

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

(19) World Intellectual Property

Organization

International Bureau

(43) International Publication Date

30 March 2017 (30.03.2017)



(10) International Publication Number

WO 2017/053482 A1

(51) International Patent Classification:

C12N 5/10 (2006.01) C12N 15/79 (2006.01)

A61K 39/00 (2006.01) C12N 15/85 (2006.01)

C07K 16/00 (2006.01)

(21) International Application Number:

PCT/US2016/052962

(22) International Filing Date:

21 September 2016 (21.09.2016)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/222,187 22 September 2015 (22.09.2015) US

(71) Applicant (for all designated States except AL, AT, BE, BG, CH, CN, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IN, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR): **GENENTECH, INC.** [US/US]; 1 DNA Way, South San Francisco, CA 94080-4990 (US).

(71) Applicant (for AL, AT, BE, BG, CH, CN, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IN, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR only): **F. HOFFMANN-LA ROCHE AG** [CH/CH]; Grenzacherstrasse 124, 4070 Basel (CH).

(72) Inventors: **HALEY, Benjamin**; c/o Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080 (US). **HU, Zhilan**; c/o Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080 (US). **JOLY, John, C.**; c/o Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080 (US). **SHEN, Amy, Y.**; c/o Genentech, Inc., 1 DNA Way, South

San Francisco, CA 94080 (US). **SNEDECOR, Bradley, Richard**; c/o Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080 (US).

(74) Agents: **CHAPMAN, John D.** et al.; Morrison & Foerster LLP, 755 Page Mill Road, Palo Alto, CA 94304-1018 (US).

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

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

Published:

- with international search report (Art. 21(3))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: EXPRESSION OF FC-CONTAINING PROTEINS

(57) Abstract: The present invention provides compositions comprising a Fc-containing protein wherein substantially all the Fc domains have a C-terminal lysine. Further provided are host cell for producing said compositions, methods of making said host cells and compositions, and method of use thereof.

WO 2017/053482 A1

EXPRESSION OF FC-CONTAINING PROTEINS

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority to U.S. Provisional Patent Application No. 62/222,187, filed September 22, 2015, the disclosure of which is incorporated herein by reference in its entirety.

SUBMISSION OF SEQUENCE LISTING ON ASCII TEXT FILE

[0002] The content of the following submission on ASCII text file is incorporated herein by reference in its entirety: a computer readable form (CRF) of the Sequence Listing (file name: 146392032640SEQLIST.TXT, date recorded: August 18, 2016, size: 5 KB).

TECHNICAL FIELD

[0003] The present invention relates to compositions comprising a Fc-containing protein, wherein substantially all Fc-containing proteins have a C-terminal lysine on each Fc domain. The present invention further relates to host cells for producing said compositions, methods of making said host cells and compositions, and methods of use thereof.

BACKGROUND

[0004] Therapeutic Fc domain-containing proteins, *e.g.*, monoclonal antibodies and Fc-containing fusion proteins, have emerged as important drugs for the treatment of diseases, such as cancer, autoimmune disease, and infection, and currently represent the fastest growing class of therapeutic agents. *See, e.g.*, Beck et al., *Nat Rev Immunol*, 10, 2010. Clinically, polypeptide-based therapeutics have demonstrated potent activity due to their high degree of target specificity and ability to engage an immune response. *See, e.g.*, Sidhu, *Nat Biotechnol*, 25, 2007. However, commercial production of a homogeneous polypeptide-based therapeutic is challenging as polypeptides are subject to numerous modifications, for example, different disulfide pairing, deamidation, oxidation, N-terminal glutamine cyclization, and different glycan structures. *See, e.g.*, Carson, *Nat Biotechnol*, 23, 2005. Furthermore, the presence and structure consistency of polypeptide modifications are highly sensitive to production conditions. Inclusion of an immunoglobulin Fc domain with a

biologically active polypeptide provides beneficial biological and pharmacological properties to the biologically active agent, including, increased plasma half-life allowing for prolonged therapeutic activity and less frequent dosing. *See*, Czajkowski et al., *Mol Med*, 4, 2012. However, presence of a Fc domain further increases the potential for heterogeneity of a Fc-containing protein product, for example, variation of C-terminal lysine presence as compared across each Fc domain.

[0005] Industry standards for the production of biologics, such as Fc-containing proteins, require demonstration of product consistency across all clinical and commercial batches. Comparability studies of the polypeptide product must be conducted to ensure consistency following changes to production parameters, for example, those that occur during phase I to phase III production optimization, commercial scaling, and post-commercialization alterations. Production parameter changes include, but are not limited to, changes to the cell culture process parameters, cell culture media, host cell, purification, storage, and formulation. Furthermore, for multi-site production, it is necessary to ensure product consistency across all production facilities. Variability and heterogeneity in the levels of a polypeptide attribute, such as Fc domain C-terminal lysine presence, are considered to be an indication of a lack of production process control. *See*, Chirino et al., *Nat Biotechnol*, 22, 2004; and Kozlowski et al., *J Pharm Sci*, 97, 2006. Differences in cell lines and their sensitivity to manufacturing processes, *e.g.*, culture conditions, affect characteristics of the final product. For example, Fc domain C-terminal lysine processing, *i.e.*, cleavage, during production of a monoclonal antibody (consisting of two heavy chains), results in a mixture of antibody isoforms bearing zero, one, or two Fc domain C-terminal lysine residues. Extensive Fc domain C-terminal lysine processing can be observed in the production of an antibody, resulting in large percentages of the total antibody population bearing zero or one Fc domain C-terminal lysine. *See*, Dick et al., *Biotechnol Bioeng*, 100, 2008. Thus, the heterogeneity of Fc domain C-terminal lysine presence presents a unique challenge in the production of therapeutic Fc-containing proteins.

[0006] Extensive cell culture condition optimization has been investigated to decrease the heterogeneity of Fc domain C-terminal lysine presence that results during the production of Fc-containing proteins. For example, the effects of controlling cell media concentrations of heavy metals, such as zinc, cell media concentrations of amino acids, pH, and temperature on antibody Fc domain C-

terminal lysine presence has been described in U.S. Patent Application No. 20130280274. However, as demonstrated in U.S. Patent Application No. 20130280274, optimization of cell culture parameters does not eliminate the heterogeneity of Fc domain C-terminal lysine presence.

[0007] Fc domains of IgG1, IgG2, and IgG4 isotype immunoglobulins contain a conserved basic amino acid lysine at the C-terminus of each Fc domain.

Heterogeneity of Fc domain C-terminal lysine presence of these immunoglobulin isotypes is believed to result from proteolysis by endogenous carboxypeptidase(s) during cell culture production. *See, Harris, J Chromatogr A*, 705, 1995; and Luo et al., *Biotechnol Bioeng*, 109, 2012. Carboxypeptidases are enzymes that hydrolytically cleave amino acids from the C-termini of proteins and peptides. There are 13 known members of carboxypeptidases found in most mammalian species. *See, Reznik et al., Cell Mol Life Sci*, 58, 2001. While different carboxypeptidase proteins and their activities have been detected in CHO cells, a commonly used production host cell, it is not known which carboxypeptidase(s) is responsible for the removal of the Fc domain C-terminal lysine. *See, Dick et al., Biotechnol Bioeng*, 100, 2008. It has previously been speculated that one or more carboxypeptidases, including CpB and CpM, may be involved in Fc domain C-terminal lysine processing. *See, Dick et al., Biotechnol Bioeng*, 100, 2008.; Harris, *J Chromatogr A*, 705, 1995; and Luo et al., *Biotechnol Bioeng*, 109, 2012.

[0008] The disclosures of all publications, patents, patent applications and published patent applications referred to herein are hereby incorporated herein by reference in their entirety.

BRIEF SUMMARY OF THE INVENTION

[0009] Provided herein are host cells for expression of a Fc-containing protein, wherein the host cell has at least a 60% reduction of a carboxypeptidase D (CpD) expression level as compared to a host cell comprising a wild type CpD gene without CpD gene inactivation.

[0010] In some embodiments according to (or as applied to) any of the embodiments above, the host cell is a eukaryotic cell. In some embodiments, the host cell is a mammalian cell. In some embodiments, the host cell is a Chinese hamster ovary (CHO) cell.

[0011] In some embodiments according to (or as applied to) any of the embodiments above, the Fc-containing protein is an antibody. In some embodiments, the antibody is a monoclonal antibody.

[0012] In some embodiments according to (or as applied to) any of the embodiments above, the Fc-containing protein is a Fc-containing fusion protein.

[0013] In some embodiments according to (or as applied to) any of the embodiments above, the CpD gene in the host cell is inactivated. In some embodiments, the CpD gene is inactivated by siRNA. In some embodiments, the CpD gene is inactivated by shRNA. In some embodiments, the CpD gene is inactivated by gene deletion. In some embodiments, the CpD gene is inactivated by gene addition or substitution. In some embodiments, the CpD gene is inactivated using a clustered, regularly interspaced, short palindromic repeats (CRISPR) system. In some embodiments, the CpD gene is inactivated using a transcription activator-like effector nuclease (TALEN) system. In some embodiments, the CpD gene is inactivated using a zinc-finger nuclease (ZFN) system. In some embodiments, the CpD gene is inactivated using a meganuclease system.

[0014] In some embodiments according to (or as applied to) any of the embodiments above, the host cell comprises an expression vector comprising a nucleic acid encoding the Fc-containing protein.

[0015] Provided herein are cell culture systems comprising a host cell of any of the embodiments above.

[0016] Provided herein are methods of producing a host cell, wherein the methods comprise inactivating the CpD gene in the host cell, thereby producing the host cell of any of the embodiments above.

[0017] In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a siRNA system. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a shRNA system. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene by gene deletion. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene by gene addition or substitution. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a CRISPR system. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a TALEN system. In some embodiments, the method of producing a host cell comprises inactivating the CpD

gene using a ZFN system. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a meganuclease system.

[0018] Provided herein are methods of making a Fc-containing protein comprising: a) culturing a host cell, and b) obtaining the Fc-containing protein expressed by the host cell.

[0019] Provided herein are compositions comprising a plurality of Fc-containing proteins, wherein substantially all Fc-containing proteins in the composition have a C-terminal lysine on each Fc domain. In some embodiments, substantially all Fc-containing proteins of the composition have the same amino acid sequence. In some embodiments, the plurality of Fc-containing proteins are substantially homogeneous in charge state. In some embodiments, the Fc-containing protein is a Fc-containing fusion protein. In some embodiments, the Fc-containing protein is an antibody. In some embodiments, the antibody is a human antibody. In some embodiments, the antibody is a humanized antibody. In some embodiments, the antibody comprises two heavy chains, and wherein each heavy chain comprises a C-terminal lysine. In some embodiments, the Fc-containing protein is conjugated to a drug. In some embodiments, the Fc-containing proteins in the composition comprise an IgG1 Fc domain. In some embodiments, the Fc-containing proteins in the composition comprise an IgG2 Fc domain. In some embodiments, the Fc-containing proteins in the composition comprise an IgG4 Fc domain.

[0020] In some embodiments according to (or as applied to) any of the embodiments above, the composition comprising a plurality of Fc-containing proteins is a pharmaceutical composition. In some embodiments, the pharmaceutical composition comprising a plurality of Fc-containing proteins is a sterile pharmaceutical composition.

[0021] In some embodiments according to (or as applied to) any of the embodiments above, the composition comprising a plurality of Fc-containing proteins is a cell culture medium.

[0022] In some embodiments according to (or as applied to) any of the embodiments above, the composition comprising a plurality of Fc-containing proteins is a cell lysate.

[0023] In some embodiments according to (or as applied to) any of the embodiments above, the composition comprising a plurality of Fc-containing proteins is an eluate from a protein purification column.

[0024] Also provided herein are methods of treating a disease in an individual in need thereof comprising administering to the individual a pharmaceutical composition or a sterile pharmaceutical composition described in the embodiments above. In some embodiments, the method of treating a disease in an individual in need thereof comprises administering the composition parenterally. In some embodiments, the method of treating a disease in an individual in need thereof comprises administering the composition intravenously or subcutaneously. In some embodiments, the method of treating a disease in an individual in need thereof comprises administering the composition locally. In some embodiments, the method of treating a disease in an individual in need thereof comprises administering the composition topically.

[0025] These and other aspects and advantages of the present invention will become apparent from the subsequent detailed description and the appended claims. It is to be understood that one, some, or all of the properties of the various embodiments described herein may be combined to form other embodiments of the present invention.

BRIEF DESCRIPTION OF THE FIGURES

[0026] **FIGS. 1A-1B** shows qPCR mRNA analysis of endogenous carboxypeptidases in cell lines. Each error bar denotes 1 standard deviation. **FIG. 1A** shows relative carboxypeptidase B (CpB), carboxypeptidase E (CpE), carboxypeptidase M (CpM), carboxypeptidase N1 (CpN1), and carboxypeptidase D (CpD) mRNA expression levels in DP12 and CHOK1 host cells. Expression levels of mRNA for each carboxypeptidase were normalized to the housekeeping gene β -2-microglobulin (b2m). White bars represent measurements from the DP12 host and black bars represent measurements from the CHOK1 host. **FIG. 1B** shows relative carboxypeptidase mRNA expression levels in antibody-expressing cell lines. Expression level of mRNA for each carboxypeptidase was normalized to the housekeeping gene b2m. White bars represent measurements from the antibody-expressing cell line A and black bars represent measurements from the antibody-expressing cell line B.

