourth nucleic acid, and c) at sites

between both first and second, and third and fourth nucleic acids. In certain embodiments, the IRES comprises the sequence of GATGATAATACCCTCGAGATCCGTGCCATCATG.

**[0016]** In certain embodiments, each of the first and second fragment of glutamine synthetase further comprises an interaction domain. In certain embodiments, the interaction domain is a leucine zipper or an antiparallel leucine zipper polypeptide. In certain embodiments, the interaction domain is a leucine zipper polypeptide of GCN4, C/EBP, c-Fos, c-Jun, c-Myc or c-Max. In certain embodiments, the interaction domain is a leucine zipper polypeptide having the following sequence: ALKKELQANKKELAQLKWELQALKKELAQ EQLEKKLQALEKKLAQLEWKNQALEKKLAQ. In certain embodiments, each of the first and second interaction domains comprise the amino acid sequence of ALKKELQANKKELAQLKWELQALKKELAQ or EQLEKKLQALEKKLAQLEWKNQALEKKLAQ.

**[0017]** In certain embodiments, each of the first and second interaction domains further comprises a linker linking the first or second interaction domain to an interaction domain. In certain embodiments, the linker comprises a sequence selected from GGPGG, GPGGG, GGGGSGGGGGS, GGGGS and GGGGSGGGGS.

**[0018]** In certain embodiments, a) the first interaction domain is fused to the N-terminal of the first glutamine synthetase fragment and has the amino acid sequence of EQLEKKLQALEKKLAQLEWKNQALEKKLAQGGGGSGGGGS and the second interaction domain is fused to the C-terminal of the second glutamine synthetase fragment and has the amino acid sequence of GGGGSGGGGSALKKELQANKKELAQLKWELQALKKELAQ; or b) the first interaction domain is fused to the C-terminal of the first glutamine synthetase fragment and has the amino acid sequence of GGGGSGGGGSALKKELQANKKELAQLKWELQALKKELAQ and the second interaction domain is fused to the N-terminal of the second glutamine synthetase fragment and has the amino acid sequence of EQLEKKLQALEKKLAQLEWKNQALEKKLAQGGGGSGGGGS.

**[0019]** In certain embodiments, the vector disclosed herein, or one or both vectors of the expression system disclosed herein further comprises a fifth nucleic acid encoding a selectable marker selected from Zeomycin, neomycin, puromycin, Blasticidin S, and GPT.

**[0020]** In certain embodiments, disclosed herein is a host cell comprising the vector or the expression system disclosed herein. In certain embodiments, the host cell is CHO, VERO,

BHK, HeLa, Cos, MDCK, 293, 3T3, WI338, or NSO cells. In certain embodiments, the host cell is CHO cells. In certain embodiments, the host cell lacks endogenous glutamine synthetase.

**[0021]** In certain embodiments, disclosed herein is a method of producing an antibody heavy chain or a fragment thereof and an antibody light chain or a fragment thereof comprising culturing the host cell under conditions wherein the heteromeric complex is expressed by the host cell. In certain embodiments, the method further comprises isolating the heteromeric complex.

## BRIEF DESCRIPTION OF THE DRAWINGS

**[0022]** Figure 1 shows sequence alignment of glutamine synthetase from difference species using Geneious Aligner.

**[0023]** Figure 2 shows the plasmid map of the vectors used in the Examples.

**[0024]** Figure 3 shows cell recovery data of various control cell lines (CS9, pGS) and cells transfected with plasmids comprising glutamine synthetase fragments generated by splitting the polypeptide at sites 1-55, 1-92, 1-187, and 1-244 of SEQ ID NO:1 in both linear (L1, L2) and circular format (C1, C2). Cells carrying split site 1-187 recovered in both the linear and circular format while the site 1-92 recovered in the linear format.

**[0025]** Figure 4 shows cells recovery data of cells carrying split site 1-186 in the linear format.

## DETAILED DESCRIPTION

**[0026]** Disclosed herein are vectors, expression systems and methods for producing heteromeric complexes (e.g., monoclonal antibodies). The invention disclosed herein is advantageous in that they are robust, express each component of a heteromeric complex in proportion and to high levels. In addition, it reduces the time required to select for cells expressing high levels of proteins.

**[0027]** The invention disclosed herein utilizes that the selectable marker glutamine synthetase can be divided into two fragments at selected amino acid positions of the glutamine synthetase polypeptide, and the two fragments interact and form a functional glutamine synthetase protein when expressed together, thereby providing a selectable activity. The individual fragment does not have significant selectable activity when expressed alone, but they provide selectable activity when co-expressed. Polynucleotide encoding each glutamine

synthetase fragment is operationally linked to a polynucleotide encoding a polypeptide that forms a heteromeric complex, and the expression of a functional glutamine synthetase in a selectable environment that requires the enzyme leads to the expression of the heteromeric complex.

**[0028]** Two fragments of glutamine synthetase generated at a splitting site may be expressed from one vector. Alternatively, two fragments of glutamine synthetase generated at a splitting site may be expressed from two vectors. Thus in some embodiments, disclosed herein is a vector comprising: a) a first nucleic acid encoding a first polypeptide, b) a second nucleic acid encoding a first fragment of glutamine synthetase, wherein the transcription of the first nucleic acid is operably linked to the transcription of the second nucleic acid, c) a third nucleic acid encoding a third polypeptide, the third polypeptide is capable of associating with the first polypeptide to form a heteromeric complex, and d) a fourth nucleic acid encoding a second fragment of glutamine synthetase, wherein the transcription of the third nucleic acid is operably linked to the transcription of the fourth nucleic acid, wherein the first fragment and the second fragment of glutamine synthetase associate to provide a selectable activity, and wherein the vector is capable of being transfected into mammalian cells and improving selection of the transfected cells.

**[0029]** In some embodiments, disclosed herein is an expression system comprising: a) a first vector comprising a first nucleic acid encoding a first polypeptide, wherein the transcription of the first nucleic acid is operably linked to the transcription of a second nucleic acid encoding a first fragment of glutamine synthetase, and b) a second vector comprising a third nucleic acid encoding a third polypeptide, wherein the transcription of the third nucleic acid is operably linked to the transcription of a fourth nucleic acid encoding a second fragment of glutamine synthetase, wherein the first polypeptide is capable of associating with the third polypeptide to form a heteromeric complex, the first and second fragments of glutamine synthetase associate to provide a selectable activity, and wherein the expression system is capable of being transfected into mammalian cells and improving selection of the transfected cells.

**[0030]** As used herein, the term "vector" is understood as expression vectors, which are DNA sequences that are required for transcription and translation of their DNAs in a eukaryotic host cell (e.g., a mammalian cell) after transfection with vector. An appropriately constructed vector usually comprises at least one expressible marker selectable in eukaryotic cells (e.g., mammalian cells) and restriction sites for insertion of the expression cassette for the

recombinant product gene under control of an upstream promoter region. Optionally, the vector can further comprise an internal ribosomal entry site (IRES) to facilitate translation. The vector may further comprise an origin of replication such as origin of Epstein Barr Virus (EBV) or SV40 virus for autonomous replication/episomal maintenance in eukaryotic host cells. Additional components may be added to facilitate replication in prokaryotic and/or eukaryotic cells, integration of the vector into a eukaryotic chromosome, and markers to aid in selection of and/or screening for cells containing the vector. Vectors include linear DNA fragments, DNA fragments encompassing nuclear targeting sequences or are specially optimized for interaction with transfection reagents, viruses, plasmids, phages, phagemids, cosmids, viruses, retroviruses and the like that can be shuttled and produced in bacteria.

**[0031]** 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.

**[0032]** As used herein, the term "operably linked" refers to when one nucleic acid is placed into a functional relationship with another nucleic acid. More specifically, operably linked includes that two different nucleic acids encoding different polypeptides have transcription induced simultaneously. Operably linked is also intended to mean that the linked nucleic acids can be contiguous in a single transcriptional unit, while translation is directed from one or more ribosomal start sites (e.g., internal ribosomal start site).

**[0033]** As used herein, the term "heteromeric complex" is understood to include a molecular complex formed by the association of at least two different molecules. The association can be non-covalent interaction or covalent attachment, e.g., disulfide bonds. The

two different molecules are typically two different polypeptides; however, the invention also contemplates heteromeric complexes between polypeptides and nucleic acids and between different nucleic acids. In some embodiments, the heteromeric complex provides a functional activity, such as, the ability to bind a substrate (e.g., an immunoglobulin capable of binding a corresponding antigen), enzymatic activity or the like. In some embodiments, the heteromeric complex is secreted into the culture medium of the host cell in which it is being produced.

**[0034]** In some embodiments, the heteromeric complex is an immunoglobulin molecule. The immunoglobulin in vertebrate systems is an antibody comprised of two identical light chains and two identical heavy chains. Each heavy and light chain has a variable region and a constant region. The four chains are joined together by disulfide bonds, such that each light chain is joined with a heavy chain and the heavy chains are connected across their tails forming a Y-shaped heteromeric complex. 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. The heavy chain constant domain comprises three constant domains (CH1, CH2 and CH3) and a hinge region. 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. Antibodies are well known in the art and may be from any origin, including human, non-human or a hybrid of both (e.g., human antibodies, humanized antibodies, chimeric antibodies, and antibodies from species other than human).

**[0035]** In some embodiments, the heteromeric complex is an antigen binding fragment of an antibody. Non-limiting examples of antigen binding fragments include Fab, Fab', $(Fab')_2$, and Fv. A Fab fragment is a monovalent fragment having the light chain variable domain (VL), heavy chain variable domain (VH), light chain constant domain (CL) and the first constant domain of the heavy chain (CH1); a F(ab')2 fragment is a bivalent fragment having two Fab' fragments linked by a disulfide bridge at the hinge region, and the F(ab')2 fragment can be split into two Fab' fragments by mild reduction; an Fv fragment has the VL and VH domains of a single arm of an antibody. Antigen binding fragments of an antibody are well known and used in the art. Numerous techniques are known by which DNA encoding immunoglobulin molecules can be manipulated to yield DNAs capable of encoding recombinant proteins such as antibodies with enhanced affinity, or antigen binding fragment thereof. See, for example, Larrick et al. (1989), Biotechnology 7:934-938; Reichmann et al. (1988), Nature 332:323-327;

Roberts et al. (1987), Nature 328:731-734; Verhoeyen et al. (1988), Science 239:1534-1536; Chaudhary et al. (1989), Nature 339:394-397.

**[0036]** In the vector or the expression system disclosed herein, the first nucleic acid encodes a first polypeptide, and the third nucleic acid encodes a third polypeptide, and the first and the third polypeptides are capable of associating with each other to form a heteromeric complex. In embodiments wherein the heteromeric complex is an antibody or an antigen binding fragment thereof, the heavy chain or a fragment thereof may be encoded by either the first or the third nucleic acid, and the light chain or a fragment thereof may be encoded by either the first or the third nucleic acid. Thus, in some embodiments, the first polypeptide is an immunoglobulin heavy chain or a fragment thereof and the third polypeptide is an immunoglobulin light chain or a fragment thereof. In some embodiments, the first polypeptide is an immunoglobulin light chain or a fragment thereof and the third polypeptide is an immunoglobulin heavy chain or a fragment thereof. When expressed using the vector or the expression system disclosed herein, the light chain or a fragment thereof is fused in frame to the first or the second glutamine synthetase fragment; while the heavy chain or a fragment thereof is fused in frame to the second or the first glutamine synthetase fragment.

**[0037]** In some embodiments, the heteromeric complex is a heterodimeric protein, e.g., a heterodimeric protein comprising two different polypeptides or hetero-oligomeric protein. In some embodiments, the heterodimeric protein is a bispecific antigen binding molecule. As used herein, the term "bispecific antigen binding molecule" is understood to include molecules that recognize two different epitopes either on the same antigen or on different antigens. Many bispecific antigen binding molecules known in the art are generated using recombinant DNA technology. See e.g., Holliger P. and Hudson P. J., Nature Biotech., 23(9): 1126-1136 (2005); Brinkmann U. and Kontermann R. E., MABS, 9(2): 182–212 (2017); Bird R, et al. Science, 242:423-6 (1988); Hudson P, Kortt A., J Immunol Methods. 231:177-89 (1999); Holliger P, et al., Proc Natl Acad Sci U S A. 90:6444-8 (1993). Any bispecific antigen binding molecule can be made using the invention disclosed herein so long as the bispecific antigen binding molecule is comprised of two different polypeptides. In certain bispecific antigen binding molecules, the two different polypeptides may be linked covalently, e.g., by a short peptide linker, or non-covalently. Non-limiting examples of bispecific antigen binding molecules include bispecific scFv (diabody) molecules, bispecific sc(Fab)$_2$ molecules, bispecific Fab fusion molecules, bispecific scFv-Fc molecules, bispecific Fab-dsFv molecules, bispecific Fab-VHH molecules (see e.g., Brinkmann U. and Kontermann R. E., MABS, 9(2): 182–212 (2017), the description

of the various bispecific antigen binding molecules is incorporated herein by reference), and bispecific X-body molecules (see WO2017/134140).