[0027] **FIGS. 2A-2B** shows the level of carboxypeptidase mRNA expression following transfection with a carboxypeptidase specific siRNA construct in antibody-expressing cell line A (white bars) and antibody-expressing cell line B (black bars). Each error bar denotes 1 standard deviation. **FIG. 2A** shows relative CpD mRNA expression levels in antibody-expressing cell line A and antibody-expressing cell line

B. The grey bars represent the CpD mRNA level in the cell lines after scramble construct transfection for both cell line A and cell line B; the white bar represents the CpD mRNA level in cell line A after CpDi construct transfection; and the black bar represents CpD mRNA in cell line B after CpDi construct transfection. The expression level of CpD mRNA in each cell line was compared to the expression level of CpD in the respective cell line following transfection with a scramble construct.

FIG. 2B shows relative CpN mRNA expression levels in antibody-expressing cell line A and antibody-expressing cell line B. The grey bars represent the mRNA level in the cell lines after scramble construct transfection; the white bar represents the mRNA level in cell line A after CpNi construct transfection; and the black bar represents mRNA level in cell line B after CpNi construct transfection. The expression level of CpN mRNA in each cell line was compared to the expression level of CpN in the respective cell line following transfection with a scramble construct.

[0028] **FIG. 3** shows C-terminal lysine levels of an expressed antibody following transfection with a carboxypeptidase specific siRNA construct in antibody-expressing cell line A (black bars) and antibody-expressing cell line B (grey bars).

[0029] **FIGS. 4A-4B** shows schematic diagrams of wild type (WT) and knockout (KO) CpD alleles. **FIG. 4A** shows a schematic diagram of the wild type CpD allele. Two guide RNA (gRNA) sequences (gRNA 1 and gRNA 2) were used for targeting exon 2 and exon 21 of the CHO CpD gene, respectively, for gene knockout via CRISPR. The distance between the two gRNAs is about 46 kb. Forward and reverse PCR primers used to detect the WT CpD sequence aimed to be removed by CRISPR are denoted as WT.F and WT.R, respectively. **FIG. 4B** shows a schematic diagram of the KO CpD allele. Forward and reverse PCR primers used to detect the KO CpD allele are denoted as KO.F and KO.R, respectively.

[0030] **FIGS. 5A-5B** shows CpD expression levels in two CpD knockout clones from the antibody-expressing cell line B. **FIG. 5A** shows CpD RNA expression levels of two CpD KO clones and two CpD WT clones. **FIG. 5B** shows Western blot analysis of two CpD KO clones and two CpD WT clones.

[0031] **FIGS. 6A-6C** shows fed-batch shake flask evaluation data of two CpD KO clones and two CpD WT clones of antibody-producing cell line B. Each bar represents the average from two replicates. Each error bar denotes 1 standard deviation. **FIG. 6A** shows the titer of expressed antibody at day 14. **FIG. 6B** shows the specific productivity (Qp) in picograms (pg) per cell per day. **FIG. 6C** shows

integrated viable cell counts in 100 millions of cells-days per liter after 14 days of culture.

[0032] FIG. 7 shows C-terminal lysine levels in two CpD KO clones and two CpD WT clones of antibody-producing cell line B.

DETAILED DESCRIPTION OF THE INVENTION

[0033] The present application is based on the surprising finding that carboxypeptidase D (CpD), a zinc-binding enzyme that hydrolyzes single C-terminal amino acids of proteins, is the sole carboxypeptidase responsible for Fc domain C-terminal lysine processing (i.e., cleavage). Specifically, using mass spectrometry analysis, we showed that Fc domain C-terminal lysine cleavage was completely abolished in CpD knockout cells, thus demonstrating that CpD is the only endogenous carboxypeptidase that cleaves antibody Fc domain C-terminal lysines in CHO cells.

[0034] The present application thus provides a method of producing Fc-containing proteins with a C-terminal lysine on each Fc domain by using a host cell that has a reduced level of CpD expression. The products produced by such methods can lead to 100% homogeneity in Fc domain C-terminal lysine presence. Such a method eliminates the need to consider the effects of changing production conditions on the status of Fc domain C-terminal lysine presence, and thus greatly facilitates the manufacturing process of Fc-containing proteins. For example, use of such a method for the production of a Fc-containing protein eliminates the need to optimize cell culture conditions to minimize variability of Fc domain C-terminal lysine presence as compared across a plurality of production batches (e.g., comparison of a plurality of clinical and commercial batches). Additionally, for example, use of such a method for the production of a Fc-containing protein eliminates the need to optimize cell culture conditions to balance minimizing the variability of Fc domain C-terminal lysine presence with the variability and status of other characteristics of the Fc-containing protein (e.g., titer of Fc-containing protein and presence of post-translational modifications). Furthermore, such a method of producing Fc-containing proteins greatly facilitates the subsequent analysis of the Fc-containing proteins. For example, use of such a method for the production of a Fc-containing protein eliminates the need to incorporate additional steps to assay for charge variants of a plurality of Fc-containing proteins with heterogeneous Fc domain C-terminal lysine presence (such

as, e.g., a step involving incubation with CpB to remove all Fc domain C-terminal lysines).

[0035] Furthermore, retention of the Fc domain C-terminal lysine can inhibit subsequent alterations of Fc domain C-terminal amino acids during Fc-containing protein production. For example, the conserved Fc domain C-terminal amino acid sequence of IgG1, IgG2, and IgG3 immunoglobulins is proline-glycine-lysine. Cleavage of the Fc domain C-terminal lysine generates a terminal glycine, allowing peptidylglycine α -amidating monooxygenase to catalyze a two-step amidation reaction to remove the glycine residue and add an amide moiety to the proline in IgG1, IgG2, and IgG3 (activity of this enzyme is also dependent on the copper concentration in the culture media). Fc domain C-terminal proline amidation, and the extent thereof, can further generate charge variants within a population of Fc-containing proteins.

[0036] Thus, the present application, in one aspect provides a host cell for expression of a Fc-containing protein, wherein the host cell has a reduced level of carboxypeptidase D (CpD) expression. The present application, in another aspect, provides a cell culture system comprising the host cell. The present application, in yet another aspect, provides a method of producing a host cell, wherein the method comprises inactivating the CpD gene in the host cell, thereby producing the host cell. The present application, in yet another aspect, provides a method of making a Fc-containing protein comprising culturing the host cell and obtaining the Fc-containing protein expressed by the host cell. The present application, in yet another aspect, provides a composition comprising a plurality of Fc-containing proteins, wherein substantially all Fc-containing proteins in the composition have a C-terminal lysine on each Fc domain. The present application, in yet another aspect, provides a method of treating a disease in an individual in need thereof comprising administering to the individual the composition, wherein the composition is a pharmaceutical composition comprising a plurality of Fc-containing proteins having a C-terminal lysine on each Fc domain.

Definitions

[0037] "Fc domain," as used herein, refers to a Fc region of an immunoglobulin heavy chain or a C-terminal fragment thereof. The term includes wild type Fc domains and variant Fc domains. In some embodiments, the human IgG heavy chain

Fc domain extends from Cys226, or from Pro230, to the carboxyl-terminus of the heavy chain (amino acid number is according to the EU numbering system, also called the EU index, as described in Kabat et al., *Sequences of Proteins of Immunological Interest*, 5th Ed. Public Health Service, National Institutes of Health, Bethesda, MD, 1991). In some embodiments, the term includes a C-terminal fragment of an immunoglobulin heavy chain and one or more constant regions. In some embodiments, for IgG, the Fc domain may comprise immunoglobulin domains CH2 and CH3 and the hinge between CH1 and CH2.

[0038] “Fc-containing protein,” as used herein, refers to a protein (e.g., an antibody or a Fc-containing fusion protein) comprising a Fc domain. In some embodiments, the Fc-containing protein comprises one or more protein subunits. In some embodiments, the Fc-containing protein comprises one or more polypeptides.

[0039] As used herein, “Fc-containing fusion protein” refers to a protein comprising a Fc domain fused to at least one other heterologous protein unit or polypeptide.

[0040] The term “polypeptide” used herein refers to a single polypeptide chain.

[0041] The term “heavy chain” used herein refers to an immunoglobulin heavy chain.

[0042] The term “antibody” herein is used in the broadest sense and encompasses various antibody structures, including but not limited to monoclonal antibodies, polyclonal antibodies, multispecific antibodies (e.g., bispecific antibodies), and antibody fragments so long as they comprise a Fc domain.

[0043] The term “chimeric” antibody refers to an antibody in which a portion of the heavy and/or light chain is derived from a particular source or species, while the remainder of the heavy and/or light chain is derived from a different source or species.

[0044] A “human antibody” is one which possesses an amino acid sequence which corresponds to that of an antibody produced by a human or a human cell or derived from a non-human source that utilizes human antibody repertoires or other human antibody-encoding sequences. This definition of a human antibody specifically excludes a humanized antibody comprising non-human antigen-binding residues.

[0045] A “humanized” antibody refers to a chimeric antibody comprising amino acid residues from non-human hypervariable regions (HVRs) and amino acid residues from human framework regions (FRs). In certain embodiments, a humanized antibody

will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the HVRs (e.g., CDRs) correspond to those of a non-human antibody, and all or substantially all of the FRs correspond to those of a human antibody. A humanized antibody optionally may comprise at least a portion of an antibody constant region derived from a human antibody. A "humanized form" of an antibody, e.g., a non-human antibody, refers to an antibody that has undergone humanization.

[0046] The term "monoclonal antibody" as used herein refers to an antibody obtained from a population of substantially homogeneous antibodies, i.e., the individual antibodies comprising the population are identical and/or bind the same epitope, except for possible variant antibodies, e.g., containing naturally occurring mutations or arising during production of a monoclonal antibody preparation, such variants generally being present in minor amounts. In contrast to polyclonal antibody preparations, which typically include different antibodies directed against different determinants (epitopes), each monoclonal antibody of a monoclonal antibody preparation is directed against a single determinant on an antigen. Thus, the modifier "monoclonal" indicates the character of the antibody as being obtained from a substantially homogeneous population of antibodies, and is not to be construed as requiring production of the antibody by any particular method.

[0047] As used herein, the term "immunoadhesin" designates molecules which combine the binding specificity of a heterologous protein (an "adhesin") with the effector functions of immunoglobulin constant domains. Structurally, the immunoadhesins comprise a fusion of an amino acid sequence with a desired binding specificity, which amino acid sequence is other than the antigen recognition and binding site of an antibody (*i.e.*, is "heterologous" compared to a constant region of an antibody), and an immunoglobulin constant domain sequence (*e.g.*, C_H2 and/or C_H3 sequence of an IgG). Exemplary adhesin sequences include contiguous amino acid sequences that comprise a portion of a receptor or a ligand that binds to a protein of interest. Adhesin sequences can also be sequences that bind a protein of interest, but are not receptor or ligand sequences (*e.g.*, adhesin sequences in peptibodies). Such polypeptide sequences can be selected or identified by various methods, include phage display techniques and high throughput sorting methods. The immunoglobulin constant domain sequence in the immunoadhesin can be obtained from any immunoglobulin, such as IgG1, IgG2, IgG3, or IgG4 subtypes, IgA (including IgA1

and IgA2), IgE, IgD, or IgM.

[0048] “Host cell,” as used herein, refers to a cell capable of producing a protein or polypeptide product. For example, a host cell can produce a Fc-containing protein.

[0049] The term “individual” refers to a mammal and includes, but is not limited to, human, bovine, horse, feline, canine, rodent, or primate.

[0050] As used herein, “treatment” or “treating” is an approach for obtaining beneficial or desired results including clinical results. For purposes of this invention, beneficial or desired clinical results include, but are not limited to, one or more of the following: alleviating one or more symptoms resulting from the disease, diminishing the extent of the disease, stabilizing the disease (e.g., preventing or delaying the worsening of the disease), preventing or delaying the spread (e.g., metastasis) of the disease, preventing or delaying the recurrence of the disease, delay or slowing the progression of the disease, ameliorating the disease state, providing a remission (partial or total) of the disease, decreasing the dose of one or more other medications required to treat the disease, delaying the progression of the disease, increasing the quality of life, and/or prolonging survival. The methods of the invention contemplate any one or more of these aspects of treatment.

[0051] A “therapeutically effective amount” is at least the minimum concentration required to effect a measurable improvement of a particular disorder. A therapeutically effective amount herein may vary according to factors such as the disease state, age, sex, and weight of the patient, and the ability of the antibody to elicit a desired response in the individual. A therapeutically effective amount is also one in which any toxic or detrimental effects of the antibody are outweighed by the therapeutically beneficial effects. A “prophylactically effective amount” refers to an amount effective, at the dosages and for periods of time necessary, to achieve the desired prophylactic result. Typically, but not necessarily, since a prophylactic dose is used in subjects prior to or at the earlier stage of disease, the prophylactically effective amount can be less than the therapeutically effective amount.

[0052] It is understood that aspects and embodiments of the invention described herein include “consisting” and/or “consisting essentially of” aspects and embodiments.

[0053] Reference to “about” a value or parameter herein includes (and describes) variations that are directed to that value or parameter per se. For example, description referring to “about X” includes description of “X.”

[0054] As used herein and in the appended claims, the singular forms “a,” “or,” and “the” include plural referents unless the context clearly dictates otherwise.

Host cell

[0055] The present application provides a host cell for expression of a Fc-containing protein, wherein the host cell has a reduced level of carboxypeptidase D (CpD) expression.

[0056] Among the host cells that may be employed are eukaryotic cells, such as yeast or higher eukaryotic cells. Higher eukaryotic cells include insect cells and established cell lines of mammalian origin.