**[0038]** Additional non-limiting examples of heterodimeric or hetero-oligomeric proteins include BMP2/BMP7, osteogenic protein, interleukin 1 converting enzyme (ICE), various interleukin receptors (e.g., the IL-18 receptor, IL-13 receptor, IL-4 receptor and IL-7 receptor), receptors of the nucleus such as retinoid receptors, T-cell receptors, integrins such as cell adhesion molecules, betal-integrins, tumor necrosis factor receptor and soluble and membrane bound forms of class I and class II major histocompatibility complex proteins (MHC). For heteromeric complexes that are receptors, the invention encompasses both soluble and membrane bound forms of the polypeptides. Descriptions of additional heteromeric proteins that can be produced according to the invention can be found in, for example, Human Cytokines: Handbook for Basic and Clinical Research, Vol. II (Aggarwal and Gutterman, eds. Blackwell Sciences, Cambridge Mass., 1998); Growth Factors: A Practical Approach (McKay and Leigh, Eds. Oxford University Press Inc., New York, 1993) and The Cytokine Handbook (A W Thompson, ed.; Academic Press, San Diego Calif.; 1991).

**[0039]** In embodiments wherein the heteromeric complex is a heterodimeric protein (e.g., bispecific scFv, bispecific sc(Fab)$_2$ or interleukin receptors), each polypeptide that forms the heterodimeric protein may be encoded by either the first or the third nucleic acid, and each is fused in frame to a first or a second glutamine synthetase fragment when expressed using the vector or the expression system.

**[0040]** Not wishing to be bound by any theory, it is believed that by using two glutamine synthetase fragments generated at a split site, the vector and expression system disclosed herein provide a more stringent selection condition and a robust expression of recombinant proteins. In addition to expressing heteromeric complexes, the vector and expression system disclosed herein are also useful for expressing proteins comprised of identical polypeptide, e.g., proteins comprised of homodimers and proteins comprised of the same polypeptide. Such proteins are known and used in the art, including, e.g., fusion proteins such as etanercept, aflibercept, epoetin alfa, darbepoetin alfa, filgrastim, pegfilgrastim and BiTE® molecules (e.g., disclosed in WO2008/119567 and WO2017/134140, the structure and sequence of which are incorporated by reference). Thus, in certain embodiments, the first and third nucleic acid are identical, and/or the first and third nucleic acid encode the same polypeptide, and/or the first and third polypeptides are identical. In such embodiments, a protein is expressed when the first and third polypeptide are expressed by the vector or the

expression system, or for a homodimeric protein, the protein is expressed when the first and third polypeptide are expressed by the vector or expression system and associate to form a homodimer.

**[0041]** Glutamine synthetase is an enzyme that catalyzes the synthesis of glutamine from glutamate and ammonia. It is a multimeric protein that can be composed of 8, 10, or 12 identical subunits stacked into two face-to-face rings. Glutamine synthetase from mouse and several other eukaryotic species were analyzed carefully to identify amino acid positions at which to split the polypeptide into a first and a second glutamine synthetase fragment that can interact/associate to form a functional multimeric glutamine synthetase protein when the two fragments are co-expressed. In the vector or the expression system disclosed herein, the first fragment of glutamine synthetase can be either an N-terminal portion or a C-terminal portion of glutamine synthetase. Similarly, the second fragment of glutamine synthetase can be either an N-terminal portion or a C-terminal portion of glutamine synthetase. In some embodiments, the first fragment of glutamine synthetase is an N-terminal fragment of glutamine synthetase and the second fragment of glutamine synthetase is a C-terminal fragment of glutamine synthetase. In some embodiments, the first fragment of glutamine synthetase is a C-terminal fragment of glutamine synthetase and the second fragment of glutamine synthetase is an N-terminal fragment of glutamine synthetase.

**[0042]** Any glutamine synthetase may be used in the invention disclosed herein so long as that, when co-expressed in a host cell (e.g., a mammalian cell), the first and second fragments of the glutamine synthetase interact/associate to form a monomer and then a functional enzyme to provide a selectable activity. In some embodiments, the glutamine synthetase is a mammalian glutamine synthetase. In some embodiment, the glutamine synthetase is a murine glutamine synthetase. In some embodiments, the glutamine synthetase is a mouse glutamine synthetase. In some embodiments, the glutamine synthetase is a non-mouse, mammalian glutamine synthetase. In some embodiments, the glutamine synthetase is a non-mouse, murine glutamine synthetase. In some embodiments, the glutamine synthetase comprises the amino acid sequence of SEQ ID NO: 1. In some embodiments, the glutamine synthetase is a mammalian glutamine synthetase having an amino acid sequence different from SEQ ID NO:1.

1 MATSASSHLNKGIKQMYMSLPQGEKVQAMYIWVDGTGEGLRCKTRTLDCE

51 PKCVEELPEWNFDGSSTFQSEGSNSDMYLHPVAMFRDPFRKDPNKLVLCE

101 VFKYNRKP AETNLRHICKRIMDMVSNQHPWFGMEQEYTLMGTDGHPFGWP

151 SNGFPGPQGPYYCGVGADKAYGRDIVEAHYRACLYAGVKITGTNAEVMPA
201 QWEFQIGPCEGIRMGDHLWI ARFILHRVCEDFGVI ATFDPKPIPGNWNGA
251 GCHTNFSTKAMREENGLKCIEEAIDKLSKRHQYHIRAYDPKGGLDNARRL
301 TGFHETSNINDFSAGVANRGASIRIPRTVGQEKKGYFEDRRPSANCDPYA
351 VTEAIVRTCLLNETGDEPFQYKN* (SEQ ID NO:1)

**[0043]** In some embodiments, the glutamine synthetase comprises the amino acid sequence of SEQ ID NO:1. In some embodiments, the first and second fragments of glutamine synthetase are generated by splitting the glutamine synthetase polypeptide at an amino acid position selected from K52, E55, D92, G187, G245, R262, K291, G302 or D311 of SEQ ID NO:1. In some embodiments, the first and second fragments of glutamine synthetase are generated by splitting the glutamine synthetase polypeptide at an amino acid position selected from K52, E55, D92, G187, or G245 of SEQ ID NO:1. In some embodiments, the first and second fragments of glutamine synthetase are generated by splitting the glutamine synthetase at the amino acid position D92 or G187 of SEQ ID NO:1.

**[0044]** In some embodiments, the glutamine synthetase is a non-mouse, mammalian glutamine synthetase or a mammalian glutamine synthetase having an amino acid sequence that is not identical to SEQ ID NO:1. Non-limiting examples of such glutamine synthetase include rat glutamine synthetase, hamster glutamine synthetase, canine glutamine synthetase, and human glutamine synthetase. Mammalian glutamine synthetases are conserved proteins, however, proteins from different species may contain some variations in amino acid sequence and/or number of amino acids. See e.g., Figure 1. When a mammalian glutamine synthetase having an amino acid sequence that is not identical to SEQ ID NO:1 is used, in some embodiments, the first and second fragments of glutamine synthetase are generated by splitting the mammalian glutamine synthetase polypeptide having an amino acid sequence different from SEQ ID NO:1 at an amino acid position equivalent to an amino acid position selected from K52, E55, D92, G187, G245, R262, K291, G302 and D311 of SEQ ID NO: 1 according to sequence alignment; in some embodiments, the first and second fragments of glutamine synthetase are generated by splitting the mammalian glutamine synthetase polypeptide having an amino acid sequence different from SEQ ID NO:1 at an amino acid position equivalent to an amino acid position selected from K52, E55, D92, G187, and G245 of SEQ ID NO: 1 according to sequence alignment; in some embodiments, the first and second fragments of glutamine synthetase are generated by splitting the mammalian glutamine synthetase polypeptide having an amino acid sequence different from SEQ ID NO:1 at an amino acid

position equivalent to the amino acid position D92 or G187 of SEQ ID NO: 1 according to sequence alignment.

**[0045]** Whether an amino acid position of a glutamine synthetase from a given species or whether an amino acid position of a glutamine synthetase having an amino acid sequence different from SEQ ID NO:1 is equivalent to a specific amino acid position of SEQ ID NO:1 can be determined by aligning that glutamine synthetase amino acid sequence and SEQ ID NO:1. When aligning two sequences, typically one sequence serves as a reference sequence, to which a test sequence is compared. When comparing an amino acid sequence of a particular glutamine synthetase amino acid sequence and SEQ ID NO:1, SEQ ID NO:1 may be the reference sequence while a glutamine synthetase having an amino acid sequence different from SEQ ID NO:1 is the test sequence.

**[0046]** Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison can be conducted, e.g., by manual alignment and visual inspection, see, e.g., Current Protocols in Molecular Biology (Ausubel et al., eds. 1995 supplement); by the local homology algorithm of Smith & Waterman, Adv. Appl. Math. 2:482 (1981), by the homology alignment algorithm of Needleman & Wunsch, J. Mol. Biol. 48:443 (1970), by the search for similarity method of Pearson & Lipman, Proc. Nat'l. Acad. Sci. USA 85:2444 (1988), or by computerized implementations of these algorithms (e.g., GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, Wis.), ALIGN, or by ALIGN-2 (Genentech, South San Francisco, Calif.), Megalign (DNASTAR), or the Geneious Aligner or ClustalW (Available from Biomatters, www.geneious.com/), Clustal Omega or T-Coffee (available from European Molecular Biology Laboratory - European Bioinformatic Institute, www.ebi.ac.uk/Tools/msa/tcoffee/)).

**[0047]** A useful example of algorithm that is suitable for sequence alignment and determining percent sequence identity and sequence similarity is the BLAST and BLAST 2.0 algorithms, which are described in Altschul et al., Nuc. Acids Res. 25:3389-3402 (1977) and Altschul et al., J. Mol. Biol. 215:403-410 (1990), respectively. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (website at www.ncbi.nlm.nih.gov/). Other examples of software programs that are suitable for multiple sequence alignment and determining percent sequence identity and sequence similarity include Geneious Aligner, ClustalW, Clustal Omega, and T-coffee. Any of these software programs may be used for sequence alignment in the invention disclosed herein.

In certain embodiments, Geneious Aligner or ClustalW is use for sequence alignment in the invention disclosed herein.

**[0048]** There are two different ways for splitting a glutamine synthetase polypeptide to generate a first and a second fragments of glutamine synthetase at an amino acid position. For example, splitting glutamine synthetase at amino acid position K52 of SEQ ID NO: 1 can be done such that the first glutamine synthetase fragment comprises amino acid residues 1 to 51 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 52 to 373 of SEQ ID NO: 1, or the first glutamine synthetase fragment comprises amino acid residues 1 to 52 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 53 to 373 of SEQ ID NO: 1. In the invention disclosed herein, splitting a glutamine synthetase polypeptide at an amino acid position include both ways of splitting the polypeptide.

**[0049]** Thus, when SEQ ID NO:1 is split at amino acid position E55, the first glutamine synthetase fragment can comprise amino acid residues 1 to 54 of SEQ ID NO: 1 and the second glutamine synthetase fragment can comprise amino acid residues 55 to 373 of SEQ ID NO: 1; or the first glutamine synthetase fragment can comprise amino acid residues 1 to 55 of SEQ ID NO: 1 and the second glutamine synthetase fragment can comprise amino acid residues 56 to 373 of SEQ ID NO: 1;

when SEQ ID NO:1 is split at amino acid position D92, the first glutamine synthetase fragment can comprise amino acid residues 1 to 91 of SEQ ID NO: 1 and the second glutamine synthetase fragment can comprise amino acid residues 92 to 373 of SEQ ID NO: 1; or the first glutamine synthetase fragment can comprise amino acid residues 1 to 92 of SEQ ID NO: 1 and the second glutamine synthetase fragment can comprise amino acid residues 93 to 373 of SEQ ID NO: 1;

when SEQ ID NO:1 is split at amino acid position G187, the first glutamine synthetase fragment can comprise amino acid residues 1 to 186 of SEQ ID NO: 1 and the second glutamine synthetase fragment can comprise amino acid residues 187 to 373 of SEQ ID NO: 1; or the first glutamine synthetase fragment can comprise amino acid residues 1 to 187 of SEQ ID NO: 1 and the second glutamine synthetase fragment can comprise amino acid residues 188 to 373 of SEQ ID NO: 1;

when SEQ ID NO:1 is split at amino acid position G245, the first glutamine synthetase fragment can comprise amino acid residues 1 to 244 of SEQ ID NO: 1 and the second glutamine synthetase fragment can comprise amino acid residues 245 to 373 of SEQ ID NO: 1;

or the first glutamine synthetase fragment can comprise amino acid residues 1 to 245 of SEQ ID NO: 1 and the second glutamine synthetase fragment can comprise amino acid residues 246 to 373 of SEQ ID NO: 1;

when SEQ ID NO:1 is split at amino acid position R262, the first glutamine synthetase fragment can comprise amino acid residues 1 to 261 of SEQ ID NO: 1 and the second glutamine synthetase fragment can comprise amino acid residues 262 to 373 of SEQ ID NO: 1; or the first glutamine synthetase fragment can comprise amino acid residues 1 to 262 of SEQ ID NO: 1 and the second glutamine synthetase fragment can comprise amino acid residues 263 to 373 of SEQ ID NO: 1;

when SEQ ID NO:1 is split at amino acid position K291, the first glutamine synthetase fragment can comprise amino acid residues 1 to 290 of SEQ ID NO: 1 and the second glutamine synthetase fragment can comprise amino acid residues 291 to 373 of SEQ ID NO: 1; or the first glutamine synthetase fragment can comprise amino acid residues 1 to 291 of SEQ ID NO: 1 and the second glutamine synthetase fragment can comprise amino acid residues 292 to 373 of SEQ ID NO: 1;

when SEQ ID NO:1 is split at amino acid position G302, the first glutamine synthetase fragment can comprise amino acid residues 1 to 301 of SEQ ID NO: 1 and the second glutamine synthetase fragment can comprise amino acid residues 302 to 373 of SEQ ID NO: 1; or the first glutamine synthetase fragment can comprise amino acid residues 1 to 302 of SEQ ID NO: 1 and the second glutamine synthetase fragment can comprise amino acid residues 303 to 373 of SEQ ID NO: 1;

when SEQ ID NO:1 is split at amino acid position D311, the first glutamine synthetase fragment can comprise amino acid residues 1 to 310 of SEQ ID NO: 1 and the second glutamine synthetase fragment can comprise amino acid residues 311 to 373 of SEQ ID NO: 1; or the first glutamine synthetase fragment can comprise amino acid residues 1 to 311 of SEQ ID NO: 1 and the second glutamine synthetase fragment can comprise amino acid residues 312 to 373 of SEQ ID NO: 1.