[0057] Examples of suitable mammalian host cell include the COS-7 line of monkey kidney cells (ATCC CRL 1651) (Gluzman *et al.*, *Cell*, 23, 1981), L cells, 293 cells, C127 cells, 3T3 cells (ATCC CCL 163), Chinese hamster ovary (CHO) cells, or their derivatives such as Veggie CHO and related cell lines which grow in serum-free media (Rasmussen *et al.*, *Cytotechnology*, 28, 1998), HeLa cells, BHK (ATCC CRL10) cell lines, and the CVI/EBNA cell line derived from the African green monkey kidney cell line CVI (ATCC CCL 70) as described by McMahan *et al.*, *EMBO J*, 10, 1991, human embryonic kidney cells such as 293, 293 EBNA, or MSR 293, human epidermal A431 cells, human Colo205 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from in vitro culture of primary tissue, primary explants, HL-60, U937, HaK, or Jurkat cells. Optionally, for example, mammalian cell lines such as HepG2/3B, KB, NIH 3T3 or S49, can be used as host cells.

[0058] In some embodiments, the host cell is a CHO cell. CHO cells are well known in the art. See, e.g., Xu *et al.*, *Nat Biotechnol*, 29, 2011. In some embodiments, the host cell is a DP12 host cell. In some embodiments, the host cell is a DUXB-11 derived DHFR-deficient DP12 host cell. In some embodiments, the host cell is a CHOK1 host cell. In some embodiments, the host cell is a DHFR-positive CHOK1 host cell. In some embodiments, the host cell is a CHOK1M host cell.

[0059] In some embodiments, the host cell is a mouse host cell. In some embodiments, the host cell is a Sp2/0 host cell. In some embodiments, the host cell is a NS0 host cell.

[0060] In some embodiments, the host cell is a hybridoma. In some embodiments, the hybridoma is an antibody-producing cell, wherein the antibody-producing cell is

collected from a host following immunization of the host with an antigen. In some embodiments, the antibody-producing cell is fused with a myeloma cell. In some embodiments, the host cell is a mouse myeloma-derived cell line.

[0061] Alternatively, the host cell can be a lower eukaryote such as yeast. Suitable yeasts include *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Kluyveromyces* strains, *Candida*, or any yeast strain capable of expressing heterologous polypeptides.

[0062] As used herein, “a reduced level of CpD expression,” refers to an at least 60% reduction of a CpD expression level, as compared to a host cell comprising a wild type CpD gene without CpD gene inactivation. In some embodiments, the host cell has at least a 60% reduction of a CpD expression level, as compared to a host cell prior to inactivation. In some embodiments, the CpD expression level is reduced by at least about 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 100%. In some embodiments, the CpD expression level is reduced by at least 95%. In some embodiments, the CpD expression level is reduced by 100%. In some embodiments, the CpD expression level is reduced by about 60% to 100%, about 70% to 100%, about 80% to 100%, about 90% to 100%, or about 95% to 100%.

[0063] In some embodiments, the CpD gene in the host cell is inactivated, wherein the reduction of a CpD expression level is based on a DNA level. In some embodiments, the CpD gene in the host cell is inactivated, wherein the reduction of a CpD expression level is based on a RNA level. In some embodiments, the CpD gene in the host cell is inactivated, wherein the reduction of a CpD expression level is based on a polypeptide level.

[0064] In some embodiments, the CpD gene in the host cell is inactivated, wherein the reduced expression level is based on a CpD expression level following gene inactivation. In some embodiments, the CpD gene in the host cell is inactivated, wherein the reduced expression level is based on a DNA level. In some embodiments, the CpD gene in the host cell is inactivated, wherein the reduced expression level is based on a RNA level. In some embodiments, the CpD gene in the host cell is inactivated, wherein the reduced expression level is based on a polypeptide level.

[0065] In some embodiments, the CpD gene in the host cell is inactivated, wherein the CpD expression level is reduced by at least 60%. In some embodiments, the CpD gene in the host cell is inactivated, wherein the CpD expression level is reduced by at least 90%. In some embodiments, the CpD gene in the host cell is inactivated,

wherein the CpD expression level is reduced by at least 95%. In some embodiments, the CpD gene in the host cell is inactivated, wherein the CpD expression level is reduced by 100%.

[0066] In some embodiments, the CpD gene in the host cell is inactivated. As used herein, “inactivated” refers to inhibiting the translation, or potential future translation, of a gene (i.e., expression of a protein). Inactivation can occur at any stage or process of gene expression, including, but not limited to, transcription, translation, and protein expression, and inactivation can affect any gene or gene product including, but not limited to, DNA, RNA, such as mRNA, and polypeptides.

[0067] Methods and techniques for inactivating the CpD gene in a host cell include, but are not limited to, small interfering RNA (siRNA), small hairpin RNA (shRNA; also referred to as a short hairpin RNA), clustered, regularly interspaced, short palindromic repeats (CRISPR), transcription activator-like effector nuclease (TALEN), zinc-finger nuclease (ZFN), homologous recombination, non-homologous end-joining, and meganuclease. *See, e.g., O’Keefe, Mater Methods*, 3, 2013; Doench *et al., Nat Biotechnol*, 32, 2014; Gaj *et al., Trends Biotechnol*, 31, 2014; and Silva *et al., Curr Gene Ther*, 11, 2011.

[0068] In some embodiments, the CpD gene is inactivated by a small interfering RNA (siRNA) system. Methods for identifying siRNA sequences suitable for CpD gene inactivation are well known in the art. For example, general consideration for developing and identifying siRNA to target the CpD gene include: a) first search sequences that are preferably 21-23 nucleotides in length (followed by reduction of sequence length as necessary), b) avoid regions within 50-100 base pairs of the start codon and the termination codon, c) avoid intron regions, d) avoid stretches of four or more bases, e.g., AAAA, e) avoid regions with GC content that is less than 30% or greater than 60%, f) avoid repeats and low sequence complexity, g) avoid single nucleotide polymorphism sites, and h) avoid sequences that are complementary to sequences in other off-target genes. *See, e.g., Rules of siRNA design for RNA interference*, Protocol Online, May 29, 2004; and Reynolds *et al., Nat Biotechnol*, 22, 2004.

[0069] In some embodiments, the siRNA system comprises a siRNA nucleotide sequence that is about 10 to 200 nucleotides in length, or about 10 to 100 nucleotides in length, or about 15 to 100 nucleotides in length, or about 10 to 60 nucleotides in length, or about 15 to 60 nucleotides in length, or about 10 to 50 nucleotides in length,

or about 15 to 50 nucleotides in length, or about 10 to 30 nucleotides in length, or about 15 to 30 nucleotides in length. In some embodiments, the siRNA nucleotide sequence is approximately 10-25 nucleotides in length. In some embodiments, the siRNA nucleotide sequence is approximately 15-25 nucleotides in length. In some embodiments, the siRNA nucleotide sequence is at least about 10, at least about 15, at least about 20, or at least about 25 nucleotides in length. In some embodiments, the siRNA system comprises a nucleotide sequence that is at least about 80%, at least about 85%, at least about 90%, at least about 95%, or 100% complementary to a region of a CpD mRNA molecule. In some embodiments, the siRNA system comprises a nucleotide sequence that is at least at least about 80%, at least about 85%, at least about 90%, at least about 95%, or 100% complementary to a region of a CpD pro-mRNA molecule. In some, embodiments, the siRNA system comprises a double stranded RNA molecule. In some embodiments, the siRNA system comprises a single stranded RNA molecule. In some embodiments, the host cell comprises a siRNA system as described in the any of the embodiments herein. In some embodiments, the host cell comprises a pro-siRNA nucleotide sequence that is processed into an active siRNA molecule as described in the any of the embodiments herein. In some embodiments, the host cell comprises a siRNA nucleotide sequence that is at least about 80%, at least about 85%, at least about 90%, at least about 95%, or 100% complementary to a region of a CpD mRNA molecule. In some embodiments, the host cell comprises an expression vector encoding a siRNA molecule as described in the any of the embodiments herein. In some embodiments, the host cell comprises an expression vector encoding a pro-siRNA molecule as described in the any of the embodiments herein.

[0070] In some embodiments, the siRNA system comprises a delivery vector. In some embodiments, the host cell comprises a delivery vector. In some embodiments, the delivery vector comprises the pro-siRNA and/or siRNA molecule.

[0071] Exemplary CpD siRNA target sequences are listed in Table 1.

Table 1. Exemplary CpD siRNA target sequences siRNA.

SEQ ID NO.:	CpD siRNA nucleotide sequence:
19	GGA AGA GAA CTG CTA CTA A

[0072] In some embodiments, the CpD gene is inactivated by a small hairpin RNA (shRNA; also referred to as a short hairpin RNA) system. Gene inactivation by shRNA systems are well known in the art. In some embodiments, the shRNA system comprises a nucleotide sequence that is about 10 to 200 nucleotides in length, or about 10 to 100 nucleotides in length, or about 15 to 100 nucleotides in length, or about 10 to 60 nucleotides in length, or about 15 to 60 nucleotides in length, or about 10 to 50 nucleotides in length, or about 15 to 50 nucleotides in length, or about 10 to 30 nucleotides in length, or about 15 to 30 nucleotides in length. In some embodiments, the shRNA nucleotide sequence is approximately 10-25 nucleotides in length. In some embodiments, the shRNA nucleotide sequence is approximately 15-25 nucleotides in length. In some embodiments, the shRNA nucleotide sequence is at least about 10, at least about 15, at least about 20, or at least about 25 nucleotides in length. In some embodiments, the shRNA system comprises a nucleotide sequence that is at least about 80%, at least about 85%, at least about 90%, at least about 95%, or 100% complementary to a region of a CpD mRNA molecule. In some embodiments, the shRNA system comprises a nucleotide sequence that is at least about 80%, at least about 85%, at least about 90%, at least about 95%, or 100% complementary to a region of a CpD pro-mRNA molecule. In some, embodiments, the shRNA system comprises a double stranded RNA molecule. In some embodiments, the shRNA system comprises a single stranded RNA molecule. In some embodiments, the host cell comprises a shRNA system as described in the any of the embodiments herein. In some embodiments, the host cell comprises a pre-shRNA nucleotide sequence that is processed in an active shRNA nucleotide sequence as described in any of the embodiments herein. In some embodiments, the pro-shRNA molecule composed of DNA. In some embodiments, the pro-shRNA molecule is a DNA construct. In some embodiments, the host cell comprises a shRNA nucleotide sequence that is at least about 80%, at least about 85%, at least about 90%, at least about 95%, or 100% complementary to a region of a CpD mRNA molecule. In some embodiments, the host cell comprises an expression vector encoding a shRNA molecule as described in the any of the embodiments herein. In some embodiments, the host cell comprises an expression vector encoding a pro-shRNA molecule as described in the any of the embodiments herein.

[0073] In some embodiments, the shRNA system comprises a delivery vector. In some embodiments, the host comprises a delivery vector. In some embodiments, the delivery vector comprises the pro-shRNA and/or shRNA molecule.

[0074] In some embodiments, the CpD gene is inactivated by a gene deletion. As used herein, “gene deletion” refers to removal of at least a portion of a DNA sequence from, or in proximity to, a gene. In some embodiments, the sequence subjected to gene deletion comprises an exonic sequence of a gene. In some embodiments, the sequence subjected to gene deletion comprises a promoter sequence of a gene. In some embodiments, the sequence subjected to gene deletion comprises a flanking sequence of a gene. In some embodiments, a portion of a gene sequence is removed from a gene. In some embodiments, a portion of the CpD gene sequence is removed from, or in proximity to, the CpD gene. In some embodiments, the complete gene sequence is removed from a chromosome. In some embodiments, the complete CpD sequence is removed from a chromosome. In some embodiments, the host cell comprises a gene deletion as described in the any of the embodiments herein. In some embodiments, the host cell comprises a gene deletion in the CpD gene. In some embodiments, the host cell comprises a gene deletion in proximity to the CpD gene.

[0075] In some embodiments, the CpD gene is inactivated by a gene addition or substitution. As used herein, “gene addition” or “gene substitution” refers to an alteration of a gene sequence, including insertion or substitution of one or more nucleotides or nucleotide base pairs. In some embodiments, the intronic sequence of the gene is altered. In some embodiments, the exonic sequence of the gene is altered. In some embodiments, the promoter sequence of the gene is altered. In some embodiments, the flanking sequence of the gene is altered. In some embodiments, one nucleotide or nucleotide base pair is added to a gene sequence. In some embodiments, at least one consecutive nucleotide or nucleotide base pair is added to a gene sequence. In some embodiments, the host cell comprises a gene addition or substitution as described in the any of the embodiments herein. In some embodiments, the host cell comprises a gene addition or substitution in the CpD gene.

[0076] In some embodiments, the CpD gene is inactivated by a gene deletion, wherein deletion of at least one nucleotide or nucleotide base pair in a gene sequence results in a non-functional gene product. In some embodiments, the CpD gene is inactivated by a gene deletion, wherein deletion of at least one nucleotide to a gene sequence results in a gene product that no longer has the original gene product

function or activity. In some embodiments, the CpD gene is inactivated by a gene deletion, wherein deletion of at least one nucleotide to a gene sequence results in a dysfunctional gene product.

[0077] In some embodiments, the CpD gene is inactivated by a gene addition or substitution, wherein addition or substitution of at least one nucleotide or nucleotide base pair into the CpD gene sequence results in a non-functional gene product. In some embodiments, the CpD gene is inactivated by a gene inactivation, wherein incorporation or substitution of at least one nucleotide to the CpD gene sequence results in a gene product that no longer has the original gene product function or activity. In some embodiments, the CpD gene is inactivated by a gene addition or substitution, wherein incorporation or substitution of at least one nucleotide into the CpD gene sequence results in a dysfunctional gene product.