**[0050]** The first or second glutamine synthetase fragment does not have significant selectable activity when expressed alone, but they interact/associate to form a monomer and then a functional glutamine synthetase when co-expressed. The optimal activity of the glutamine synthetase can depend upon their interaction and association, and as such can be facilitated by interaction domains. Such interaction domains can be endogenous or heterologous to the first and second glutamine synthetase fragment.

**[0051]** In some embodiments, each of the first and second fragment of glutamine synthetase may be expressed as a fusion protein to an interaction domain. This can be achieved by, e.g., fusing the nucleic acid encoding each of the first and second fragment of glutamine synthetase in frame with a nucleic acid encoding an interaction domain. An interaction domain can be either N-terminal or C-terminal to the first or the second domain. When expressed, the interaction domain promotes interaction/association of the two fragments thereby allowing the formation of a functional glutamine synthetase and providing a selectable activity.

**[0052]** As used herein, an "interaction domain" is understood to include a domain (e.g., a polypeptide) capable of facilitating the interaction or association of two or more polypeptides. In some embodiments, the interaction domain is a dimerization domain. A dimerization domain can be a polypeptide capable of inducing interaction or association of two polypeptides. There are two types of dimers, those capable of forming homodimers (with the same sequence), or heterodimers (with another sequence).

**[0053]** In some embodiments, the interaction domain is a leucine zipper coiled coil polypeptide. A leucine zipper typically comprises about 35 amino acids containing a characteristic seven residue repeat with hydrophobic residues at the first and fourth residues of the repeat. Harbury et al., Science 262: 1401(1993). The two-stranded coiled-coil motif is characterized by two amphipathic α-helical chains wrapping around each other into a left-handed superhelix. Crick, F. H. C., Acta Crystallogr. 6, 689-69 (1953). Although there are two possible orientations of the α-helical chains to form a coiled-coil, parallel or antiparallel, a particular coiled-coil exists only in one specific orientation. See e.g., Monera, O. D., Kay, C. M., and Hodges, R. S. Biochemistry 33, 3862-3871 (1994). A leucine zipper is amenable to fusion to a polypeptide for oligomerizing the polypeptide as it is relatively small and is less likely to disrupt the polypeptide's normal function than would be a larger interaction domain.

**[0054]** In some embodiments, the interaction domain is a parallel leucine zipper. Parallel leucine zippers include those naturally exist as well as those designed and synthesized de novo based on the study of naturally occurring parallel coiled coils. See e.g., Lau SYM et al., J Biol Chem 259:13253-13261(1984); Betz SF et al., Curr Opin Struct Biol, 5:457-463 (1995). In some embodiments, the interaction domain is an anti-parallel leucine zipper. Anti-parallel leucine zippers include those naturally exist as well as those designed and synthesized de novo based on the study of naturally occurring anti-parallel coiled coils. See e.g., Oakley, M. G., and Hollenbeck, J. J. Curr Opin Struct Biol 11, 450-457 (2001); Ghosh, Hamilton and Regan, JACS 122, 5658-5659 (2000).

**[0055]** In some embodiments, the interaction domain is selected from a leucine zipper domain of GCN4, C/EBP, c-Fos, c-Jun, c-Myc and c-Max. In some preferred embodiments, the interaction domain comprises the sequence of ALKKELQANKKELAQLKWELQALKKELAQ EQLEKKLQALEKKLAQLEWKNQALEKKLAQ. Typically, when used in the inventions disclosed herein, ALKKELQANKKELAQLKWELQALKKELAQ is fused in frame to the C-terminal of a glutamine synthetase fragment, while EQLEKKLQALEKKLAQLEWKNQALEKKLAQ is fused in frame to the N-terminal of another glutamine synthetase fragment.

**[0056]** In some embodiments, the interaction domain is a dimerization domain such as a helix-loop-helix dimerization domain. Non-limiting examples of helix-loop-helix dimerization domains include those disclosed in Murre et al. Cell 58:537-544 (1989); the dimerization domain in the retinoic acid receptor, thyroid hormone receptor, other nuclear hormone receptors, see e.g., Kurokawa et al., Genes Dev. 7: 1423-1435(1993); and the dimerization domain in yeast transcription factors GAL4 and HAP1, see e.g., Marmonstein et al., Nature 356:408-414 (1992); Zhang et al., Proc. Natl. Acad. Sci. USA 90:2851-2855 (1993); and U.S. Patent No. 5,624,818).

**[0057]** In some embodiments, the interaction domain further comprises a linker. The linker links an interaction domain and the first or second fragment of glutamine synthetase. This can be achieved by, e.g., fusing a nucleic acid encoding the first or second glutamine synthetase in frame to a nucleic acid encoding a linker, which is then fused in frame to a nucleic acid encoding an interaction domain. Linkers can be any relatively short, flexible sequence that allows the interaction domain to interact such that the first and second fragment of glutamine synthetase associate to form a functional enzyme and provide a selectable activity.

**[0058]** Non-limiting examples of linkers are known in the art and include those having a sequence of GGPGG, GPGGG, or (GGGGS)n, where n is an integer of 1-4, and G (glycine), P (proline) and S (serine) are single letter amino acid codes. In some embodiments, the linker is a series of glycine and serine residues, for example, that described by Curtis et al. (1991; Proc Natl Acad Sci 88(13):5809-5813). In some embodiments, the linker comprises, consisting essentially of, or consisting of a sequence of GGPGG, GPGGG, GGGGSGGGGGS, (GGGGS)n, where n is an integer of 1-4. In some embodiments, the linker comprises, consisting essentially of, or consisting of a sequence of GGPGG, GPGGG, GGGGSGGGGGS or (GGGGS)n, where n is an integer of 1 or 2.

**[0059]** In some preferred embodiments, the interaction domain is an anti-parallel leucine zipper domain having the following sequence:
ALKKELQANKKELAQLKWELQALKKELAQ
EQLEKKLQALEKKLAQLEWKNQALEKKLAQ,
and the linker comprises the sequence of GGGGSGGGGS. In these embodiments, the configuration of the first and second fragment of glutamine synthetase, linker and interaction domain can be, from the N-terminal to C-terminal, the first fragment of glutamine synthetase-GGGGSGGGGS-ALKKELQANKKELAQLKWELQALKKELAQ; and EQLEKKLQALEKKLAQLEWKNQALEKKLAQ-GGGGSGGGGS-second fragment of glutamine synthetase. Each of the first and the third polypeptide (e.g., light and heavy chain of an antibody or a fragment thereof) can be linked in both orientations (e.g., N-terminal or C-terminal) to each of the first or the second fragment of glutamine synthetase. For example, in some embodiments, one chain of an antibody is N-terminal to a first fragment of glutamine synthetase and is fused in frame with the first fragment, and the other chain is C-terminal to a second fragment of glutamine synthetase and is fused in frame with the second fragment; in some embodiments, one chain of an antibody is C-terminal to a first fragment of glutamine synthetase and is fused in frame with the first fragment, and the other chain is N-terminal to a second fragment of glutamine synthetase and is fused in frame with the second fragment. Alternatively, the configuration of the first and second fragment of glutamine synthetase, linker and interaction domain can be, from the N-terminal to C-terminal, EQLEKKLQALEKKLAQLEWKNQALEKKLAQ-GGGGSGGGGS-first fragment of glutamine synthetase; and the second fragment of glutamine synthetase-GGGGSGGGGS-ALKKELQANKKELAQLKWELQALKKELAQ. Similarly, each of the first and the third polypeptide (e.g., light and heavy chain of an antibody or a fragment thereof) can be linked in both orientations (e.g., N-terminal or C-terminal) to each of the first or the second fragment of glutamine synthetase.

**[0060]** In some embodiments, the vector disclosed herein or one or both vectors in the expression system disclosed herein may further comprise an internal ribosomal entry site (IRES). IRES facilitates the initiation of translation of an mRNA from an internal site (i.e., a site other than the 5 'end of the mRNA). In some embodiments, an IRES occurs at a site between the first nucleic acid and the second nucleic acid; in some embodiments, an IRES occurs at a site between the third nucleic acid and the fourth nucleic acid; in some

embodiments, an IRES occurs at sites between both first and second, and third and fourth nucleic acids.

**[0061]** IRES is well known and used in the art. One example of a suitable IRES is the IRES of encephalomyocarditis virus (ECMV), as described in Jang and Wimmer Genes &Development 4 1560 (1990) and Jang, Davies, Kaufman and Wimmer J. Vir. 63 1651 (1989). The residues 335-848 of EMCV form a suitable IRES; other variants or portions of ECMV IRES are known and will be suitable for use in the present invention. A suitable portion or variant of an IRES is one that will confer sufficient translation of the second open reading frame (ORF). Additionally, the 3' end of an IRES may be altered (or mutated) to reduce the efficiency of translation, thereby providing a means to enhance selection and/or amplification methods. For example, the efficiency of the IRES can be decreased by using a sequence previously shown to allow efficient selection and amplification. Aldrich et al., Biotechnol Prog 19, 1433 (2003). In some embodiments, the IRES comprises the sequence of GATGATAATACCCTCGAGATCCGTGCCATCATG. Alternative sequences are known, or can be determined by one of ordinary skill in the art.

**[0062]** In some embodiments, the vector disclosed herein or the one or both vectors in the expression system disclosed herein may further comprise an expression augmenting sequence element (EASE). When used in expression vectors, EASE allows the development of stable CHO cell pools for a period of time (e.g., five to seven weeks) that express high levels of recombinant protein. See e.g., Aldrich, T. L., Cytotechnology, 28(1-3): 9–17 (1998). EASE sequences known in the art may be used in the vector or expression system disclosed herein.

**[0063]** The nucleic acids encoding a component of the desired heteromeric complex can be obtained as a cDNA or as a genomic DNA by methods known in the art. For example, messenger RNA coding for a desired component can be isolated from a suitable source employing standard techniques of RNA isolation, and the use of oligo-dT cellulose chromatography to segregate the poly-A mRNA. When the heteromeric complex to be expressed is an antibody, suitable sources of desired nucleic acids can be isolated from mature B cells or a hybridoma culture. In addition, the nucleic acids for use in the invention can be obtained by chemical synthesis.

**[0064]** In some embodiments, one or both vectors of the expression system disclosed herein may further comprise a nucleic acid encoding a different functional selectable marker, in addition to a first and/or second fragment of glutamine synthetase and a polypeptide of a heteromeric complex. As used herein, a “different functional selectable marker” is a protein

with a selectable activity different from glutamine synthetase. Well known selectable markers such as zeomycin (zeo), neomycin, which confers resistance to the aminoglycoside G-418, Colberre-Garapin et al., J. Mol. Biol. 150:1(1981), puromycin (PAC), Blasticidin S (BlaS), or GPT which confers resistance to mycophenolic acid, Mulligan & Berg, Proc. Natl. Acad. Sci. USA 78:2072(1981), etc., and hygro, which confers resistance to hygromycin, Santerre et al., Gene 30: 147(1984) can be used as different functional selectable markers.

**[0065]** Metabolic enzymes that confer cell survival or induce cell death under prescribed conditions can also be used as a different functional selectable marker. Examples include but are not limited to: dihydrofolate reductase (DHFR); herpes simplex virus thymidine kinase (TK), Wigler et al., Cell 11:223(1977), hypoxanthine-guanine phosphoribosyltransferase (HGPRT), Szybalska & Szybalski, Proc. Natl. Acad. Sci. USA 48:2026 (1962), and adenine phosphoribosyltransferase (APRT), Lowy et al., Cell 22:817(1980), which are genes which can be employed in cells lacking TK, HGPRT or APRT, respectively.

**[0066]** Selectable markers that are based on color selection can also be used as a different functional selectable marker. In a particular example, beta-galactosidase can be used, Blau et al., WO 98/44350. Fluorescence markers can also be used in the methods of the present invention, for example, GFP has been used for clonal selection of cells to measure protein interactions in protein-fragment complementation assays, Remy and Michnick, Proc. Natl. Acad. Sci., 96:5394- 5399(1999). Similarly, fluorescein-conjugated methotrexate can be used to detect cells expressing functional glutamine synthetase. An advantage for fluorescent markers is that this selection can be done in any animal cell type and is not restricted to those having a deficiency in a metabolic pathway or does not require a drug sensitivity, e.g., to neomycin.

**[0067]** Thus, in such embodiments, each of the vectors of the expression system disclosed herein comprises a nucleic acid encoding one of two polypeptides that can form a heteromeric complex operably linked to a nucleic acid encoding one of two fragments of glutamine synthetase, as well as a nucleic acid encoding a different, functional selectable marker. Further, the two polypeptides encoded by the nucleic acid of each vector can associate to form a complex, and the two fragments of glutamine synthetase can associate to provide a selectable activity, and the additional selectable marker from one or each vector provides selectable activities different than glutamine synthetase.