[0078] In some embodiments, the host cell comprises a non-functional CpD gene product. In some embodiments, the host cell comprises a CpD gene product that does not have the original CpD gene product function or activity. In some embodiments, the host cell comprises a dysfunctional CpD gene product.

[0079] In some embodiments, the host cell comprises an inactivated CpD gene, wherein the inactivated CpD gene will not express a full length, and functional, CpD gene product (e.g., a full length CpD polypeptide sequence). In some embodiments, the host cell comprises an inactivated CpD gene, wherein the inactivated CpD gene will not express an endogenous CpD gene product sequence. In some embodiments, the host cell comprises an inactivated CpD gene, wherein the inactivated CpD gene will express a variant CpD gene product. In some embodiments, the host cell comprises a variant CpD gene product.

[0080] In some embodiments, the host cell comprises a delivery vector. In some embodiments, the delivery vector is a virus vector. In some embodiments, the delivery vector is a lentivirus. In some embodiments, the delivery vector is an adenovirus. In some embodiments, the vector comprises a promoter.

[0081] In some embodiments, the host cell is a stable knockdown host cell. In some embodiments, the host cell is a stable CpD knockdown cell line. In some embodiments, the host cell is a transient knockdown cell line. In some embodiments, the host cell is a transient CpD knockdown cell line.

[0082] In some embodiments, the host cell further comprises an inactivated gene other than CpD.

[0083] In general, host cells are transformed with a recombinant expression vector that comprises DNA encoding a desired Fc-containing protein. Additionally, the host cells of the present application can be a blank host cell. As used herein, “blank host” refers to a cell that does not contain an expression vector encoding a Fc-containing protein. In some embodiments, the blank host cell is a CHO cell. In some embodiments, the blank host cell is a mouse cell.

[0084] Also provided by the present application are host cells comprising nucleic acids encoding Fc-containing proteins described herein. Nucleic acid molecules provided by the invention include DNA and RNA in both single-stranded and double-stranded form, as well as the corresponding complementary sequences. DNA includes, for example, cDNA, genomic DNA, chemically synthesized DNA, DNA amplified by PCR, and combinations thereof. The nucleic acid molecules of the invention include full-length genes or cDNA molecules as well as a combination of fragments thereof. The nucleic acids provided herein are preferentially derived from human sources.

[0085] As noted elsewhere herein, an “isolated” nucleic acid is a nucleic acid that has been separated from adjacent genetic sequences present in the genome of the organism from which the nucleic acid was isolated, in the case of nucleic acids isolated from naturally-occurring sources. In the case of nucleic acids synthesized enzymatically from a template or chemically, such as PCR products, cDNA molecules, or oligonucleotides for example, it is understood that the nucleic acids resulting from such processes are isolated nucleic acids. An isolated nucleic acid molecule refers to a nucleic acid molecule in the form of a separate fragment or as a component of a larger nucleic acid construct.

[0086] In certain embodiments, the nucleic acids are substantially free from contaminating endogenous material. The nucleic acid molecule has preferably been derived from DNA or RNA isolated at least once in substantially pure form and in a quantity or concentration enabling identification, manipulation, and recovery of its component nucleotide sequences by standard biochemical methods (such as those outlined in Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989)). Such sequences are preferably provided and/or constructed in the form of an open reading frame uninterrupted by internal non-translated sequences, or introns, that are typically present in eukaryotic genes. Sequences of nontranslated DNA can be present 5' or 3'

from an open reading frame, where the same do not interfere with manipulation or expression of the coding region.

[0087] In some embodiments, the host cell is capable of expressing a Fc-containing protein. In some embodiments, the host cell comprises a Fc-containing protein. In some embodiments, the host cell is capable of secreting a Fc-containing protein.

[0088] In some embodiments, the host cell is capable of expressing a Fc-containing protein at a similar output rate of the host cell prior to CpD gene inactivation. In some embodiments, the host cell is capable of expressing a Fc-containing protein at the same output rate of the host cell prior to CpD gene inactivation. In some embodiments, the host cell is capable of expressing a Fc-containing protein at about 70%, 75%, 80%, 85%, 90%, 95%, or 100% of the output rate of the host cell prior to CpD gene inactivation.

[0089] In some embodiments, the host cell further comprises a gene modification. Optimization of a host cell for the purposes of producing a Fc-containing protein via gene modification is well known in the art and includes considerations pertaining to, for example, glycosylation patterns, vector selection properties for integration methods, and any other cellular property that would be desirable to manipulate for the production of a Fc-containing protein. In some embodiments, the gene modification is a targeted gene modification. In some embodiments, the gene modification is a knockout gene modification. In some embodiments, the gene modification is a knock-in gene modification.

Fc-containing protein

[0090] The present application provides a Fc-containing protein, wherein each Fc domain of the Fc-containing protein has a C-terminal lysine.

[0091] In some embodiments, the Fc-containing protein comprises a Fc domain. In some embodiments, the Fc-containing protein comprises one or more Fc domains. In some embodiments, the Fc-containing protein comprises two Fc domains.

[0092] In some embodiments, the Fc-containing protein comprises a heavy chain. In some embodiments, the Fc-containing protein comprises at least one heavy chain. In some embodiments, the Fc-containing protein comprises one or more heavy chains. In some embodiments, the Fc-containing protein comprises two heavy chains.

[0093] In some embodiments, the Fc-containing protein comprises a Fc domain, wherein the Fc domain comprises a C-terminal lysine. In some embodiments, the Fc-

containing protein comprises at least one Fc domain, wherein each Fc domain comprises a C-terminal lysine. In some embodiments, the Fc-containing protein comprises one or more Fc domain, wherein each Fc domain comprises a C-terminal lysine. In some embodiments, the Fc-containing protein comprises two Fc domains, wherein each Fc domain comprises a C-terminal lysine.

[0094] In some embodiments, the Fc-containing protein comprises a heavy chain, wherein the heavy chain comprises a C-terminal lysine. In some embodiments, the Fc-containing protein comprises at least one heavy chain, wherein each heavy chain comprises a C-terminal lysine. In some embodiments, the Fc-containing protein comprises one or more heavy chains, wherein each heavy chain comprises a C-terminal lysine. In some embodiments, the Fc-containing protein comprises two heavy chains, wherein each heavy chain comprises a C-terminal lysine.

[0095] In some embodiments, the Fc-containing protein is a multimeric protein.

[0096] In some embodiments, the Fc-containing protein is an antibody. In some embodiments, the antibody is a human antibody. In some embodiments, the antibody is a humanized antibody. In some embodiments, the antibody is a monoclonal antibody. In some embodiments, the antibody is a chimeric antibody. In some embodiments, the antibody is a bispecific antibody. In some embodiments, the antibody is a multispecific antibody.

[0097] In some embodiments, the Fc-containing protein is a Fc-containing fusion protein. In some embodiments, the Fc-containing fusion protein comprises one or more Fc domains. In some embodiments, the Fc-containing fusion protein comprises one or more Fc domains, each comprising a C-terminal lysine.

[0098] In some embodiments, the Fc domain of the Fc-containing fusion protein prolongs a plasma half-life of the Fc-containing fusion protein. In some embodiments, the Fc domain of the Fc-containing fusion protein prolongs the biological activity of the Fc-containing fusion protein. In some embodiments, the Fc domain of the Fc-containing fusion protein decreases the rate of renal clearance of the Fc-containing fusion protein. In some embodiments, the Fc domain of the Fc-containing fusion protein increases the solubility of the Fc-containing fusion protein. In some embodiments, the Fc domain of the Fc-containing fusion protein increases the stability of the Fc-containing fusion protein.

[0099] Fc-containing fusion proteins are well known in the art. *See*, e.g., Czajkowski *et al.*, *EMBO Mol Med*, 4, 2012. In some embodiments, the Fc-

containing fusion protein is an immunoadhesin. In some embodiments, the Fc-containing fusion protein is a cytokine-Fc fusion protein.

[0100] In some embodiments, the Fc-containing protein is conjugated to an agent. In some embodiments, the Fc-containing protein is conjugated to at least about 2, 3, 4, 5, 6, 7, 8, 9, or 10 molecules of an agent. In some embodiments, the Fc-containing protein is conjugated to about 2-10, about 4-10, about 6-10, or about 8-10 molecules of an agent. In some embodiments, the Fc-containing protein is conjugated to an agent, wherein the agent is conjugated to the Fc domain of the Fc-containing protein. In some embodiments, the agent is a therapeutic agent. In some embodiments, the therapeutic agent is a small molecule therapeutic agent. In some embodiments, the therapeutic agent is a chemotherapeutic agent. In some embodiments, the agent is a detection agent. In some embodiments, the detection agent is a radiolabel. In some embodiments, the detection agent is a fluorescent label. In some embodiments, the detection agent is an immunolabel. In some embodiments, the Fc-containing protein is a companion diagnostic.

[0101] In some embodiments, the Fc-containing protein comprises an IgG1 Fc domain. In some embodiments, the Fc-containing protein comprises an IgG1 Fc domain comprising a C-terminal lysine. In some embodiments, the Fc-containing protein comprises one or more IgG1 Fc domains. In some embodiments, the Fc-containing protein comprises one or more IgG1 Fc domains, each Fc domain comprising a C-terminal lysine. In some embodiments, the Fc-containing protein comprises an IgG1 heavy chain. In some embodiments, the Fc-containing protein comprises an IgG1 heavy chain comprising a C-terminal lysine. In some embodiments, the Fc-containing protein comprises one or more IgG1 heavy chains. In some embodiments, the Fc-containing protein comprises one or more IgG1 heavy chains, each heavy chain comprising a C-terminal lysine.

[0102] In some embodiments, the Fc-containing protein comprises an IgG2 Fc domain. In some embodiments, the Fc-containing protein comprises an IgG2 Fc domain comprising a C-terminal lysine. In some embodiments, the Fc-containing protein comprises one or more IgG2 Fc domains. In some embodiments, the Fc-containing protein comprises one or more IgG2 Fc domains, each Fc domain comprising a C-terminal lysine. In some embodiments, the Fc-containing protein comprises an IgG2 heavy chain. In some embodiments, the Fc-containing protein comprises an IgG2 heavy chain comprising a C-terminal lysine. In some

embodiments, the Fc-containing protein comprises one or more IgG2 heavy chains. In some embodiments, the Fc-containing protein comprises one or more IgG2 heavy chains, each heavy chain comprising a C-terminal lysine.

[0103] In some embodiments, the Fc-containing protein comprises an IgG4 Fc domain. In some embodiments, the Fc-containing protein comprises an IgG4 Fc domain comprising a C-terminal lysine. In some embodiments, the Fc-containing protein comprises one or more IgG4 Fc domains. In some embodiments, the Fc-containing protein comprises one or more IgG4 Fc domains, each Fc domain comprising a C-terminal lysine. In some embodiments, the Fc-containing protein comprises an IgG4 heavy chain. In some embodiments, the Fc-containing protein comprises an IgG4 heavy chain comprising a C-terminal lysine. In some embodiments, the Fc-containing protein comprises one or more IgG4 heavy chains. In some embodiments, the Fc-containing protein comprises one or more IgG4 heavy chains, each heavy chain comprising a C-terminal lysine.

[0104] In some embodiments, the Fc-containing protein comprises a post-translational modification. In some embodiments, the post-translational modification is non-enzymatically produced. In some embodiments, the post-translational modification is enzymatically produced. In some embodiments, the post-translational modification is selected from the group consisting of a disulfide pairing, a deamidation, an oxidation, a N-terminal glutamine cyclization, and a glycosylation.

Compositions comprising a plurality Fc-containing proteins

[0105] The present application, in another aspect, provides compositions comprising a plurality of Fc-containing proteins, wherein substantially all Fc-containing proteins of the plurality of Fc-containing proteins has a C-terminal lysine on each Fc domain.

[0106] As used herein, the term “substantially all” refers to at least about 90%, including for example, 95%, or 100%. Thus, for example, the composition comprises a plurality of Fc-containing proteins, wherein at least about 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% of the Fc-containing proteins of the plurality of Fc-containing proteins have a C-terminal lysine on each Fc domain. In some embodiments, the composition comprises a plurality of Fc-containing proteins, wherein 100% of the Fc-containing proteins of the plurality of Fc-containing proteins have a C-terminal lysine on each Fc domain.

[0107] In some embodiments, the composition comprises a plurality of Fc-containing proteins, wherein about 90% to 100% of the Fc-containing proteins of the plurality of Fc-containing proteins have a C-terminal lysine on each Fc domain.

[0108] In some embodiments, the composition comprises a plurality of Fc-containing proteins, wherein 100% of the Fc-containing proteins of the plurality of Fc-containing proteins has a C-terminal lysine on each Fc domain.

[0109] In some embodiments, the composition comprises a plurality of Fc-containing proteins, wherein substantially all of the Fc-containing proteins of the plurality of Fc-containing proteins have the same amino acid sequence. In some embodiments, the composition comprises a plurality of Fc-containing proteins, wherein substantially all of the Fc-containing proteins of the plurality of Fc-containing proteins have the same Fc domain amino acid sequence. In some embodiments, the composition comprises a plurality of Fc-containing proteins, wherein substantially all of the Fc-containing proteins of the plurality of Fc-containing proteins have the same heavy chain amino acid sequence.

[0110] In some embodiments, the composition comprises a plurality of Fc-containing proteins, wherein at least about 80%, 85%, 90%, 95%, or 100% of the Fc-containing proteins of the plurality of Fc-containing proteins have the same amino acid sequence. In some embodiments, the composition comprises a plurality of Fc-containing proteins, wherein at least about 80%, 85%, 90%, 95%, or 100% of the Fc-containing proteins of the plurality of Fc-containing proteins have the same Fc domain amino acid sequence. In some embodiments, the composition comprises a plurality of Fc-containing proteins, wherein at least about 80%, 85%, 90%, 95%, or 100% of the Fc-containing proteins of the plurality of Fc-containing proteins have the same heavy chain amino acid sequence.

[0111] In some embodiments, the composition comprises a plurality of Fc-containing proteins, wherein at least about 90% to 100% of the Fc-containing proteins of the plurality of Fc-containing proteins have the same amino acid sequence. In some embodiments, the composition comprises a plurality of Fc-containing proteins, wherein at least about 90% to 100% of the Fc-containing proteins of the plurality of Fc-containing proteins have the same Fc domain amino acid sequence. In some embodiments, the composition comprises a plurality of Fc-containing proteins, wherein at least about 90% to 100% of the Fc-containing proteins of the plurality of Fc-containing proteins have the same heavy chain amino acid sequence.

[0112] In some embodiments, the composition comprises a plurality of Fc-containing proteins, wherein 100% of the Fc-containing proteins of the plurality of Fc-containing proteins have the same amino acid sequence. In some embodiments, the composition comprises a plurality of Fc-containing proteins, wherein 100% of the Fc-containing proteins of the plurality of Fc-containing proteins have the same Fc domain amino acid sequence. In some embodiments, the composition comprises a plurality of Fc-containing proteins, wherein 100% of the Fc-containing proteins of the plurality of Fc-containing proteins have the same heavy chain amino acid sequence.

[0113] In some embodiments, the composition comprises a plurality of Fc-containing proteins, wherein the plurality of Fc-containing proteins is substantially homogeneous in charge state. As used herein, “substantially homogeneous in charge state” refers to at least about 80%, 85%, 90%, 95%, or 100% of the plurality of Fc-containing proteins having the same charge. It is well appreciated in the art that the charge (e.g., surface charge and net charge) of a Fc-containing protein is highly dependent on the molecular composition of said Fc-containing protein. In some embodiments, the composition comprises a plurality of Fc-containing proteins that is substantially homogeneous in charge state, wherein at least about 80%, 85%, 90%, 95%, or 100% of the plurality of Fc-containing proteins have the same surface charge. In some embodiments, the composition comprises a plurality of Fc-containing proteins that is substantially homogeneous in charge state, wherein at least about 80%, 85%, 90%, 95%, or 100% of the plurality of Fc-containing proteins have the same net charge. In some embodiments, the composition comprises a plurality of Fc-containing proteins that is substantially homogeneous in charge state, wherein each Fc-containing protein of the plurality of Fc-containing proteins has the same charge on an amino acid residue in a position of an amino acid sequence, as compared across each of the plurality of Fc-containing proteins, (e.g., -1, 0, +1). In some embodiments, the composition comprises a plurality of Fc-containing proteins that is substantially homogeneous in charge state, wherein each Fc-containing protein of the plurality of Fc-containing proteins has the same surface charge. In some embodiments, the composition comprises a plurality of Fc-containing proteins that is substantially homogeneous in charge state, wherein each Fc-containing protein of the plurality of Fc-containing proteins has the same net charge. In some embodiments, the composition comprises a plurality of Fc-containing proteins that is substantially homogeneous in charge state, wherein at least about 80%, 85%, 90%, 95%, or 100%

of the plurality of Fc-containing proteins have the same pI. In some embodiments, the charge is associated with a Fc-containing protein (e.g., an amino acid residue). In some embodiments, the charge is associated with a post-translational modification of a Fc-containing protein.

[0114] In some embodiments, the composition is a pharmaceutical composition. In some embodiments, the pharmaceutical composition is in a form for storage. In some embodiments, the pharmaceutical composition is in a form for product transportation. In some embodiments, the pharmaceutical composition is frozen. In some embodiments, the pharmaceutical composition is lyophilized. In some embodiments, the pharmaceutical composition is reconstituted. In some embodiments, the pharmaceutical composition is an administration composition. In some embodiments, the pharmaceutical composition is in a form for administration to an individual in need thereof.

[0115] In some embodiments, the pharmaceutical composition is a sterile pharmaceutical composition. Sterile pharmaceutical formulations are compounded or manufactured according to pharmaceutical-grade sterilization standards (e.g., United States Pharmacopeia Chapters 797, 1072, and 1211; California Business & Professions Code 4127.7; 16 California Code of Regulations 1751, 21 Code of Federal Regulations 21, or ex-U.S. counterparts to such regulations) known to those of skill in the art.

[0116] In some embodiments, the pharmaceutical composition is a stable formulation. As used herein, “stable” formulation is one in which the Fc-containing proteins therein essentially retain physical and chemical stability and integrity upon storage. Various analytical techniques for measuring protein stability are available in the art and are reviewed in, for example, Jones, *Adv Drug Delivery Rev*, 10, 1993. Stability can be assessed at a selected temperature for a selected time period. For example, the extent of aggregation during storage can be used as an indicator of protein stability. Thus, a “stable” formulation may be one wherein less than about 10% and preferably less than about 5% of the Fc-containing protein are present as an aggregate in the formulation.

[0117] In some embodiments, the pharmaceutical composition is a reconstituted formulation. As used herein, a “reconstituted” formulation is one which has been prepared by dissolving a lyophilized Fc-containing protein formulation in a diluent such that the Fc-containing protein is dispersed throughout. The reconstituted

formulation is suitable for administration (e.g. intravenous or subcutaneous administration) to an individual in need there.

[0118] In some embodiments, the pharmaceutical composition is an isotonic formulation. As used herein, an “isotonic” formulation is one which has essentially the same osmotic pressure as human blood. Isotonic formulations will generally have an osmotic pressure from about 250 to 350 mOsm. The term “hypotonic” describes a formulation with an osmotic pressure below that of human blood. Correspondingly, the term “hypertonic” is used to describe a formulation with an osmotic pressure above that of human blood.

[0119] In some embodiments, the pharmaceutical composition is at a specified pH. In some embodiments, the pharmaceutical composition is at a pH of about 5-7, about 5-6, or about 5-5.5. In some embodiments, the pharmaceutical composition is at a pH of about 5.3. In some embodiments, the pharmaceutical composition is at a pH of about 5.4. In some embodiments, the pharmaceutical composition is pH adjusted.

[0120] In some embodiments, the pharmaceutical composition further comprises a pharmaceutically acceptable carrier. As used herein, a “pharmaceutically acceptable carrier” refers to an ingredient in a pharmaceutical composition, other than an active ingredient. A pharmaceutically acceptable carrier includes, but is not limited to, a buffer, excipient, stabilizer, or preservative. Pharmaceutically acceptable carriers are generally nontoxic to recipients at the dosages and concentrations employed, and include, but are not limited to: buffers such as phosphate, citrate, and other organic acids; antioxidants including ascorbic acid and methionine; preservatives (such as octadecyldimethylbenzyl ammonium chloride; hexamethonium chloride; benzalkonium chloride; benzethonium chloride; phenol, butyl or benzyl alcohol; alkyl parabens such as methyl or propyl paraben; catechol; resorcinol; cyclohexanol; 3-pentanol; and m-cresol); low molecular weight (less than about 10 residues) polypeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids such as glycine, glutamine, asparagine, histidine, arginine, or lysine; monosaccharides, disaccharides, and other carbohydrates including glucose, mannose, or dextrans; chelating agents such as EDTA; sugars such as sucrose, mannitol, trehalose or sorbitol; salt-forming counter-ions such as sodium; metal complexes (*e.g.*, Zn-protein complexes); and/or non-ionic surfactants such as polyethylene glycol (PEG). Exemplary pharmaceutically acceptable carriers herein further include interstitial drug dispersion agents such as

soluble neutral-active hyaluronidase glycoproteins (sHASEGP), for example, human soluble PH-20 hyaluronidase glycoproteins, such as rHuPH20 (HYLENEX[®], Baxter International, Inc.). In some embodiments, the pharmaceutically acceptable carrier is selected from the group consisting of sodium acetate, sucrose, polysorbate (e.g., polysorbate 20), sodium succinate, histidine HCl, and sodium chloride.

[0121] In some embodiments, the pharmaceutical composition further comprises a pharmaceutically acceptable acid. As used herein, a “pharmaceutically acceptable acid” includes inorganic and organic acids which are non-toxic at the concentration and manner in which they are formulated. For example, suitable inorganic acids include hydrochloric, perchloric, hydrobromic, hydroiodic, nitric, sulfuric, sulfonic, sulfinic, sulfanilic, phosphoric, carbonic, etc. Suitable organic acids include straight and branched-chain alkyl, aromatic, cyclic, cycloaliphatic, arylaliphatic, heterocyclic, saturated, unsaturated, mono, di- and tri-carboxylic, including for example, formic, acetic, 2-hydroxyacetic, trifluoroacetic, phenylacetic, trimethylacetic, t-butyl acetic, anthranilic, propanoic, 2-hydroxypropanoic, 2-oxopropanoic, propandioic, cyclopentanepropionic, cyclopentane propionic, 3-phenylpropionic, butanoic, butandioic, benzoic, 3-(4-hydroxybenzoyl)benzoic, 2-acetoxy-benzoic, ascorbic, cinnamic, lauryl sulfuric, stearic, muconic, mandelic, succinic, embonic, fumaric, malic, maleic, hydroxymaleic, malonic, lactic, citric, tartaric, glycolic, glyconic, gluconic, pyruvic, glyoxalic, oxalic, mesylic, succinic, salicylic, phthalic, palmoic, palmeic, thiocyanic, methanesulphonic, ethanesulphonic, 1,2-ethanedisulfonic, 2-hydroxyethanesulfonic, benzenesulphonic, 4-chlorobenzenesulfonic, naphthalene-2-sulphonic, p-toluenesulphonic, camphorsulphonic, 4-methylbicyclo [2,2,2]-oct-2-ene-1-carboxylic, glucoheptonic, 4,4'-methylenebis-3-(hydroxy-2-ene-1-carboxylic acid), hydroxynaphthoic.

[0122] In some embodiments, the pharmaceutical composition further comprises a pharmaceutically acceptable base. As used herein, a “pharmaceutically acceptable base” includes inorganic and organic bases which are non-toxic at the concentration and manner in which they are formulated. For example, suitable bases include those formed from inorganic base forming metals such as lithium, sodium, potassium, magnesium, calcium, ammonium, iron, zinc, copper, manganese, aluminum, N-methylglucamine, morpholine, piperidine and organic nontoxic bases including, primary, secondary and tertiary amines, substituted amines, cyclic amines and basic ion exchange resins, for example, isopropylamine, trimethylamine, diethylamine,

triethylamine, tripropylamine, ethanolamine, 2-diethylaminoethanol, trimethamine, dicyclohexylamine, lysine, arginine, histidine, caffeine, procaine, hydrabamine, choline, betaine, ethylenediamine, glucosamine, methylglucamine, theobromine, purines, piperazine, piperidine, N-ethylpiperidine, polyamine resins and the like. Particularly preferred organic non-toxic bases are isopropylamine, diethylamine, ethanolamine, trimethamine, dicyclohexylamine, choline, and caffeine.

[0123] Additional pharmaceutically acceptable acids and bases useable with the present invention include those which are derived from the amino acids, for example, histidine, glycine, phenylalanine, aspartic acid, glutamic acid, lysine and asparagine.

[0124] In some embodiments, the pharmaceutical composition further comprises a pharmaceutically acceptable buffer or salt, for example, those derived from both acid and base addition salts of the above indicated acids and bases. Specific buffers and/or salts include histidine, succinate and acetate.

[0125] In some embodiments, the pharmaceutical composition further comprises a pharmaceutically acceptable sugar. As used herein, a “pharmaceutically acceptable sugar” is a molecule which, when combined with a Fc-containing protein, significantly prevents or reduces chemical and/or physical instability of the Fc-containing protein upon storage. When the formulation is intended to be lyophilized and then reconstituted, “pharmaceutically acceptable sugars” may also be known as a “lyoprotectant”. Exemplary sugars and their corresponding sugar alcohols include: an amino acid such as monosodium glutamate or histidine; a methylamine such as betaine; a lyotropic salt such as magnesium sulfate; a polyol such as trihydric or higher molecular weight sugar alcohols, e.g. glycerin, dextran, erythritol, glycerol, arabitol, xylitol, sorbitol, and mannitol; propylene glycol; polyethylene glycol; PLURONICS®; and combinations thereof. Additional exemplary lyoprotectants include glycerin and gelatin, and the sugars mellibiose, melczitose, raffinose, mannotriose and stachyose. Examples of reducing sugars include glucose, maltose, lactose, maltulose, iso-maltulose and lactulose. Examples of non-reducing sugars include non-reducing glycosides of polyhydroxy compounds selected from sugar alcohols and other straight chain polyalcohols. Preferred sugar alcohols are monoglycosides, especially those compounds obtained by reduction of disaccharides such as lactose, maltose, lactulose and maltulose. The glycosidic side group can be either glucosidic or galactosidic. Additional examples of sugar alcohols are glucitol, maltitol, lactitol and iso-maltulose. The preferred pharmaceutically-acceptable sugars

are the non-reducing sugars trehalose or sucrose. Pharmaceutically acceptable sugars are added to the formulation in a “protecting amount” (e.g. pre-lyophilization) which means that the protein essentially retains its physical and chemical stability and integrity during storage (e.g., after reconstitution and storage).