**[0068]** For example, in some embodiments, the first vector can further comprise a nucleic acid encoding a first different functional selectable marker (e.g., resistance to

neomycin) and the second vector can further comprise a nucleic acid encoding a second different functional selectable marker (e.g., resistance to zeomycin) or only one vector can contain the additional different functional selectable marker. Thus, one vector is transfected into a host cell and selection is applied (e.g., the drug G418 is added to neomycin resistant cells). After selection, conventional methods can be used to determine the presence of the vector and the expression level of the polypeptides encoded by the nucleic acids on the vector, for example by PCR, Southern blot, ELISA, western blot, and the like. Once high-level expression has been obtained, the second vector is transfected into the cell line. While maintaining selection for the first vector, selection is applied for the second selectable marker (e.g., zeomycin resistance) and the presence of the second vector and expression of the respective vector encoded proteins are assessed. In such embodiments, once it has been determined that both vectors are present, selection is applied for expression of functional glutamine synthetase that have associated in the cell to provide a selectable activity as described herein.

**[0069]** Alternatively, in some embodiments, both vectors are transfected simultaneously, and selection is applied at the same time. In some embodiments, only one vector further comprises a different functional selectable marker and both vectors are transfected simultaneously, and selection for the different functional selectable marker is applied. Once it has been determined that both vectors are present, selection is applied for expression of functional glutamine synthetase that have associated in the cell to provide a selectable activity, as described herein.

**[0070]** In yet some other embodiments, vectors can further comprise a nucleic acid encoding different functional selectable markers are each transfected into separate cell lines. Once selection is applied and clones have been identified that express high levels of the proteins encoded by each desired vector, the cells are fused as described in Hori et al. (U.S. Patent No. 5,916,771). Once fusion is complete, selection is applied for the selectable activity provided by the subunits.

**[0071]** In yet some other embodiments, the first and second vectors that do not comprise a different functional selectable marker are transfected simultaneously with a third vector. The third vector encodes for a separate selectable activity, such as for example, neomycin resistance or beta galactosidase that can allow for a preliminary selection of cells that were successfully transfected. Once this preliminary selection has been performed, selection can be applied for the selectable activity of glutamine synthetase. In these embodiments, equal

quantities of the two vectors are transfected while the third vector is transfected at one-third the concentration of the first two vectors (e.g., a ratio of 3:3:1 or 6:6:1 or the like). One of skill in the art will recognize that variations in the ratios are within the scope of the invention.

**[0072]** Vectors and expression systems disclosed herein can be used for producing heteromeric complexes (e.g., antibodies and bispecific antigen binding molecules). For example, vectors or expression systems disclosed herein can be transfected into host cells using methods described above and other methods known in the art. Once it has determined that the vector or the expression system (two vectors) are present in the host cell, the cells then grow under appropriate conditions for the selectable activity of glutamine synthetase such that components of a heteromeric protein are expressed in appropriate amounts to form heteromeric proteins. For host cells that do not express endogenous glutamine synthetase (e.g., glutamine synthetase deficient cells), transfectants can be selected by, e.g., growing the cells in a cell culture medium free of glutamine synthetase. For host cells that express endogenous glutamine synthetase, transfectants may be selected by, e.g., growing the cells in a cell culture medium containing the glutamine synthetase inhibitor methionine sulfoximine (MSX) (e.g., growing the cells in the presence of a toxic level of MSX). The expressed proteins can be harvested and purified from the cells and/or the culture medium.

**[0073]** As used herein, the term "cell culture" is understood to include the growth and propagation of cells outside of a multicellular organism or tissue. Typically, cell culture is performed under sterile, controlled temperature and atmospheric conditions in tissue culture plates (e.g., 10-cm plates, 96 well plates, etc.), or other adherent culture (e.g., on microcarrier beads) or in suspension culture such as in roller bottles. Cultures can be grown in shake flasks, small scale bioreactors, and/or large-scale bioreactors. A bioreactor is a device used to culture cells in which environmental conditions such as temperature, atmosphere, agitation, and/or pH can be monitored and adjusted. A number of companies (e.g., ABS Inc., Wilmington, Del.; Cell Trends, Inc., Middletown, Md.) as well as university and/or government- sponsored organizations (e.g., The Cell Culture Center, Minneapolis, Minn.) offer cell culture services on a contract basis.

**[0074]** Optimal periods for which the cultures are in contact with agents that select for the selectable activity are for longer than the typical period for one normal growth cycle (e.g., for Chinese hamster ovary cells (CHO cells), where one growth cycle has been reported to be approximately 20-22 hours (Rasmussen et al. (1998), Cytotechnology, 28:31-42)). As such, in some embodiments, the cultures comprise selectable conditions, e.g., drugs, metabolites, or

color substrates, preferably for at least about one day, more preferably for at least about 3 days, and even more preferably for at least about 5 days or at least about 7 days.

**[0075]** In some embodiments, vectors or expression systems disclosed herein can be transfected into a suitable host cell. In some embodiments, the host cell is a mammalian cell. A wide variety of mammalian cells suitable for growth in culture are available from, for example, the American Type Culture Collection (ATCC, Manassas, Va.) and NRRL (Peoria, 111.). Non-limiting examples of mammalian cells typically used in the industrial or academic laboratory include CHO, VERO, BHK, HeLa, Cos, CV1, MDCK, 293, 3T3, PC12, myeloma (e.g., NSO), and WI38 cell lines. In addition, new mammalian cell lines can be established using methods well known by those skilled in the art (e.g., by transformation, viral infection, and/or selection). In some embodiments, the host cell is CHO, VERO, BHK, HeLa, Cos, CV1, MDCK, 293, 3T3, PC12, or NSO cells. In some embodiments, the host cell is CHO cells. CHO cells are easy to manipulate as adherent or suspension cultures and exhibit relatively good genetic stability.

**[0076]** In some embodiments, the host cell is a cell that does not express endogenous glutamine synthetase (e.g., a mammalian cell with endogenous glutamine synthetase knocked out). Cell lines that do not express endogenous glutamine synthetase include the glutamine synthetase deficient CHO cells such as CHOZN® GS-/- ZFN-modified CHO cells (Sigma Aldrich Fine Chemicals, St. Louis MO). Glutamine synthetase deficient CHO cells (or other mammalian cells) can also be prepared using methods that are known in the art.

**[0077]** In some embodiments, the host cell is a non-mammalian cell. Non-mammalian cell lines that can be used, for example, plant cell lines, insect cell lines (e.g., sf9), yeast cells or bacterial cells such as E. coli.

**[0078]** Heteromeric complexes (e.g., monoclonal antibodies) expressed using the vectors and expression systems disclosed herein can be recovered from the cell culture, e.g., from the host cell in cases where the heteromeric complexes are not secreted, and from the culture media in cases where the heteromeric complexes are secreted by the cells. However, the expression systems also encompass engineered host cells that express the heteromeric complexes anchored in the cell membrane.

**[0079]** Heteromeric complexes (e.g., monoclonal antibodies) expressed by the methods of the invention can be harvested. In addition, the complexes can be purified, or partially purified, from such culture or component (e.g., from culture medium or cell extracts) using known processes. The phrase "partially purified" means that some fractionation procedure, or procedures, have been carried out, but that more polypeptide species (at least 10%) than the

desired protein is present. By "purified" is meant that the protein is essentially homogeneous, e.g., less than 1%, less than 0.5%, less than 0.3%, or less than 0.1% contaminating proteins are present. Purification procedures can include but are not limited to one or more steps of filtration, centrifugation, precipitation, phase separation, affinity purification, gel filtration, ion exchange chromatography, size exclusion chromatography (SEC), hydrophobic interaction chromatography (HIC); using such resins as phenyl ether, butyl ether, or propyl ether), HPLC, or some combination of above.

**[0080]** The invention also optionally encompasses further formulating the proteins. By the term "formulating" is meant that the proteins can be buffer exchanged, sterilized, bulk-packaged and/or packaged for a final user. For purposes of the invention, the term "sterile bulk form" means that a formulation is free, or essentially free, of microbial contamination (to such an extent as is acceptable for food and/or drug purposes), and is of defined composition and concentration.

**[0081]** The term "sterile unit dose form" means a form that is appropriate for the customer and/or patient administration or consumption. Such compositions can comprise an effective amount of the protein, in combination with other components such as a physiologically acceptable diluent, carrier, or excipient. The term "physiologically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s).

**[0082]** The invention will be more fully understood by reference to the following examples. The examples should not, however, be construed as limiting the scope of the invention.

## EXAMPLES

**Example 1.** Identification of Suitable Amino Acid Positions of Glutamine Synthetase as Sites for Splitting the Polypeptide

**[0083]** Highly solvent exposed residues on multiple green (and yellow) fluorescence proteins have been targeted for splitting the fluorescence proteins, and polypeptide fragments generated by splitting at those amino acid positions were found to be able to interact/associate to form a functional protein. Similarly, highly solvent exposed residues were also used to split the protein Luciferase. See Ishikawa et. al., Protein Eng Des Sel. Dec;25(12):813-20 (2012). However, contrary to what was reported in literature, we found that solvent accessibility of the

backbone atoms (e.g., N, CA, C, O) only, as opposed to the complete residue, can be used to rank exposure of residues and such ranking was used to identify splitting sites.

**[0084]** Structures of selected glutamine synthetase (GS) from different species, including GS from maize, human, canine, and were obtained from the Protein Data Bank to help building a model of the mouse GS. The methionine sulfoximine (MSX) bound structure of human GS was used to build a homology model of the decamer of the mouse GS. Using the model, we analyzed solvent accessibility of the backbone atoms of the mouse glutamine synthetase and identified nine potential amino acid positions at which to split the polypeptide such that two glutamine synthetase fragments obtained by splitting the polypeptide at each position can interact/associate to form a functional glutamine synthetase. The nine amino acid positions were K52, E55, D92, G187, G245, R262, K291, G302 and D311 of the mouse glutamine synthetase (SEQ ID NO:1), with two possible ways to split the polypeptide at each position as listed in Table 1 below.

Table 1. List of amino acid positions for splitting glutamine synthetase

| AA position | N-terminal half | C-terminal half |
|---|---|---|
| 52 | 1-51 K52 | 52-373 |
| | 1-52 | 53-373 |
| | | |
| 55 | 1-54 E55 | 55-373 |
| | 1-55 | 56-373 |
| | | |
| 92 | 1-91 D92 | 92-373 |
| | 1-92 | 93-373 |
| | | |
| 245 | 1-244 G245 | 245-373 |
| | 1-245 | 246-373 |
| | | |
| 187 | 1-186 G187 | 187-373 |
| | 1-187 | 188-373 |
| | | |
| 262 | 1-261 R262 | 262-373 |
| | 1-263 | 264-373 |
| | | |
| 311 | 1-310 D311 | 311-373 |
| | 1-311 | 312-373 |
| | | |
| 302 | 1-301 G302 | 302-373 |
| | 1-302 | 303-373 |
| | | |

| | | |
|---|---|---|
| 291 | 1-291 K291 | 292-373 |
| | 1-292 | 293-373 |

**EXAMPLE 2**. Expression of Glutamine Synthetase using Two Vectors

**[0085]** Among the nine splitting sites identified in Example 1, five (K52, E55, G187, D92 and G245) were tested in this experiment. For each splitting site, a pair of linear plasmids as well as a pair of circular plasmids were constructed. An anti-parallel leucine zipper was used as the interaction domain together with a linker having the sequence of GGGGSGGGGS. Table 2 below shows the list of vectors constructed in this experiment. In Table 2, 51-L1 stands for a pair of linear plasmids carrying amino acid residues 1-51 and 52-373 respectively, of glutamine synthetase (SEQ ID NO:1); 51-L2 refers to the duplicate of 51-L1. 51-C1 refers to a pair of circular plasmids carrying residues 1-51 and 52-373 respectively, of glutamine synthetase (SEQ ID NO:1); 51-C2 refers to the duplicate of 51-C1. Similar to 51-L1, L2, C1, and C2, constructs 52-L1, L2, C1, C2 were also made since split sites separated by one amino acid may have an impact on interaction/association of two glutamine synthetase fragments. See Ishikawa et. al. Protein Eng Des Sel., 25(12):813-20 (2012).

**[0086]** Figure 2 shows the map of the two vectors of the expression system used in the experiments. Table 3 below shows the configuration of the two glutamine synthetase fragments together with the interaction domain (the underlined sequences) and linker sequence in each plasmid prepared.

Table 2 List of vectors constructed

| | CLD Abbreviation | N-terminal GS* fragment | C-terminal GS fragment |
|---|---|---|---|
| 1 | 51-L1, L2, C1, C2 | 1-51 K52 | 52-373 |

| | | | |
|---|---|---|---|
| | 52-L1, L2, C1, C2 | 1-52 | 53-373 |
| | | | |
| 2 | 54-L1, L2, C1, C2 | 1-54 E55 | 55-373 |
| | 55-L1, L2, C1, C2 | 1-55 | 56-373 |
| | | | |
| 3 | 186-L1, L2, C1, C2 | 1-186 G187 | 187-373 |
| | 187-L1, L2, C1, C2 | 1-187 | 188-373 |
| | | | |
| 4 | 91-L1, L2, C1, C2 | 1-91 D92 | 92-373 |
| | 92-L1, L2, C1, C2 | 1-92 | 93-373 |
| | | | |
| 5 | 244-L1, L2, C1, C2 | 1-244 G245 | 245-373 |
| | 245-L1, L2, C1, C2 | 1-245 | 246-373 |

*GS: glutamine synthetase

Table 3 Configuration of Glutamine synthetase fragment, interaction domain and linker

| | **N-terminal GS fragment** | **C-terminal fragment** |
|---|---|---|
| 1 | >1-51 K52<br>MATSASSHLNKGIKQMYMSLPQGEKVQAM<br>YIWVDGTGEGLRCKTRTLDCEPGGGGSGG<br>GGSALKKELQANKKELAQLKWELQALKKE<br>LAQ | >52-373<br>EQLEKKLQALEKKLAQLEWKNQALEK<br>KLAQGGGGSGGGGSKCVEELPEWNFD<br>GSSTFQSEGSNSDMYLHPVAMFRDPFR<br>KDPNKLVLCEVFKYNRKPAETNLRHIC<br>KRIMDMVSNQHPWFGMEQEYTLMGTD<br>GHPFGWPSNGFPGPQGPYYCGVGADKA<br>YGRDIVEAHYRACLYAGVKITGTNAEV<br>MPAQWEFQIGPCEGIRMGDHLWIARFIL<br>HRVCEDFGVIATFDPKPIPGNWNGAGC<br>HTNFSTKAMREENGLKCIEEAIDKLSKR<br>HQYHIRAYDPKGGLDNARRLTGFHETS<br>NINDFSAGVANRGASIRIPRTVGQEKKG<br>YFEDRRPSANCDPYAVTEAIVRTCLLNE<br>TGDEPFQYKN |
| | >1-52<br>MATSASSHLNKGIKQMYMSLPQGEKVQAM<br>YIWVDGTGEGLRCKTRTLDCEPKGGGGSG<br>GGGSALKKELQANKKELAQLKWELQALK<br>KELAQ | >53-373<br>EQLEKKLQALEKKLAQLEWKNQALEK<br>KLAQGGGGSGGGGSCVEELPEWNFDGS<br>STFQSEGSNSDMYLHPVAMFRDPFRKD<br>PNKLVLCEVFKYNRKPAETNLRHICKRI<br>MDMVSNQHPWFGMEQEYTLMGTDGH<br>PFGWPSNGFPGPQGPYYCGVGADKAYG<br>RDIVEAHYRACLYAGVKITGTNAEVMP<br>AQWEFQIGPCEGIRMGDHLWIARFILHR<br>VCEDFGVIATFDPKPIPGNWNGAGCHT<br>NFSTKAMREENGLKCIEEAIDKLSKRHQ<br>YHIRAYDPKGGLDNARRLTGFHETSNIN |

| | N-terminal GS fragment | C-terminal fragment |
|---|---|---|
| | | DFSAGVANRGASIRIPRTVGQEKKGYFE DRRPSANCDPYAVTEAIVRTCLLNETGD EPFQYKN |
| 2 | >1-54 E55<br>MATSASSHLNKGIKQMYMSLPQGEKVQAM YIWVDGTGEGLRCKTRTLDCEPKCV GGGGSGGGGSALKKELQANKKELAQLKW ELQALKKELAQ | >55-373<br>EQLEKKLQALEKKLAQLEWKNQALEK KLAQGGGGSGGGGSEELPEWNFDGSST FQSEGSNSDMYLHPVAMFRDPFRKDPN KLVLCEVFKYNRKPAETNLRHICKRIMD MVSNQHPWFGMEQEYTLMGTDGHPFG WPSNGFPGPQGPYYCGVGADKAYGRDI VEAHYRACLYAGVKITGTNAEVMPAQ WEFQIGPCEGIRMGDHLWIARFILHRVC EDFGVIATFDPKPIPGNWNGAGCHTNFS TKAMREENGLKCIEEAIDKLSKRHQYHI RAYDPKGGLDNARRLTGFHETSNINDFS AGVANRGASIRIPRTVGQEKKGYFEDRR PSANCDPYAVTEAIVRTCLLNETGDEPF QYKN |
| | >1-55<br>MATSASSHLNKGIKQMYMSLPQGEKVQAM YIWVDGTGEGLRCKTRTLDCEPKCVEGGG GSGGGGSALKKELQANKKELAQLKWELQA LKKELAQ | >56-373<br>EQLEKKLQALEKKLAQLEWKNQALEK KLAQGGGGSGGGGSELPEWNFDGSSTF QSEGSNSDMYLHPVAMFRDPFRKDPNK LVLCEVFKYNRKPAETNLRHICKRIMD MVSNQHPWFGMEQEYTLMGTDGHPFG WPSNGFPGPQGPYYCGVGADKAYGRDI VEAHYRACLYAGVKITGTNAEVMPAQ WEFQIGPCEGIRMGDHLWIARFILHRVC EDFGVIATFDPKPIPGNWNGAGCHTNFS TKAMREENGLKCIEEAIDKLSKRHQYHI RAYDPKGGLDNARRLTGFHETSNINDFS AGVANRGASIRIPRTVGQEKKGYFEDRR PSANCDPYAVTEAIVRTCLLNETGDEPF QYKN |
| 3 | >1-91 D92<br>MATSASSHLNKGIKQMYMSLPQGEKVQAM YIWVDGTGEGLRCKTRTLDCEPKCVEELPE WNFDGSSTFQSEGSNSDMYLHPVAMFRDP FRKGGGGSGGGGSALKKELQANKKELAQL KWELQALKKELAQ | >92-373<br>EQLEKKLQALEKKLAQLEWKNQALEK KLAQGGGGSGGGGSDPNKLVLCEVFKY NRKPAETNLRHICKRIMDMVSNQHPWF GMEQEYTLMGTDGHPFGWPSNGFPGPQ GPYYCGVGADKAYGRDIVEAHYRACL YAGVKITGTNAEVMPAQWEFQIGPCEGI RMGDHLWIARFILHRVCEDFGVIATFDP KPIPGNWNGAGCHTNFSTKAMREENGL KCIEEAIDKLSKRHQYHIRAYDPKGGLD NARRLTGFHETSNINDFSAGVANRGASI RIPRTVGQEKKGYFEDRRPSANCDPYA VTEAIVRTCLLNETGDEPFQYKN |

| | N-terminal GS fragment | C-terminal fragment |
|---|---|---|
| | >1-92<br>MATSASSHLNKGIKQMYMSLPQGEKVQAM<br>YIWVDGTGEGLRCKTRTLDCEPKCVEELPE<br>WNFDGSSTFQSEGSNSDMYLHPVAMFRDP<br>FRKDGGGGSGGGGSALKKELQANKKELAQ<br>LKWELQALKKELAQ | >93-373<br>EQLEKKLQALEKKLAQLEWKNQALEK<br>KLAQGGGGSGGGGSPNKLVLCEVFKYN<br>RKPAETNLRHICKRIMDMVSNQHPWFG<br>MEQEYTLMGTDGHPFGWPSNGFPGPQG<br>PYYCGVGADKAYGRDIVEAHYRACLY<br>AGVKITGTNAEVMPAQWEFQIGPCEGIR<br>MGDHLWIARFILHRVCEDFGVIATFDPK<br>PIPGNWNGAGCHTNFSTKAMREENGLK<br>CIEEAIDKLSKRHQYHIRAYDPKGGLDN<br>ARRLTGFHETSNINDFSAGVANRGASIRI<br>PRTVGQEKKGYFEDRRPSANCDPYAVT<br>EAIVRTCLLNETGDEPFQYKN |
| 4 | >1-186 G187<br>MATSASSHLNKGIKQMYMSLPQGEKVQAM<br>YIWVDGTGEGLRCKTRTLDCEPKCVEELPE<br>WNFDGSSTFQSEGSNSDMYLHPVAMFRDP<br>FRKDPNKLVLCEVFKYNRKPAETNLRHICK<br>RIMDMVSNQHPWFGMEQEYTLMGTDGHP<br>FGWPSNGFPGPQGPYYCGVGADKAYGRDI<br>VEAHYRACLYAGGGGSGGGGSALKKELQA<br>NKKELAQLKWELQALKKELAQ | >187-373<br>EQLEKKLQALEKKLAQLEWKNQALEK<br>KLAQGGGGSGGGGSGVKITGTNAEVMP<br>AQWEFQIGPCEGIRMGDHLWIARFILHR<br>VCEDFGVIATFDPKPIPGNWNGAGCHT<br>NFSTKAMREENGLKCIEEAIDKLSKRHQ<br>YHIRAYDPKGGLDNARRLTGFHETSNIN<br>DFSAGVANRGASIRIPRTVGQEKKGYFE<br>DRRPSANCDPYAVTEAIVRTCLLNETGD<br>EPFQYKN |
| | >1-187<br>MATSASSHLNKGIKQMYMSLPQGEKVQAM<br>YIWVDGTGEGLRCKTRTLDCEPKCVEELPE<br>WNFDGSSTFQSEGSNSDMYLHPVAMFRDP<br>FRKDPNKLVLCEVFKYNRKPAETNLRHICK<br>RIMDMVSNQHPWFGMEQEYTLMGTDGHP<br>FGWPSNGFPGPQGPYYCGVGADKAYGRDI<br>VEAHYRACLYAGGGGSGGGGSALKKELQ<br>ANKKELAQLKWELQALKKELAQ | >188-373<br>EQLEKKLQALEKKLAQLEWKNQALEK<br>KLAQGGGGSGGGGSVKITGTNAEVMPA<br>QWEFQIGPCEGIRMGDHLWIARFILHRV<br>CEDFGVIATFDPKPIPGNWNGAGCHTNF<br>STKAMREENGLKCIEEAIDKLSKRHQYH<br>IRAYDPKGGLDNARRLTGFHETSNINDF<br>SAGVANRGASIRIPRTVGQEKKGYFEDR<br>RPSANCDPYAVTEAIVRTCLLNETGDEP<br>FQYKN |
| 5 | >1-244 G245<br>MATSASSHLNKGIKQMYMSLPQGEKVQAM<br>YIWVDGTGEGLRCKTRTLDCEPKCVEELPE<br>WNFDGSSTFQSEGSNSDMYLHPVAMFRDP<br>FRKDPNKLVLCEVFKYNRKPAETNLRHICK<br>RIMDMVSNQHPWFGMEQEYTLMGTDGHP<br>FGWPSNGFPGPQGPYYCGVGADKAYGRDI<br>VEAHYRACLYAGVKITGTNAEVMPAQWEF<br>QIGPCEGIRMGDHLWIARFILHRVCEDFGVI<br>ATFDPKPIPGGGGSGGGGSALKKELQANKK<br>ELAQLKWELQALKKELAQ | >245-373<br>EQLEKKLQALEKKLAQLEWKNQALEK<br>KLAQGGGGSGGGGSGNWNGAGCHTNF<br>STKAMREENGLKCIEEAIDKLSKRHQYH<br>IRAYDPKGGLDNARRLTGFHETSNINDF<br>SAGVANRGASIRIPRTVGQEKKGYFEDR<br>RPSANCDPYAVTEAIVRTCLLNETGDEP<br>FQYKN |

| | N-terminal GS fragment | C-terminal fragment |
|---|---|---|
| | >1-245<br>MATSASSHLNKGIKQMYMSLPQGEKVQAM<br>YIWVDGTGEGLRCKTRTLDCEPKCVEELPE<br>WNFDGSSTFQSEGSNSDMYLHPVAMFRDP<br>FRKDPNKLVLCEVFKYNRKPAETNLRHICK<br>RIMDMVSNQHPWFGMEQEYTLMGTDGHP<br>FGWPSNGFPGPQGPYYCGVGADKAYGRDI<br>VEAHYRACLYAGVKITGTNAEVMPAQWEF<br>QIGPCEGIRMGDHLWIARFILHRVCEDFGVI<br>ATFDPKPIPGGGGGSGGGGSALKKELQANK<br>KELAQLKWELQALKKELAQ | >246-373<br>EQLEKKLQALEKKLAQLEWKNQALEK<br>KLAQGGGGSGGGGSNWNGAGCHTNFS<br>TKAMREENGLKCIEEAIDKLSKRHQYHI<br>RAYDPKGGLDNARRLTGFHETSNINDFS<br>AGVANRGASIRIPRTVGQEKKGYFEDRR<br>PSANCDPYAVTEAIVRTCLLNETGDEPF<br>QYKN |

**[0087]** Each pair of the plasmid constructs were transfected into glutamine synthetase knock out host CHO cells (endogenous glutamine synthetase gene knocked out using recombinant DNA technology), which is a glutamine synthetase deficient, serum free suspension growth adapted CHO cell line derived from CHO-K1 (Kao and Puck, 1968). These host cells are auxotrophic for glutamine, and therefore require presence of glutamine in the growth medium for survival. The host cells were co-transfected with each of the two plasmid constructs using a standard electroporation procedure. Transfected cells that had successfully integrated a fully-functional GS enzyme were able to grow in media lacking glutamine. After transfection, the cells were grown in selective growth media, lacking glutamine, until viability reached >90%. The resulting cell population was referred to as the stable pool. The recovery data is presented in Figures 3 and 4. In both figures, CS9 represents a CHO DHFR-cell line derived from DuxB11, pGS is a vector carrying the DNA of full Glutamine Synthetase, which was used as a positive control.

**[0088]** Glutamine synthetase knock out CHO cells transfected with both linear and circular plasmids comprising glutamine synthetase fragments generated by splitting at amino acid position G187 were able to grow in cell culture medium lacking glutamine (Figures 3 and 4). In addition, glutamine synthetase knock out CHO cells transfected with linear plasmids comprising glutamine synthetase fragments generated by splitting at amino acid position D92 were able to grow in cell culture medium lacking glutamine (Figure 3).