[0126] In some embodiments, the pharmaceutical composition further comprises a pharmaceutically acceptable preservative. As used herein, a “pharmaceutically acceptable preservative” is a compound which can be added to the formulations herein to reduce bacterial activity. Examples of potential preservatives include, but are not limited to, octadecyldimethylbenzyl ammonium chloride, hexamethonium chloride, benzalkonium chloride (a mixture of alkylbenzyltrimethylammonium chlorides in which the alkyl groups are long-chain compounds), and benzethonium chloride. Other types of preservatives include aromatic alcohols such as phenol, butyl and benzyl alcohol, alkyl parabens such as methyl or propyl paraben, catechol, resorcinol, cyclohexanol, 3-pentanol, and m-cresol.

[0127] In some embodiments, the composition comprising a plurality of Fc-containing proteins is a cell culture medium, wherein substantially all Fc-containing proteins have a C-terminal lysine on each Fc domain. In some embodiments, the cell culture medium is a nutrient media. Nutrient media contains all elements needed for host cell growth. In some embodiments, the cell culture medium is a minimal medium. Minimal media contains the minimum nutrients possible for host cell growth, for example, generally without the presence of amino acids. In some embodiments, the cell culture medium is a selective medium. Selective media comprises an agent that inhibits growth of a select organism.

[0128] In some embodiments, the cell culture medium further comprises a cell culture medium nutrient for cell support and/or growth. In some embodiments, the cell culture medium nutrient is selected from, for example: proteins; peptides; amino acids; carbohydrates; metals and minerals, for example calcium, magnesium, iron; trace metals, for example, phosphates and sulphates; buffers; pH indicators, for example, phenol red, bromo-cresol purple; and antimicrobial agents.

[0129] In some embodiments, the cell culture medium further comprises a host cell. In some embodiment, the culture medium is substantially devoid of a host cell.

[0130] In some embodiments, the composition is a cell lysate. In some embodiments, the cell lysate comprises a plurality of Fc-containing proteins and host cell components. In some embodiments, the cell lysate is a centrifuged cell lysate. In

some embodiments, the cell lysate comprises a precipitated portion of the cell lysate and a supernatant portion of the cell lysate. In some embodiments, the cell lysate comprises a pelleted portion of the cell lysate and a supernatant portion of the cell lysate.

[0131] In some embodiments, the composition is an eluate from a protein purification column. As used herein, “eluate” refers to any fluid that passes through a protein purification column. In some embodiments, the eluate comprises a fluid that is isolated from a flow-through fluid. In some embodiments, the eluate comprises a fluid that is isolated from a wash fluid. In some embodiments, the eluate comprises a fluid that is isolated from one or more wash fluids. In some embodiments, the eluate comprises a fluid that is isolated from an elution fluid.

[0132] Protein purification columns for enriching Fc-containing proteins are known in the art. The following are exemplary types of protein purification columns: immunoaffinity, protein A, protein G, ion-exchange, reverse phase, cation-exchange, strong cation-exchange, anion exchange, hydrophobic, mixed modal, hydroxylapatite, and gel filtration.

[0133] In some embodiments, the composition is a library of Fc-containing proteins, wherein at least two of the Fc-containing proteins of the plurality of Fc-containing proteins are different. In some embodiments, the library comprises at least two Fc-containing proteins that bind to different antigens. In some embodiments, the library comprises at least two Fc-containing proteins that bind to different epitopes. In some embodiments, the different Fc-containing proteins are contained in different vessels.

[0134] In some embodiments, the present invention provides libraries comprising at least 2, 3, 4, 5, 10, 30, 100, 250, 500, 750, 1000, 2500, 5000, 7500, 10000, 25000, 50000, 75000, 100000, 250000, 500000, 750000, 1000000, 2500000, 5000000, 7500000, 10000000, or more than 10000000 different Fc-containing proteins.

Batch

[0135] The present application provides large-scale batches (e.g., commercial batches or batches at manufacture scale) of any of the compositions described in the embodiments herein.

[0136] In some embodiments, the batch comprises at least about 5 g, 10 g, 50 g, 100 g, 200 g, 300 g, 400 g, 500 g, 600 g, 700 g, 800 g, 900 g, 1,000 g, 1,500 g, 2,000 g,

2,500 g, 3,000 g, 3,500 g, 4,000 g, 4,500 g, or 5,000 g of a plurality of Fc-containing proteins, wherein each Fc-containing protein of the plurality of Fc-containing proteins has a C-terminal lysine on each Fc domain. In some embodiments, the batch comprises at least about 5-5,000 g, 50-4,000 g, or about 100-1,000 g of a plurality of Fc-containing proteins, wherein each Fc-containing protein of the plurality of Fc-containing proteins has a C-terminal lysine on each Fc domain.

[0137] In some embodiments, the batch is in a form for drug storage. In some embodiments, the batch is in a form for product transportation. In some embodiments, the batch is in a form for administration to an individual in need thereof. In some embodiments, the batch is lyophilized. In some embodiments, the batch is not conjugated to an agent. In some embodiments, the batch is conjugated to an agent. In some embodiments, the batch further comprises a formulation component.

[0138] In some embodiments, the batch is a cell culture medium. In some embodiments the batch is a cell lysate.

[0139] In some embodiments, the batch, or a portion thereof, is in a vessel. In some embodiments, the batch, or portion thereof, is in a vial. In some embodiments, the batch, or portion thereof, is in a plurality of vials. In some embodiments, the batch, or portion thereof, is in a syringe.

[0140] In some embodiments, the batch, or a portion thereof, is in a plurality of vials, wherein each vial comprises a plurality of Fc-containing proteins, wherein each Fc-containing protein has a C-terminal lysine on each Fc domain. In some embodiments, the batch, or a portion thereof, is in a plurality of vials, wherein each vial comprises a plurality of Fc-containing proteins, wherein at least about 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% of the plurality of Fc-containing protein in each vial has a C-terminal lysine on each Fc domain. In some embodiments, the batch, or a portion thereof, is in a plurality of vials, wherein at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 100% of the vials comprise at least about 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% of the plurality of Fc-containing proteins comprising a C-terminal lysine on each Fc domain.

[0141] In some embodiments, the batch is aliquoted into a unit dosage. As used herein, a “unit dosage” is the amount of Fc-containing protein intended for administration as a single unit dose. In some embodiments, the single unit dose is about 1 to about 500 mg of a Fc-containing protein. In some embodiments, the unit

dosage is packaged in a container. In some embodiments, the unit dosage is packaged in a vial.

[0142] In some embodiments, the size of the commercial batch is no greater than about 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 times the size of the clinical batch. As used herein, “commercial batch” refers to an amount of Fc-containing protein produced during one or more production runs completed for purposes of commercial production and/or distribution. As used herein, “clinical batch” refers to an amount of Fc-containing protein produced during one or more production runs completed for purposes of clinical testing. In some embodiments, the size of the commercial batch is no greater than 10 times the size of the clinical batch.

[0143] In some embodiments, the vessel comprises an aliquot of a commercial batch as described herein, wherein the commercial batch comprises a composition comprising a Fc-containing protein, wherein each Fc domain of the Fc-containing protein has a C-terminal lysine. In some embodiments, the vial comprises an aliquot of a commercial batch as described herein, wherein the commercial batch comprises a composition comprising a Fc-containing protein, wherein each Fc domain of the Fc-containing protein has a C-terminal lysine. In some embodiments, the syringe comprises an aliquot of a commercial batch as described herein, wherein the commercial batch comprises a composition comprising a Fc-containing protein, wherein each Fc domain of the Fc-containing protein has a C-terminal lysine.

Cell culture system

[0144] The invention provides a cell culture system comprising any host cell described in the embodiments herein.

[0145] In some embodiments, the cell culture system maintains a host cell in an environment for the production of a Fc-containing protein. In some embodiments, the cell culture system maintains a host cell in an environment for growth.

[0146] In some embodiments, the cell culture system comprises a seed culture. In some embodiments, the cell culture system comprises an inoculum culture. In some embodiments, the inoculum culture is a primary inoculum culture. In some embodiments, the inoculum culture is a secondary inoculum. In some embodiments, the cell culture system comprises a production culture.

[0147] In some embodiments, the cell culture system comprises a mechanism for maintaining a host cell in a specified environment. In some embodiments, the

specified environment is a specified temperature. In some embodiments, the specified temperature is about 15 °C to about 45 °C. In some embodiments, the specified temperature is about 30 °C. In some embodiments, the specified environment is a specified pH. In some embodiments, the specified environment is a specified dissolved oxygen concentration. In some embodiments, the specified environment is a specified nutrient level.

[0148] In some embodiments, the cell culture further comprises a culture medium. In some embodiments, the cell culture system comprises a seed culture. In some embodiments, the culture medium is an inoculum culture. In some embodiments, the inoculum culture medium is a primary inoculum culture medium. In some embodiments, the inoculum culture medium is a secondary inoculum medium. In some embodiments, the culture medium is a production culture medium.

[0149] In some embodiments, the culture system comprises a Fc-containing protein as described in the embodiments herein. In some embodiments, the culture system comprises a plurality of Fc-containing protein as described in the embodiments herein. In some embodiments, the culture system comprises a composition comprising a Fc-containing protein as described in the embodiments herein.

[0150] In some embodiments, the culture system comprises a Fc-containing protein, wherein each Fc domain of the Fc-containing protein has a C-terminal lysine. In some embodiments, the culture system comprises a plurality of Fc-containing protein, wherein each Fc-containing protein of the plurality of Fc-containing proteins has a C-terminal lysine Fc domain. In some embodiments, the culture system comprises a composition comprising a plurality of Fc-containing protein, wherein each Fc-containing protein of the plurality of Fc-containing proteins has a C-terminal lysine on each Fc domain.

Methods of producing a host cell with a reduction of a CpD expression level

[0151] The present application provides methods of producing any host cell described in the embodiments herein, wherein the methods comprise inactivating the CpD gene of the host cell.

[0152] Generally, the methods of producing a host cell comprise inactivating the CpD gene of the host cell. In some embodiments, the method of producing a host cell, wherein the host cell has a reduction of a CpD expression level, comprises inactivating the CpD gene using a siRNA system. In some embodiments, the method

of producing a host cell comprises inactivating the CpD gene using a siRNA system comprising a siRNA nucleotide sequence that is about 10 to 200 nucleotides in length, or about 10 to 100 nucleotides in length, or about 15 to 100 nucleotides in length, or about 10 to 60 nucleotides in length, or about 15 to 60 nucleotides in length, or about 10 to 50 nucleotides in length, or about 15 to 50 nucleotides in length, or about 10 to 30 nucleotides in length, or about 15 to 30 nucleotides in length. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a siRNA system comprising a siRNA nucleotide sequence that is approximately 10-25 nucleotides in length. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a siRNA system comprising a siRNA nucleotide sequence that is approximately 15-25 nucleotides in length. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a siRNA system comprising a siRNA nucleotide sequence that is at least about 10, at least about 15, at least about 20, or at least about 25 nucleotides in length. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a siRNA system comprising a nucleotide sequence that is at least about 80%, at least about 85%, at least about 90%, at least about 95%, or 100% complementary to a region of a CpD mRNA molecule. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a siRNA system comprising a siRNA nucleotide sequence that is at least at least about 80%, at least about 85%, at least about 90%, at least about 95%, or 100% complementary to a region of a CpD pro-mRNA molecule. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a siRNA system comprising a double stranded RNA molecule. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a siRNA system comprising a single stranded RNA molecule. Exemplary CpD siRNA nucleotide sequences are listed in Table 1.

[0153] In some embodiments, the method of producing a host cell comprises delivering a siRNA system to a host cell. In some embodiments, the method of producing a host cell comprises delivering a siRNA system to a host cell, wherein the siRNA system comprises a siRNA nucleotide sequence. In some embodiments, the method of producing a host cell comprises delivering a siRNA system to the host cell using electroporation. In some embodiments, the method of producing a host cell comprises delivering a siRNA system to the host cell using transfection techniques. In

some embodiments, the method of producing a host cell comprises delivering a siRNA system to the host cell using a virus.

[0154] In some embodiments, the method of producing a host cell, wherein the host cell has a reduced level of CpD expression, further comprises determining the level of CpD gene inactivation in the host cell. In some embodiments, determining the level of CpD gene inactivation comprises determining a level of CpD expression prior to delivering a siRNA nucleotide sequence to the host cell. In some embodiments, determining the level of CpD gene inactivation comprises determining a level of CpD expression after delivering a siRNA nucleotide sequence to the host cell. In some embodiments, the level of CpD expression is determined at the RNA level. In some embodiments, the method of determining the level of CpD gene inactivation comprises determining the level of CpD expression using PCR. In some embodiments, the level of CpD expression is determined at the protein level. In some embodiments, the method of determining the level of CpD gene inactivation comprises determining the CpD expression level using immunohistochemistry. In some embodiments, the method of determining the level of CpD gene inactivation comprises determining the CpD expression level using Western blot. In some embodiments, the method of determining the level of CpD gene inactivation comprises determining the CpD expression level using flow cytometry. In some embodiments, the level of CpD gene inactivation is determined by comparing the level of CpD expression after delivery of a siRNA to a control value. In some embodiments, the level of CpD gene inactivation is determined by comparing the level of CpD expression after delivery of a siRNA to a wild type value. In some embodiments, the level of CpD gene inactivation is determined by comparing a level of CpD expression after delivery of a siRNA to a level of CpD expression prior to delivery of a CpD siRNA.