**[0089]** All references cited in this application are incorporated by reference herein.

# CLAIMS

What is claimed:

1. A vector comprising:

a) a first nucleic acid encoding a first polypeptide,

b) a second nucleic acid encoding a first fragment of glutamine synthetase, wherein the transcription of the first nucleic acid is operably linked to the transcription of the second nucleic acid,

c) a third nucleic acid encoding a third polypeptide, the third polypeptide is capable of associating with the first polypeptide to form a heteromeric complex, and

d) a fourth nucleic acid encoding a second fragment of glutamine synthetase, wherein the transcription of the third nucleic acid is operably linked to the transcription of the fourth nucleic acid, and

wherein the first fragment and the second fragment of glutamine synthetase associate to provide a selectable activity, and wherein the vector is capable of being transfected into mammalian cells and improving selection of transfected cells

2. The vector of claim 1, wherein

a) the first fragment of glutamine synthetase is an N-terminal fragment of glutamine synthetase and the second fragment of glutamine synthetase is a C-terminal fragment of glutamine synthetase, or

b) the first fragment of glutamine synthetase is a C-terminal fragment of glutamine synthetase and the second fragment of glutamine synthetase is an N-terminal fragment of glutamine synthetase.

3. The vector of claim 1 or claim 2, wherein the glutamine synthetase comprising the amino acid sequence of SEQ ID NO: 1, and the first and second fragments of glutamine synthetase are generated by splitting the glutamine synthetase polypeptide at an amino acid position selected from K52, E55, D92, G187, G245, R262, K291, G302 and D311 of SEQ ID NO: 1.

4. The vector of claim 3, wherein the first and second fragments of glutamine synthetase are generated by splitting the glutamine synthetase polypeptide at an amino acid position selected from K52, E55, D92, G187, and G245 of SEQ ID NO: 1.

5. The vector of claim 4, wherein the first and second fragments of glutamine synthetase are generated by splitting the glutamine synthetase polypeptide at the amino acid position D92 or G187 of SEQ ID NO: 1.

6. The vector of claim 1 or claim 2, wherein the glutamine synthetase is a mammalian glutamine synthetase having an amino acid sequence different from SEQ ID NO:1, and the first and second fragments of glutamine synthetase are generated by splitting the mammalian glutamine synthetase polypeptide at an amino acid position equivalent to an amino acid position selected from K52, E55, D92, G187, G245, R262, K291, G302 and D311 of SEQ ID NO: 1 according to sequence alignment.

7. The vector of claim 3, wherein:

a) the first glutamine synthetase fragment comprises amino acid residues 1 to 51 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 52 to 373 of SEQ ID NO: 1;

b) the first glutamine synthetase fragment comprises amino acid residues 1 to 52 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 53 to 373 of SEQ ID NO: 1;

c) the first glutamine synthetase fragment comprises amino acid residues 1 to 54 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 55 to 373 of SEQ ID NO: 1;

d) the first glutamine synthetase fragment comprises amino acid residues 1 to 55 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 56 to 373 of SEQ ID NO: 1;

e) the first glutamine synthetase fragment comprises amino acid residues 1 to 91 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 92 to 373 of SEQ ID NO: 1;

f) the first glutamine synthetase fragment comprises amino acid residues 1 to 92 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 93 to 373 of SEQ ID NO: 1;

g) the first glutamine synthetase fragment comprises amino acid residues 1 to 186 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 187 to 373 of SEQ ID NO: 1;

h) the first glutamine synthetase fragment comprises amino acid residues 1 to 187 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 188 to 373 of SEQ ID NO: 1;

i) the first glutamine synthetase fragment comprises amino acid residues 1 to 244 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 245 to 373 of SEQ ID NO: 1;

j) the first glutamine synthetase fragment comprises amino acid residues 1 to 245 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 246 to 373 of SEQ ID NO: 1;

k) the first glutamine synthetase fragment comprises amino acid residues 1 to 261 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 262 to 373 of SEQ ID NO: 1;

l) the first glutamine synthetase fragment comprises amino acid residues 1 to 262 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 263 to 373 of SEQ ID NO: 1;

m) the first glutamine synthetase fragment comprises amino acid residues 1 to 301 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 302 to 373 of SEQ ID NO: 1;

n) the first glutamine synthetase fragment comprises amino acid residues 1 to 302 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 303 to 373 of SEQ ID NO: 1;

o) the first glutamine synthetase fragment comprises amino acid residues 1 to 310 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 311 to 373 of SEQ ID NO: 1; or

p) the first glutamine synthetase fragment comprises amino acid residues 1 to 311 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 312 to 373 of SEQ ID NO: 1.

8. The vector of claim 7, wherein

a) the first fragment of glutamine synthetase comprises amino acid residues 1 to 91 of SEQ ID NO: 1 and the second fragment of glutamine synthetase comprises amino acid residues 92 to 373 of SEQ ID NO:1;

b) the first fragment of glutamine synthetase comprises amino acid residues 1 to 92 of SEQ ID NO: 1 and the second fragment of glutamine synthetase comprises amino acid residues 93 to 373 of SEQ ID NO:1;

c) the first fragment of glutamine synthetase comprises amino acid residues 1 to 186 SEQ ID NO: 1 and the second fragment of glutamine synthetase comprises amino acid residues 187 to 373 of SEQ ID NO:1; or

d) the first fragment of glutamine synthetase comprises amino acid residues 1 to 187 SEQ ID NO: 1 and the second fragment of glutamine synthetase comprises amino acid residues 188 to 373 of SEQ ID NO:1.

9. The vector of any one of claims 1-8, wherein the heteromeric complex is an antibody, an antigen binding fragment of an antibody or a bispecific antigen binding molecule.

10. The vector of any one of claims 1-9, wherein:

a) the first nucleic acid encodes an antibody heavy chain or a fragment thereof and the third nucleic acid encodes an antibody light chain or a fragment thereof; or

b) the first nucleic acid encodes an antibody light chain or a fragment thereof and the third nucleic acid encodes an antibody heavy chain or a fragment thereof.

11. The vector of any one of claims 1-10, wherein the vector further comprises an internal ribosomal entry site (IRES) and/or expression augmenting sequence element (EASE).

12. The vector of claim 11, wherein the IRES occurs at a site selected from:

a) a site between the first nucleic acid and the second nucleic acid;

b) a site between the third nucleic acid and the fourth nucleic acid, and

c) at sites between both first and second, and third and fourth nucleic acids.

13. The vector of any one of claims 1-12, wherein the IRES comprises the sequence of GATGATAATACCCTCGAGATCCGTGCCATCATG.

14. The vector of any one of claims 1-13, wherein the first fragment of glutamine synthetase is fused to a first interaction domain, and the second fragment of glutamine synthetase is fused to a second interaction domain.

15. The vector of claim 14, wherein each of the first and second interaction domains is a leucine zipper or an anti-parallel leucine zipper polypeptide.

16. The vector of claim 14, wherein each of the first and second interaction domains comprise the amino acid sequence of ALKKELQANKKELAQLKWELQALKKELAQ or EQLEKKLQALEKKLAQLEWKNQALEKKLAQ.

17. The vector of any one of claims 14-16, wherein each of the first and second interaction domains further comprises a linker.

18. The vector of claim 17, wherein the linker comprises a sequence selected from GGPGG, GPGGG, GGGGSGGGGGS, GGGGS and GGGGSGGGGS.

19. The vector of claim 18, wherein the linker comprises the sequence of GGGGSGGGGS.

20. The vector of claim 17 or 18, wherein

a) the first interaction domain is fused to the N-terminal of the first glutamine synthetase fragment and has the amino acid sequence of EQLEKKLQALEKKLAQLEWKNQALEKKLAQGGGGSGGGGS and the second interaction domain is fused to the C-terminal of the second glutamine synthetase fragment and has the amino acid sequence of GGGGSGGGGSALKKELQANKKELAQLKWELQALKKELAQ; or

b) the first interaction domain is fused to the C-terminal of the first glutamine synthetase fragment and has the amino acid sequence of GGGGSGGGGSALKKELQANKKELAQLKWELQALKKELAQ and the second interaction domain is fused to the N-terminal of the second glutamine synthetase fragment and has the amino acid sequence of EQLEKKLQALEKKLAQLEWKNQALEKKLAQGGGGSGGGGS.

21. The vector of any one of claims 1-20, wherein the vector further comprises a fifth nucleic acid encoding a selectable marker selected from Zeomycin, neomycin, puromycin, Blasticidin S, and GPT.

22. A host cell comprising the vector of any one of claims 1-21.

23. The host cell of claim 22, wherein the host cell is CHO, VERO, BHK, HeLa, Cos, MDCK, 293, 3T3, WI338, or NSO cells.

24. The host cell of claim 22 or 23, wherein the host cell lacks endogenous glutamine synthetase.

25. A method of producing an antibody heavy chain or a fragment thereof and an antibody light chain or a fragment thereof comprising culturing the host cell of any one of claims 22-24 under conditions wherein the heteromeric complex is expressed by the host cell.

26. The method of claim 25, wherein the heteromeric complex is an antibody.

27. The method of claim 26, further comprising isolating the heteromeric complex.

28. An expression system comprising:

a) a first vector comprising a first nucleic acid encoding a first polypeptide, wherein the transcription of the first nucleic acid is operably linked to the transcription of a second nucleic acid encoding a first fragment of glutamine synthetase, and

b) a second vector comprising a third nucleic acid encoding a third polypeptide, wherein the transcription of the third nucleic acid is operably linked to the transcription of a fourth nucleic acid encoding a second fragment of glutamine synthetase,

wherein the first polypeptide is capable of associating with the third polypeptide to form a heteromeric complex,

wherein the first and second fragments of glutamine synthetase associate to provide a selectable activity, and

wherein the expression system is capable of being transfected into mammalian cells and improving selection of the cells.

29. The expression system of claim 28, wherein

a) the first fragment of glutamine synthetase is an N-terminal fragment of glutamine synthetase and the second fragment of glutamine synthetase is a C-terminal fragment of glutamine synthetase, or

b) the first fragment of glutamine synthetase is a C-terminal fragment of glutamine synthetase and the second fragment of glutamine synthetase is an N-terminal fragment of glutamine synthetase.

30. The expression system of claim 28 or 29, wherein the glutamine synthetase comprises the amino acid sequence of SEQ ID NO: 1, and the first and second fragments of glutamine synthetase are generated by splitting the glutamine synthetase polypeptide at an amino acid position selected from K52, E55, D92, G187, G245, R262, K291, G302 and D311 of SEQ ID NO: 1.

31. The expression system of claim 30, wherein the first and second fragments of glutamine synthetase are generated by splitting the glutamine synthetase polypeptide at an amino acid position selected from K52, E55, D92, G187 and G245 of SEQ ID NO: 1.

32. The expression system of claim 30, wherein the first and second fragments of glutamine synthetase are generated by splitting the glutamine synthetase polypeptide at the amino acid position D92 or G187 of SEQ ID NO: 1.

33. The expression system of claim 28 or 29, wherein the glutamine synthetase is a mammalian glutamine synthetase having an amino acid sequence different from SEQ ID NO:1, and the first and second fragments of glutamine synthetase are generated by splitting the glutamine synthetase polypeptide at an amino acid position equivalent to an amino acid position selected from K52, E55, D92, G187, G245, R262, K291, G302 and D311 of SEQ ID NO:1 according to sequence alignment.

34. The expression system of claim 30, wherein

a) the first glutamine synthetase fragment comprises amino acid residues 1 to 51 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 52 to 373 of SEQ ID NO: 1;

b) the first glutamine synthetase fragment comprises amino acid residues 1 to 52 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 53 to 373 of SEQ ID NO: 1;

c) the first glutamine synthetase fragment comprises amino acid residues 1 to 54 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 55 to 373 of SEQ ID NO: 1;

d) the first glutamine synthetase fragment comprises amino acid residues 1 to 55 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 56 to 373 of SEQ ID NO: 1;

e) the first glutamine synthetase fragment comprises amino acid residues 1 to 91 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 92 to 373 of SEQ ID NO: 1;

f) the first glutamine synthetase fragment comprises amino acid residues 1 to 92 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 93 to 373 of SEQ ID NO: 1;

g) the first glutamine synthetase fragment comprises amino acid residues 1 to 186 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 187 to 373 of SEQ ID NO: 1;

h) the first glutamine synthetase fragment comprises amino acid residues 1 to 187 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 188 to 373 of SEQ ID NO: 1;

i) the first glutamine synthetase fragment comprises amino acid residues 1 to 244 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 245 to 373 of SEQ ID NO: 1;

j) the first glutamine synthetase fragment comprises amino acid residues 1 to 245 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 246 to 373 of SEQ ID NO: 1**;**

k) the first glutamine synthetase fragment comprises amino acid residues 1 to 262 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 263 to 373 of SEQ ID NO: 1;

l) the first glutamine synthetase fragment comprises amino acid residues 1 to 263 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 264 to 373 of SEQ ID NO: 1;

m) the first glutamine synthetase fragment comprises amino acid residues 1 to 301 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 302 to 373 of SEQ ID NO: 1;

n) the first glutamine synthetase fragment comprises amino acid residues 1 to 302 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 303 to 373 of SEQ ID NO: 1;

o) the first glutamine synthetase fragment comprises amino acid residues 1 to 310 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 311 to 373 of SEQ ID NO: 1; or

p) the first glutamine synthetase fragment comprises amino acid residues 1 to 311 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 312 to 373 of SEQ ID NO: 1.