[0155] Exemplary primers and probes for determining the level of CpD at the RNA level are provided in Table 2.

Table 2. Exemplary nucleotide sequences for determining the RNA expression level of CpD.

	SEQ ID NO.:	Nucleotide sequence:
CpD forward primer	1	CCC ACA CAT TAC AAA TCT TAC CA
CpD reverse primer	2	GAG ATT TCG AGG GAC CAA AT

CpD probe	3	TTG GGA CAG AGT GCT GAG TAT CGT CA
-----------	---	---------------------------------------

[0156] In some embodiments, the method of producing a host cell, wherein the host cell has a reduced level of CpD expression, comprises inactivating the CpD gene using a shRNA system. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a shRNA system comprising a shRNA nucleotide sequence that is about 10 to 200 nucleotides in length, or about 10 to 100 nucleotides in length, or about 15 to 100 nucleotides in length, or about 10 to 60 nucleotides in length, or about 15 to 60 nucleotides in length, or about 10 to 50 nucleotides in length, or about 15 to 50 nucleotides in length, or about 10 to 30 nucleotides in length, or about 15 to 30 nucleotides in length. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a shRNA system comprising a shRNA nucleotide sequence that is approximately 10-25 nucleotides in length. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a shRNA system comprising a shRNA nucleotide sequence that is approximately 15-25 nucleotides in length. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a shRNA system comprising a shRNA nucleotide sequence that is at least about 10, at least about 15, at least about 20, or at least about 25 nucleotides in length. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a shRNA system comprising a nucleotide sequence that is at least about 80%, at least about 85%, at least about 90%, at least about 95%, or 100% complementary to a region of a CpD mRNA molecule. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a shRNA system comprising a shRNA nucleotide sequence that is at least at least about 80%, at least about 85%, at least about 90%, at least about 95%, or 100% complementary to a region of a CpD pro-mRNA molecule. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a shRNA system comprising a double stranded RNA molecule. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a shRNA system comprising a single stranded RNA molecule. Exemplary CpD shRNA nucleotide sequences are listed in Table 1.

[0157] In some embodiments, the method of producing a host cell comprises delivering a shRNA system to a host cell. In some embodiments, the method of producing a host cell comprises delivering a shRNA system to a host cell, wherein the shRNA system comprises a shRNA nucleotide sequence. In some embodiments, the method of producing a host cell comprises delivering a shRNA system to the host cell using electroporation. In some embodiments, the method of producing a host cell comprises delivering a shRNA system to the host cell using transfection techniques. In some embodiments, the method of producing a host cell comprises delivering a shRNA system to the host cell using a virus.

[0158] In some embodiments, the method of producing a host cell, wherein the host cell has a reduced level of CpD expression, further comprises determining the level of CpD gene inactivation in the host cell. In some embodiments, determining the level of CpD gene inactivation comprises determining a level of CpD expression prior to delivering a shRNA nucleotide sequence to the host cell. In some embodiments, determining the level of CpD gene inactivation comprises determining a level of CpD expression after delivering a shRNA nucleotide sequence to the host cell. In some embodiments, the level of CpD expression is determined at the RNA level. In some embodiments, the method of determining the level of CpD gene inactivation comprises determining the level of CpD expression using PCR. In some embodiments, the level of CpD expression is determined at the protein level. In some embodiments, the method of determining the level of CpD gene inactivation comprises determining the CpD expression level using immunohistochemistry. In some embodiments, the method of determining the level of CpD gene inactivation comprises determining the CpD expression level using Western blot. In some embodiments, the method of determining the level of CpD gene inactivation comprises determining the CpD expression level using flow cytometry. In some embodiments, the level of CpD gene inactivation is determined by comparing the level of CpD expression after delivery of a shRNA to a control value. In some embodiments, the level of CpD gene inactivation is determined by comparing the level of CpD expression after delivery of a shRNA to a wild type value. In some embodiments, the level of CpD gene inactivation is determined by comparing a level of CpD expression after delivery of a shRNA to a level of CpD expression prior to delivery of a CpD shRNA.

[0159] Exemplary primers and probes for determining the level of CpD at the RNA level are provided in Table 2.

[0160] Additionally provided in the present application are methods of producing a host cell by gene deletion or gene addition or substitution. For example, methods include, but are not limited to use of CRISPR, TALEN, ZFN, and meganuclease systems.

[0161] Generally, a CRISPR system comprises a caspase protein, such as Cas9, and an RNA sequence comprising a nucleotide sequence, referred to as a guide sequence, that is complementary to a sequence of interest. The caspase and RNA sequence form a complex that identify a DNA sequence of a host cell, and subsequently the nuclease activity of the caspase allows for cleavage of the DNA strand. Caspases isotypes have single-stranded DNA or double-stranded DNA nuclease activity. Design of guide RNA sequences and number of guide RNA sequences used in a CRISPR system allow for removal of a specific stretch of a gene and/or addition of a DNA sequence.

[0162] In some embodiments, the method of producing a host cell, wherein the host cell has a reduced level of CpD expression, comprises inactivating the CpD gene using a CRISPR system. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a CRISPR system comprising a coding vector. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a CRISPR system comprising a coding vector comprising a DNA endonuclease gene. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a CRISPR system comprising a coding vector comprising a CAS gene. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a CRISPR system comprising a coding vector comprising a CAS9 gene. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a CRISPR system comprising a coding vector encoding a CAS9 gene. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a CRISPR system comprising a Cas protein. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a CRISPR system comprising a Cas9 protein. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a CRISPR system comprising a coding vector encoding a RNA molecule capable of interacting with the Cas9 protein. In some embodiments, the method of producing a host cell comprises inactivating the

CpD gene using a CRISPR system comprising a coding vector encoding a RNA molecule comprising a guide RNA (gRNA) unit, wherein the gRNA unit comprises a nucleotide sequence that is complementary to a portion of a CpD gene sequence. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a CRISPR system comprising a RNA molecule comprising a gRNA unit, wherein the gRNA unit comprises a nucleotide sequence that is complementary to a portion of a CpD gene sequence. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a CRISPR system comprising a coding vector encoding a RNA molecule comprising a trans-activating crRNA (tracrRNA) unit. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a CRISPR system comprising a RNA molecule comprising a tracrRNA unit. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a CRISPR system comprising a coding vector encoding a RNA molecule comprising a gRNA unit and a tracrRNA unit, wherein the gRNA unit comprises a nucleotide sequence that is complementary to a portion of a gene sequence. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a CRISPR system comprising a RNA molecule comprising a gRNA unit and a tracrRNA unit, wherein the gRNA unit comprises a nucleotide sequence that is complementary to a portion of a CpD gene sequence. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a CRISPR system comprising: a) a first RNA molecule comprising a gRNA unit, wherein the gRNA unit comprises a first nucleotide sequence that is complementary to a portion of a CpD gene sequence; and b) a second RNA molecule comprising a gRNA unit, wherein the gRNA unit comprises a second nucleotide sequence that is complementary to a portion of a CpD gene sequence. In some embodiments, the first nucleotide sequence and second nucleotide sequence are different. In some embodiments, the first nucleotide sequence is complementary to a portion of a CpD gene sequence that is in a different location than the region of the portion of the CpD gene that is complementary to the second nucleotide sequence.

[0163] Exemplary CpD target sequences included in gRNA sequences are listed in Table 3.

Table 3. Exemplary CpD target sequences included in gRNA.

SEQ ID NO.:	CpD target sequence:
21	GAA GAC GAG ACT TTC AAA GAC GG
22	TCA GTT AGG TGG CTT AGG TTC GG

[0164] In some embodiments, the method of producing a host cell comprises delivering a CRISPR system to the host cell. In some embodiments, the method of producing a host cell comprises delivering a vector comprising a CRISPR system using a delivery vector. In some embodiments, the delivery vector is a virus vector. In some embodiments, the delivery vector is a lentivirus. In some embodiments, the delivery vector is an adenovirus. In some embodiments, the vector comprises a promoter.

[0165] Generally, a TALEN system comprises one or more restriction nucleases and two or more protein complexes that allow for recognition of a DNA sequence and subsequent double-stranded DNA cleavage. A protein complex of the TALEN system comprises a number of transcription activator-like effectors (TALEs), each recognizing a specific nucleotide, and a domain of a restriction nuclease. Generally, a TALEN system is designed so that two protein complexes, each comprising TALEs and a domain of a restriction nuclease, will individually bind to DNA sequences in a manner to allow for the two domains (one from each protein complex) of a restriction nuclease to form an active nuclease and cleave a specific DNA sequence. Design of number of protein complexes and sequences to be cleaved in a TALEN system allows for removal of a specific stretch of a gene and/or addition of a DNA sequence.

[0166] In some embodiments, the method of producing a host cell, wherein the host cell has a reduced level of CpD expression, comprises inactivating the CpD gene using a TALEN system. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a TALEN system comprising a first TALEN unit. In some embodiments, the first TALEN unit comprises a first TALEN binding unit. In some embodiments, the first TALEN binding unit comprises at least one transcription activator-like effector (TALE) and a first nuclease domain. In some embodiments, the first TALEN binding unit comprises at least 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15 TALEs and a first nuclease domain, wherein the TALEs are linked together, and wherein the linked TALEs recognize a portion of a CpD

nucleotide sequence. In some embodiments, the first TALEN binding unit comprises at least 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15 TALEs and a first nuclease domain, wherein the TALEs are linked together, wherein the linked TALEs recognize a portion of a CpD nucleotide sequence, and wherein the linked TALEs are further linked to the first nuclease domain. In some embodiments, the first TALEN unit further comprises a second TALEN binding unit. In some embodiments, the second TALEN binding unit comprises at least one transcription activator-like effector (TALE) and a second nuclease domain. In some embodiments, the second TALEN binding unit comprises at least 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15 TALEs and a second nuclease domain, wherein the TALEs are linked together, and wherein the linked TALEs recognize a portion of a CpD nucleotide sequence. In some embodiments, the second TALEN binding unit comprises at least 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15 TALEs and a second nuclease domain, wherein the TALEs are linked together, wherein the linked TALEs recognize a portion of a CpD nucleotide sequence, and wherein the linked TALEs are further linked to the second nuclease domain. In some embodiments, the first TALEN binding unit and second TALEN binding unit bind to different sequences of the CpD gene. In some embodiments, the first nuclease domain is a domain of an endonuclease. In some embodiments, the first nuclease domain is a domain of a restriction endonuclease. In some embodiments, the first nuclease domain is a domain of FokI. In some embodiments, the second nuclease domain is a domain of an endonuclease. In some embodiments, the second nuclease domain is a domain of a restriction endonuclease. In some embodiments, the second nuclease domain is a domain of FokI. In some embodiments, the first nuclease domain and second nuclease domain associate to comprise an active restriction endonuclease. In some embodiments, the first nuclease domain and second nuclease domain associate to comprise an active FokI enzyme.

[0167] In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a TALEN system further comprising a second TALEN unit. In some embodiments, the second TALEN unit comprises a third TALEN binding unit. In some embodiments, the third TALEN binding unit comprises at least one transcription activator-like effector (TALE) and a third nuclease domain. In some embodiments, the third TALEN binding unit comprises at least 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15 TALEs and a third nuclease domain, wherein the TALEs are linked together, and wherein the linked TALEs recognize a portion of a

CpD nucleotide sequence. In some embodiments, the third TALEN binding unit comprises at least 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15 TALEs and a third nuclease domain, wherein the TALEs are linked together, wherein the linked TALEs recognize a portion of a CpD nucleotide sequence, and wherein the linked TALEs are further linked to the third nuclease domain. In some embodiments, the second TALEN unit further comprises a fourth TALEN binding unit. In some embodiments, the fourth TALEN binding unit comprises at least one transcription activator-like effector (TALE) and a fourth nuclease domain. In some embodiments, the fourth TALEN binding unit comprises at least 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15 TALEs and a fourth nuclease domain, wherein the TALEs are linked together, and wherein the linked TALEs recognize a portion of a CpD nucleotide sequence. In some embodiments, the fourth TALEN binding unit comprises at least 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15 TALEs and a fourth nuclease domain, wherein the TALEs are linked together, wherein the linked TALEs recognize a portion of a CpD nucleotide sequence, and wherein the linked TALEs are further linked to the fourth nuclease domain. In some embodiments, the third TALEN binding unit and fourth TALEN binding unit bind to different sequences of the CpD gene. In some embodiments, the third nuclease domain is a domain of an endonuclease. In some embodiments, the third nuclease domain is a domain of a restriction endonuclease. In some embodiments, the third nuclease domain is a domain of FokI. In some embodiments, the fourth nuclease domain is a domain of an endonuclease. In some embodiments, the fourth nuclease domain is a domain of a restriction endonuclease. In some embodiments, the fourth nuclease domain is a domain of FokI. In some embodiments, the third nuclease domain and fourth nuclease domain associate to comprise an active restriction endonuclease. In some embodiments, the third nuclease domain and fourth nuclease domain associate to comprise an active FokI enzyme.

[0168] In some embodiments, the method of producing a host cell, wherein the host cell has a reduced level of CpD expression, comprises inactivating the CpD gene using a TALEN system, wherein the TALEN system comprises a first TALEN unit and a second TALEN unit that bind to different, non-overlapping portions of a CpD gene sequence.