35. The expression system of claim 34, wherein

a) the first glutamine synthetase fragment comprises amino acid residues 1 to 91 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 92 to 373 of SEQ ID NO: 1;

b) the first glutamine synthetase fragment comprises amino acid residues 1 to 92 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 93 to 373 of SEQ ID NO: 1;

c) the first glutamine synthetase fragment comprises amino acid residues 1 to 186 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 187 to 373 of SEQ ID NO: 1; or

d) the first glutamine synthetase fragment comprises amino acid residues 1 to 187 of SEQ ID NO: 1 and the second glutamine synthetase fragment comprises amino acid residues 188 to 373 of SEQ ID NO: 1.

36. The expression system of any one of claims 28-35, wherein one or both the first and second vectors further comprise an IRES and/or an EASE.

37. The expression system of any one of claims 28-36, wherein the first fragment of glutamine synthetase is fused to a first interaction domain, and the second glutamine synthetase is fused to a second interaction domain.

38. The expression system of claim 37, wherein each of the first and second interaction domains is a leucine zipper or an anti-parallel leucine zippier polypeptide.

39. The expression system of claim 37 or 38, wherein each of the first and second interaction domains further comprises a linker having a sequence selected from GGPGG, GPGGG, GGGGSGGGGGS, GGGGS and GGGGSGGGGS.

40. The expression system of claims 39, wherein

a) The first interaction domain is fused to the N-terminal of the first glutamine synthetase fragment and has the amino acid sequence of GGGGSGGGGSALKKELQANKKELAQLKWELQALKKELAQ and the second interaction

domain is fused to the C-terminal of the second glutamine synthetase fragment and has the amino acid sequence of EQLEKKLQALEKKLAQLEWKNQALEKKLAQGGGGSGGGGS; or

b) The first interaction domain is fused to the C-terminal of the first glutamine synthetase fragment and has the amino acid sequence of EQLEKKLQALEKKLAQLEWKNQALEKKLAQGGGGSGGGGS and the second interaction domain is fused to the N-terminal of the second glutamine synthetase fragment and has the amino acid sequence of EQLEKKLQALEKKLAQLEWKNQALEKKLAQGGGGSGGGGS.

41. The expression system of any one of claims 28-40, wherein the heteromeric complex is an antibody, an antigen binding fragment of an antibody or a bispecific antigen binding molecule.

42. The expression system of any one of claims 28-41, wherein

a) the first polypeptide is a heavy chain of an antibody or a fragment thereof and the third polypeptide is a light chain of an antibody or a fragment thereof; or

b) the first polypeptide is a light chain of an antibody or a fragment thereof and the third polypeptide is a heavy chain of an antibody or a fragment thereof.

43. The expression system of any one of claims 28-42, wherein one or both the first and second vector further comprises an additional nucleic acid encoding a selectable marker selected from Zeomycin, neomycin, puromycin, Blasticidin S, and GPT.

44. A host cell comprising the expressing system of any one of claims 28-43.

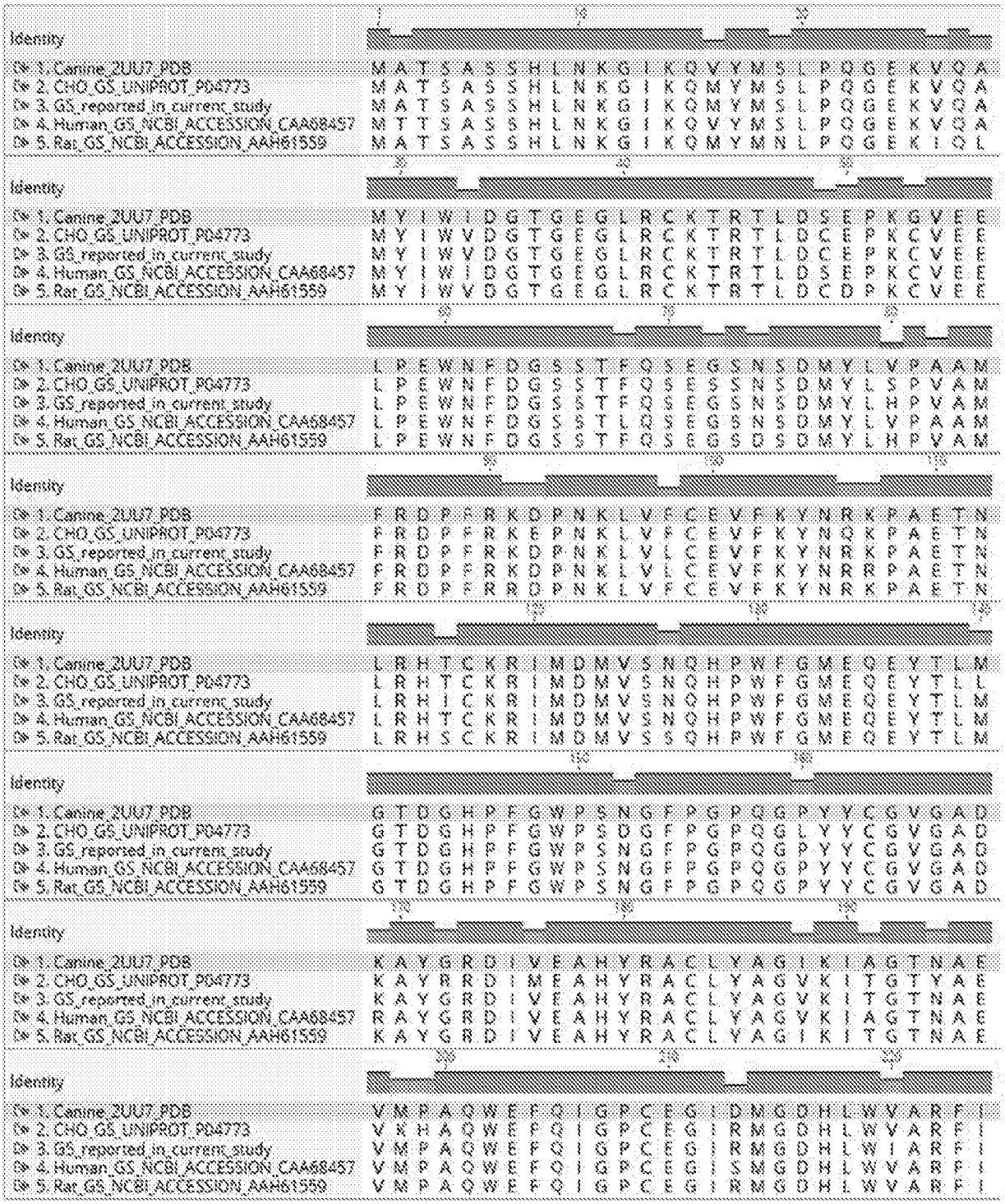
45. The host cell of claim 44, wherein the host cell is CHO, VERO, BHK, HeLa, Cos, MDCK, 293, 3T3, NSO, or WI38 cells.

46. The host cell of claim 44 or 45 wherein the host cell lacks endogenous glutamine synthetase.

47. A method of producing an antibody heavy chain or a fragment thereof and an antibody light chain or a fragment thereof comprising culturing the host cell of any one of claims 44-46 under conditions wherein the heteromeric complex is expressed by the host cell.

48. The method of claim 47, further comprising isolating the heteromeric complex.

Figure 1


Identity
1. Canine_2UU7_PDB
2. CHO_GS_UNIPROT_P04773
3. GS_reported_in_current_study
4. Human_GS_NCBI_ACCESSION_CAA68457
5. Rat_GS_NCBI_ACCESSION_AAH61559

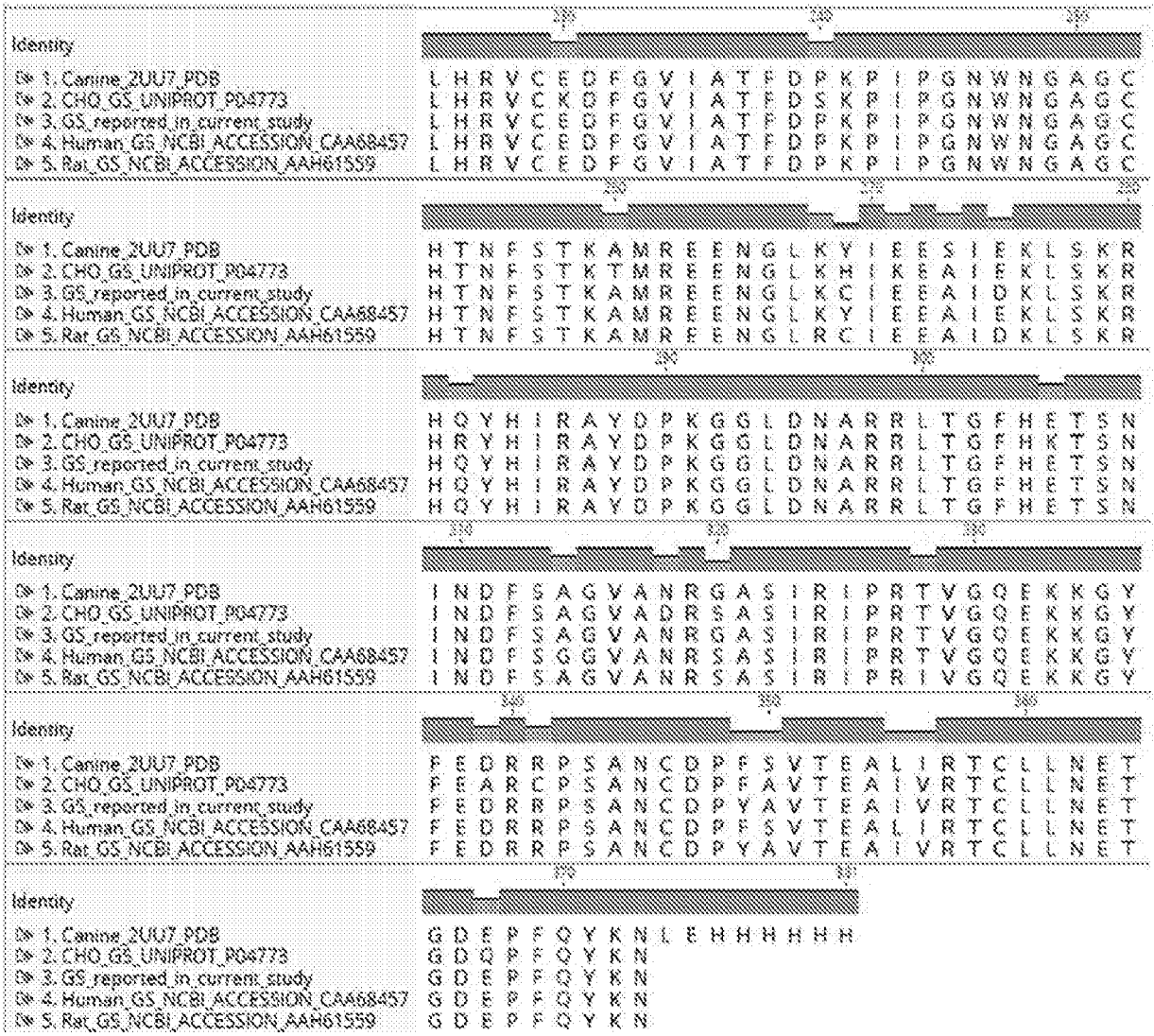

Identity
1. Canine_2UU7_PDB L H R V C E D F G V I A T F D P K P I P G N W N G A G C
2. CHO_GS_UNIPROT_P04773 L H R V C K D F G V I A T F D S K P I P G N W N G A G C
3. GS_reported_in_current_study L H R V C E D F G V I A T F D P K P I P G N W N G A G C
4. Human_GS_NCBI_ACCESSION_CAA68457 L H R V C E D F G V I A T F D P K P I P G N W N G A G C
5. Rat_GS_NCBI_ACCESSION_AAH61559 L H R V C E D F G V I A T F D P K P I P G N W N G A G C
Identity
1. Canine_2UU7_PDB H T N F S T K A M R E E N G L K Y I E E S I E K L S K R
2. CHO_GS_UNIPROT_P04773 H T N F S T K T M R E E N G L K H I K E A I E K L S K R
3. GS_reported_in_current_study H T N F S T K A M R E E N G L K C I E E A I D K L S K R
4. Human_GS_NCBI_ACCESSION_CAA68457 H T N F S T K A M R E E N G L K Y I E E A I E K L S K R
5. Rat_GS_NCBI_ACCESSION_AAH61559 H T N F S T K A M R E E N G L R C I E E A I D K L S K R
Identity
1. Canine_2UU7_PDB H Q Y H I R A Y D P K G G L D N A R R L T G F H E T S N
2. CHO_GS_UNIPROT_P04773 H R Y H I R A Y D P K G G L D N A R R L T G F H K T S N
3. GS_reported_in_current_study H Q Y H I R A Y D P K G G L D N A R R L T G F H E T S N
4. Human_GS_NCBI_ACCESSION_CAA68457 H Q Y H I R A Y D P K G G L D N A R R L T G F H E T S N
5. Rat_GS_NCBI_ACCESSION_AAH61559 H Q Y H I R A Y D P K G G L D N A R R L T G F H E T S N
Identity
1. Canine_2UU7_PDB I N D F S A G V A N R G A S I R I P R T V G Q E K K G Y
2. CHO_GS_UNIPROT_P04773 I N D F S A G V A D R S A S I R I P R T V G Q E K K G Y
3. GS_reported_in_current_study I N D F S A G V A N R G A S I R I P R T V G Q E K K G Y
4. Human_GS_NCBI_ACCESSION_CAA68457 I N D F S G G V A N R S A S I R I P R T V G Q E K K G Y
5. Rat_GS_NCBI_ACCESSION_AAH61559 I N D F S A G V A N R S A S I R I P R I V G Q E K K G Y
Identity
1. Canine_2UU7_PDB F E D R R P S A N C D P F S V T E A L I R T C L L N E T
2. CHO_GS_UNIPROT_P04773 F E A R C P S A N C D P F A V T E A I V R T C L L N E T
3. GS_reported_in_current_study F E D R R P S A N C D P Y A V T E A I V R T C L L N E T
4. Human_GS_NCBI_ACCESSION_CAA68457 F E D R R P S A N C D P F S V T E A L I R T C L L N E T
5. Rat_GS_NCBI_ACCESSION_AAH61559 F E D R R P S A N C D P Y A V T E A I V R T C L L N E T
Identity
1. Canine_2UU7_PDB G D E P F Q Y K N L E H H H H H H
2. CHO_GS_UNIPROT_P04773 G D Q P F Q Y K N
3. GS_reported_in_current_study G D E P F Q Y K N
4. Human_GS_NCBI_ACCESSION_CAA68457 G D E P F Q Y K N
5. Rat_GS_NCBI_ACCESSION_AAH61559 G D E P F Q Y K N

Figure 2

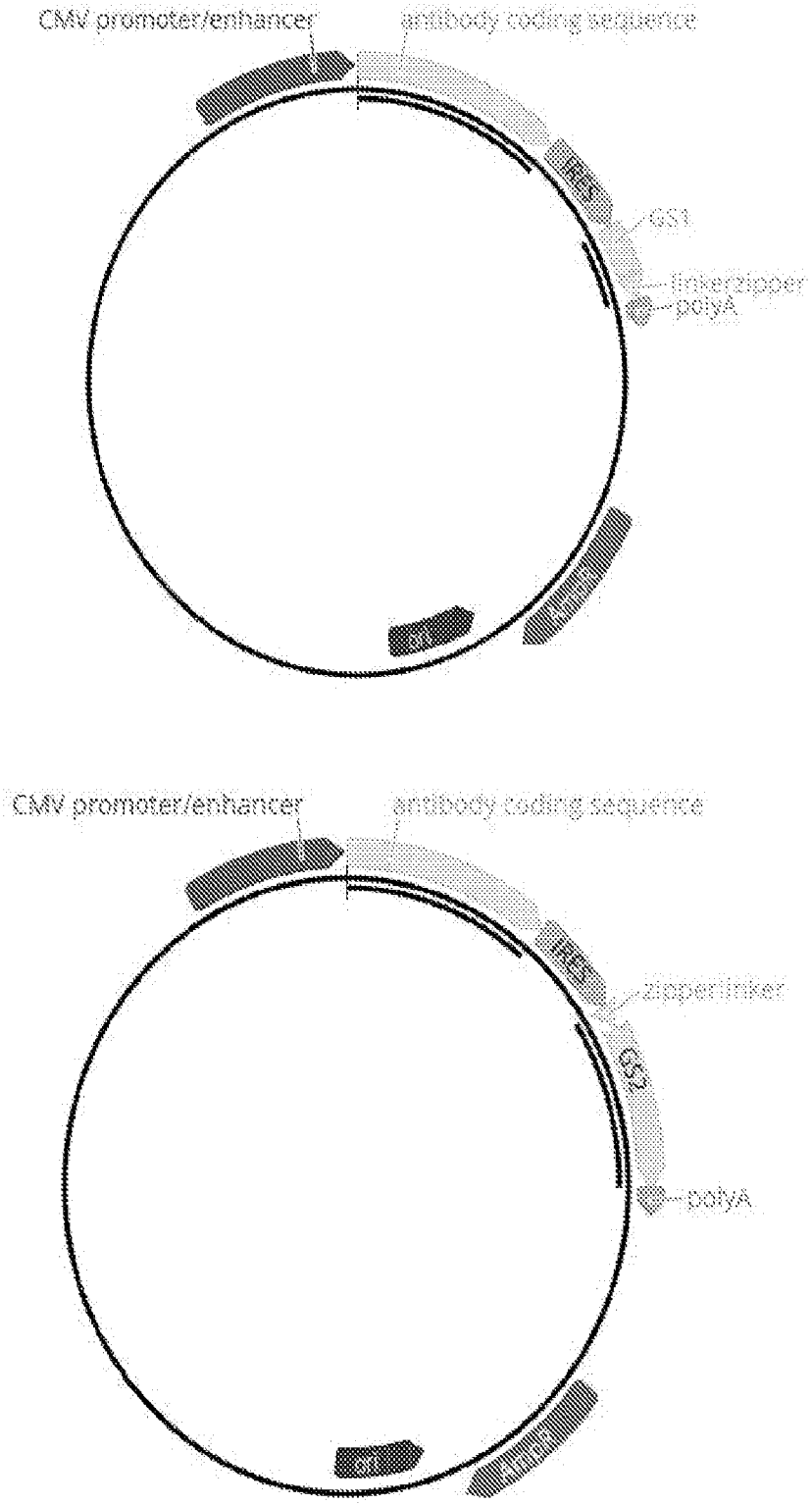

CMV promoter/enhancer
antibody coding sequence
Fc2
GS1
linker/zipper
polyA
CH1
CMV promoter/enhancer
antibody coding sequence
Fc2
zipper/linker
GS2
polyA
CH1

Figure 3

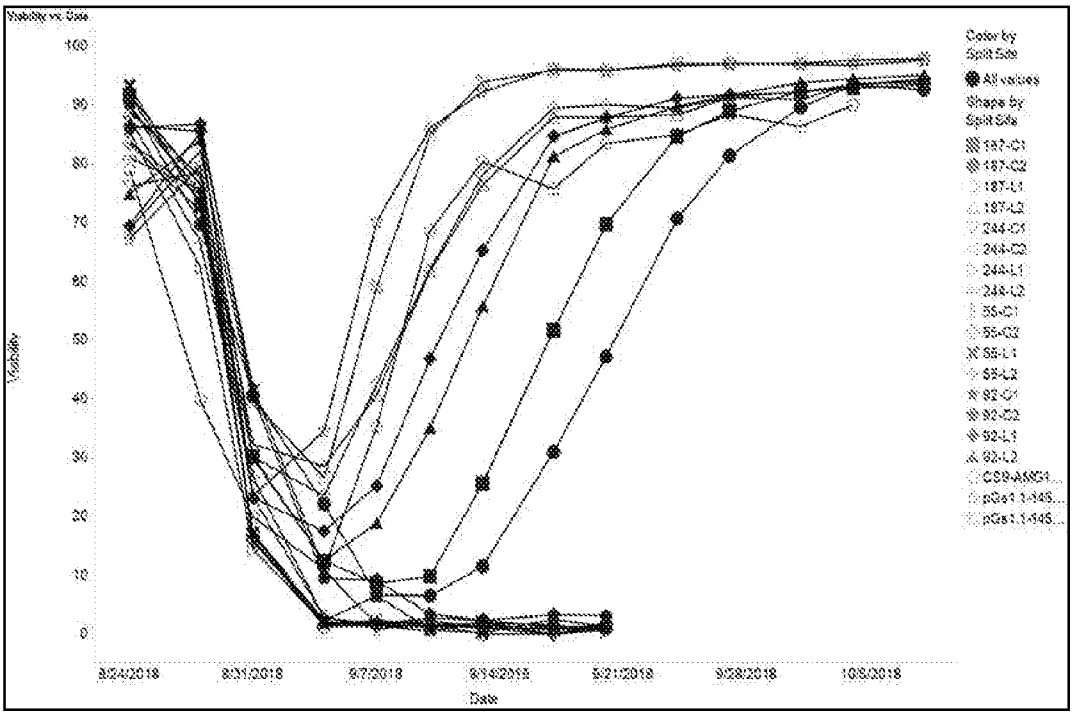
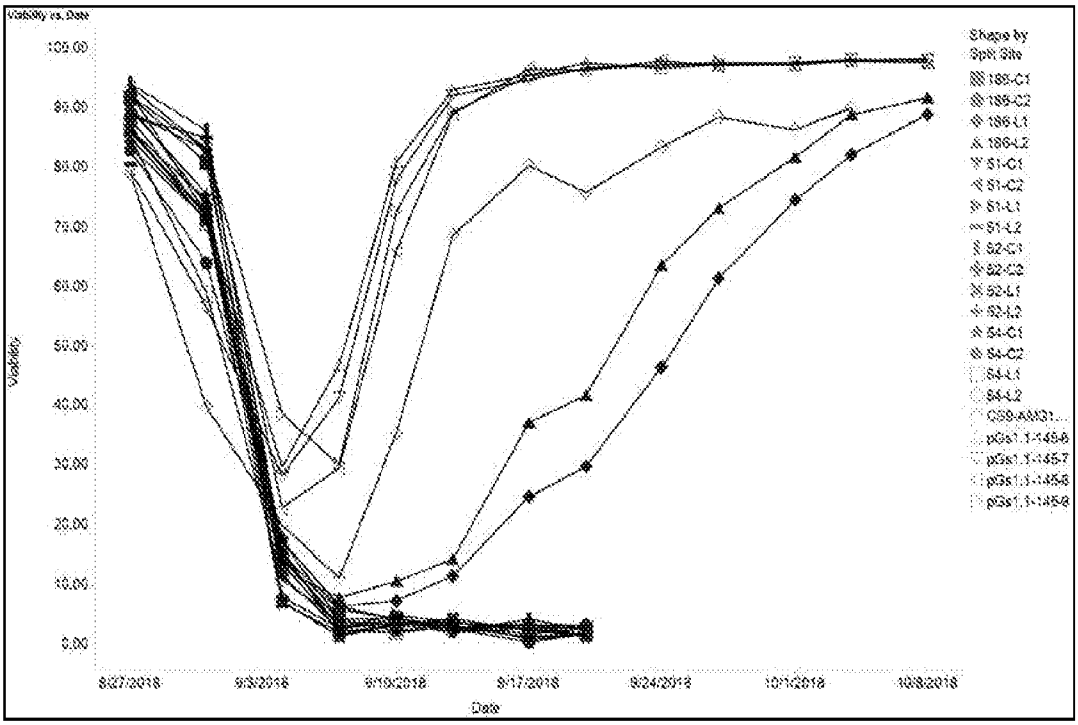



Figure 4



# INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2020/031309

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - C12N 15/52; C12N 15/62; C12N 15/67; C12N 15/79 (2020.01)
CPC - C07K 2317/10; C12N 15/52; C12N 15/62; C12N 15/67; C12N 15/79; C12Y 603/01002 (2020.05)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

see Search History document

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

see Search History document

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

see Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | US 2019/0127452 A1 (AMGEN INC.) 02 May 2019 (02.05.2019) entire document | 1, 2, 28, 29 |
| A | WO 2018/162517 A1 (F. HOFFMANN-LA ROCHE AG et al) 13 September 2018 (13.09.2018) entire document | 1, 2, 28, 29 |
| A | US 2015/0337331 A1 (JCR PHARMACEUTICALS CO., LTD.) 26 November 2015 (26.11.2015) entire document | 1, 2, 28, 29 |
| A | WO 2004/050879 A1 (BOEHRINGER INGELHEIM PHARMA GMBH & CO. KG et al) 17 June 2004 (17.06.2004) entire document | 1, 2, 28, 29 |
| A | WO 2003/035887 A1 (IMMUNEX CORPORATION et al) 01 May 2003 (01.05.2003) entire document | 1, 2, 28, 29 |

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

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

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 13 July 2020 | 04 AUG 2020 |
| Name and mailing address of the ISA/US<br>Mail Stop PCT, Attn: ISA/US, Commissioner for Patents<br>P.O. Box 1450, Alexandria, VA 22313-1450<br>Facsimile No. 571-273-8300 | Authorized officer<br>Blaine R. Copenheaver<br>Telephone No. PCT Helpdesk: 571-272-4300 |

# INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2020/031309

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

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

   a. ☐ forming part of the international application as filed:

      ☐ in the form of an Annex C/ST.25 text file.

      ☐ on paper or in the form of an image file.

   b. ☐ furnished together with the international application under PCT Rule 13*ter*.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.

   c. ☐ furnished subsequent to the international filing date for the purposes of international search only:

      ☐ in the form of an Annex C/ST.25 text file (Rule 13*ter*.1(a)).

      ☐ on paper or in the form of an image file (Rule 13*ter*.1(b) and Administrative Instructions, Section 713).

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

3. Additional comments:

ISA/225 was mailed on 15 May 2020. No approved electronic sequence listing was submitted in response to the ISA/225.

# INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2020/031309

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

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

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: 3-8, 30-35
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

Claims 3-8 and 30-35 are held unsearchable as a meaningful opinion could not be formed without the sequence listing; the applicant did not, within the prescribed time limit, furnish a sequence listing in the form of an Annex C/ST.25 text file, and such listing was not available to the International Searching Authority in the form and manner acceptable to it; or the sequence listing furnished did not comply with the standard provided for in Annex C of the Administrative Instructions.

3. ☒ Claims Nos.: 9-27, 36-48
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

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

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

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

**Remark on Protest**

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.