[0169] In some embodiments, the method of producing a host cell, wherein the host cell has a reduced level of CpD expression, comprises inactivating the CpD gene using a TALEN system, wherein the TALEN system comprises a coding vector

encoding a first TALEN unit. In some embodiments, the method of producing a host cell, wherein the host cell has a reduced level of CpD expression, comprises inactivating the CpD gene using a TALEN system, wherein the TALEN system comprises a coding vector encoding a first TALEN unit, wherein the TALEN system comprises a coding vector encoding a second TALEN unit. In some embodiments, the method of producing a host cell, wherein the host cell has a reduced level of CpD expression, comprises inactivating the CpD gene using a TALEN system, wherein the TALEN system comprises a coding vector encoding a first TALEN unit, wherein the TALEN system comprises a coding vector encoding a first TALEN unit and a second TALEN unit.

[0170] In some embodiments, the method of producing a host cell, wherein the host cell has a reduced level of CpD expression, comprises inactivating the CpD gene using a TALEN system, wherein the TALEN system comprises a first TALEN unit. In some embodiments, the method of producing a host cell, wherein the host cell has a reduced level of CpD expression, comprises inactivating the CpD gene using a TALEN system, wherein the TALEN system comprises a second TALEN unit. In some embodiments, the method of producing a host cell, wherein the host cell has a reduced level of CpD expression, comprises inactivating the CpD gene using a TALEN system, wherein the TALEN system comprises a first TALEN unit and a second TALEN unit.

[0171] In some embodiments, the first TALEN binding unit comprises a group of linked TALEs, wherein the group of TALEs recognize a nucleotide sequence. In some embodiments, the nucleotide sequence is a sequence comprising a portion of a CpD gene. In some embodiments, the nucleotide sequence is a sequence comprising a portion of a CpD gene promoter. In some embodiments, the nucleotide sequence is a sequence comprising a portion of a sequence flanking a CpD gene. In some embodiments, the sequence is at least about 80%, at least about 85%, at least about 90%, at least about 95%, or 100% homologous to a portion of a CpD gene. In some embodiments, the sequence is at least about 80%, at least about 85%, at least about 90%, at least about 95%, or 100% homologous to a portion of a CpD gene promoter. In some embodiments, the sequence is at least about 80%, at least about 85%, at least about 90%, at least about 95%, or 100% homologous to a portion of a sequence flanking a CpD gene.

[0172] In some embodiments, the method of producing a host cell comprises delivering a TALEN system to the host cell. In some embodiments, the method of producing a host cell comprises delivering a vector comprising a TALEN system using a delivery vector. In some embodiments, the delivery vector is a virus vector. In some embodiments, the delivery vector is a lentivirus. In some embodiments, the delivery vector is an adenovirus.

[0173] Generally, a ZFN system comprises one or more restriction nucleases and two or more protein complexes that allow for recognition of a DNA sequence and subsequent double-stranded DNA cleavage. A protein complex of the ZFN system comprises a number of zinc fingers, each recognizing a specific nucleotide codon, and a domain of a restriction nuclease. Generally, a ZFN system is designed so that two protein complexes, each comprising zinc fingers and a domain of a restriction nuclease, will individually bind to DNA sequences in a manner to allow for the two domains (one from each protein complex) of a restriction nuclease to form an active nuclease and cleave a specific DNA sequence. Design of number of protein complexes and sequences to be cleaved in a ZFN system allows for removal of a specific stretch of a gene and/or addition of a DNA sequence.

[0174] In some embodiments, the method of producing a host cell, wherein the host cell has a reduced level of CpD expression, comprises inactivating the CpD gene using a ZFN system. In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a ZFN system comprising a first ZFN unit. In some embodiments, the first ZFN unit comprises a first ZFN binding unit. In some embodiments, the first ZFN binding unit comprises at least one zinc finger and a first nuclease domain. In some embodiments, the first ZFN binding unit comprises at least 2, 3, 4, 5, 6, 7, 8, 9, or 10 zinc fingers and a first nuclease domain, wherein the zinc fingers are linked together, and wherein the linked zinc fingers recognize a portion of a CpD nucleotide sequence. In some embodiments, the first ZFN binding unit comprises at least 2, 3, 4, 5, 6, 7, 8, 9, or 10 zinc fingers and a first nuclease domain, wherein the zinc fingers are linked together, wherein the linked zinc fingers recognize a portion of a CpD nucleotide sequence, and wherein the linked zinc fingers are further linked to the first nuclease domain. In some embodiments, the first ZFN unit further comprises a second ZFN binding unit. In some embodiments, the second ZFN binding unit comprises at least one zinc finger and a second nuclease domain. In some embodiments, the second ZFN binding unit comprises at least 2, 3, 4, 5, 6, 7, 8, 9, or

10 zinc fingers and a second nuclease domain, wherein the zinc fingers are linked together, and wherein the linked zinc fingers recognize a portion of a CpD nucleotide sequence. In some embodiments, the second ZFN binding unit comprises at least 2, 3, 4, 5, 6, 7, 8, 9, or 10 zinc fingers and a second nuclease domain, wherein the zinc fingers are linked together, wherein the linked zinc fingers recognize a portion of a CpD nucleotide sequence, and wherein the linked zinc fingers are further linked to the second nuclease domain. In some embodiments, the first ZFN binding unit and second ZFN binding unit bind to different sequences of the CpD gene. In some embodiments, the first nuclease domain is a domain of an endonuclease. In some embodiments, the first nuclease domain is a domain of a restriction endonuclease. In some embodiments, the first nuclease domain is a domain of Fok1. In some embodiments, the second nuclease domain is a domain of an endonuclease. In some embodiments, the second nuclease domain is a domain of a restriction endonuclease. In some embodiments, the second nuclease domain is a domain of Fok1. In some embodiments, the first nuclease domain and second nuclease domain associate to comprise an active restriction endonuclease. In some embodiments, the first nuclease domain and second nuclease domain associate to comprise an active Fok1 enzyme.

[0175] In some embodiments, the method of producing a host cell comprises inactivating the CpD gene using a ZFN system further comprising a second ZFN unit. In some embodiments, the second ZFN unit comprises a third ZFN binding unit. In some embodiments, the third ZFN binding unit comprises at least one zinc finger and a third nuclease domain. In some embodiments, the third ZFN binding unit comprises at least 2, 3, 4, 5, 6, 7, 8, 9, or 10 zinc fingers and a third nuclease domain, wherein the zinc fingers are linked together, and wherein the linked zinc fingers recognize a portion of a CpD nucleotide sequence. In some embodiments, the third ZFN binding unit comprises at least 2, 3, 4, 5, 6, 7, 8, 9, or 10 zinc fingers and a third nuclease domain, wherein the zinc fingers are linked together, wherein the linked zinc fingers recognize a portion of a CpD nucleotide sequence, and wherein the linked zinc fingers are further linked to the third nuclease domain. In some embodiments, the second ZFN unit further comprises a fourth ZFN binding unit. In some embodiments, the fourth ZFN binding unit comprises at least one zinc finger and a fourth nuclease domain. In some embodiments, the fourth ZFN binding unit comprises at least 2, 3, 4, 5, 6, 7, 8, 9, or 10 zinc fingers and a fourth nuclease domain, wherein the zinc fingers are linked together, and wherein the linked zinc fingers recognize a portion of a CpD

nucleotide sequence. In some embodiments, the fourth ZFN binding unit comprises at least 2, 3, 4, 5, 6, 7, 8, 9, or 10 zinc fingers and a fourth nuclease domain, wherein the zinc fingers are linked together, wherein the linked zinc fingers recognize a portion of a CpD nucleotide sequence, and wherein the linked zinc fingers are further linked to the fourth nuclease domain. In some embodiments, the third ZFN binding unit and fourth ZFN binding unit bind to different sequences of the CpD gene. In some embodiments, the third nuclease domain is a domain of an endonuclease. In some embodiments, the third nuclease domain is a domain of a restriction endonuclease. In some embodiments, the third nuclease domain is a domain of FokI. In some embodiments, the fourth nuclease domain is a domain of an endonuclease. In some embodiments, the fourth nuclease domain is a domain of a restriction endonuclease. In some embodiments, the fourth nuclease domain is a domain of FokI. In some embodiments, the third nuclease domain and fourth nuclease domain associate to comprise an active restriction endonuclease. In some embodiments, the third nuclease domain and fourth nuclease domain associate to comprise an active FokI enzyme.

[0176] In some embodiments, the method of producing a host cell, wherein the host cell has a reduced level of CpD expression, comprises inactivating the CpD gene using a ZFN system, wherein the ZFN system comprises a first ZFN unit and a second TALEN unit that bind to different, non-overlapping portions of a CpD gene sequence.

[0177] In some embodiments, the method of producing a host cell, wherein the host cell has a reduced level of CpD expression, comprises inactivating the CpD gene using a ZFN system, wherein the ZFN system comprises a coding vector encoding a first ZFN unit. In some embodiments, the method of producing a host cell, wherein the host cell has a reduced level of CpD expression, comprises inactivating the CpD gene using a ZFN system, wherein the ZFN system comprises a coding vector encoding a first ZFN unit, wherein the ZFN system comprises a coding vector encoding a second ZFN unit. In some embodiments, the method of producing a host cell, wherein the host cell has a reduced level of CpD expression, comprises inactivating the CpD gene using a ZFN system, wherein the ZFN system comprises a coding vector encoding a first ZFN unit, wherein the ZFN system comprises a coding vector encoding a first ZFN unit and a second ZFN unit.

[0178] In some embodiments, the method of producing a host cell, wherein the host cell has a reduced level of CpD expression, comprises inactivating the CpD gene using a ZFN system, wherein the ZFN system comprises a first ZFN unit. In some

embodiments, the method of producing a host cell, wherein the host cell has a reduced level of CpD expression, comprises inactivating the CpD gene using a ZFN system, wherein the ZFN system comprises a second ZFN unit. In some embodiments, the method of producing a host cell, wherein the host cell has a reduced level of CpD expression, comprises inactivating the CpD gene using a ZFN system, wherein the ZFN system comprises a first ZFN unit and a second ZFN unit.

[0179] In some embodiments, the first ZFN binding unit comprises a group of linked zinc fingers, wherein the group of zinc fingers recognize a nucleotide sequence. In some embodiments, the nucleotide sequence is a sequence comprising a portion of a CpD gene. In some embodiments, the nucleotide sequence is a sequence comprising a portion of a CpD gene promoter. In some embodiments, the nucleotide sequence is a sequence comprising a portion of a sequence flanking a CpD gene. In some embodiments, the sequence is at least about 80%, at least about 85%, at least about 90%, at least about 95%, or 100% homologous to a portion of a CpD gene. In some embodiments, the sequence is at least about 80%, at least about 85%, at least about 90%, at least about 95%, or 100% homologous to a portion of a CpD gene promoter. In some embodiments, the sequence is at least about 80%, at least about 85%, at least about 90%, at least about 95%, or 100% homologous to a portion of a sequence flanking a CpD gene.

[0180] In some embodiments, the method of producing a host cell comprises delivering a ZFN system to the host cell. In some embodiments, the method of producing a host cell comprises delivering a vector comprising a ZFN system using a delivery vector. In some embodiments, the delivery vector is a virus vector. In some embodiments, the delivery vector is a lentivirus. In some embodiments, the delivery vector is an adenovirus.

[0181] Generally, a meganuclease system comprises one or more meganucleases that allow for recognition of a DNA sequence and subsequent double-stranded DNA cleavage.

[0182] In some embodiments, the method of producing a host cell, wherein the host cell has a reduced level of CpD expression, comprises inactivating the CpD gene using a meganuclease system. In some embodiments, the meganuclease has a DNA recognition sequence that is about 8 to about 35 nucleotide base pairs in length. In some embodiments, the meganuclease has a DNA recognition sequence that is about 12 to about 30 nucleotide base pairs in length. In some embodiments, the DNA

recognition sequence is a sequence comprising a portion of a CpD gene. In some embodiments, the DNA recognition sequence is a sequence comprising a portion of a CpD gene promoter. In some embodiments, the DNA recognition sequence is a sequence comprising a portion of a sequence flanking a CpD gene.

[0183] In some embodiments, the method of producing a host cell comprises delivering a meganuclease system to the host cell. In some embodiments, the method of producing a host cell comprises delivering a vector comprising a meganuclease system using a delivery vector. In some embodiments, the delivery vector is a virus vector. In some embodiments, the delivery vector is a lentivirus. In some embodiments, the delivery vector is an adenovirus.

[0184] In some embodiments, the method of producing a host cell, wherein the host cell has a reduced level of CpD expression, further comprises determining a level of CpD gene inactivation, such as gene deletion or gene addition or substitution. In some embodiments, the method of producing a host cell comprises determining the level of CpD gene inactivation, wherein a CpD gene deletion is detected. In some embodiments, the method of producing a host cell comprises determining the level of CpD gene inactivation, wherein a CpD gene addition or substitution is detected. In some embodiments, the method of producing a host cell comprises determining the level of CpD gene inactivation, wherein a level of CpD expression is determined prior to inactivating a gene in a host cell. In some embodiments, the method of producing a host cell comprises determining the level of CpD gene inactivation, wherein a level of CpD expression is determined after using a CRISPR system to inactivate the CpD gene in a host cell. In some embodiments, the method of producing a host cell comprises determining the level of CpD gene inactivation, wherein a level of CpD expression is determined after using a TALEN system to inactivate the CpD gene in a host cell. In some embodiments, the method of producing a host cell comprises determining the level of CpD gene inactivation, wherein a level of CpD expression is determined after using a ZFN system to inactivate the CpD gene in a host cell. In some embodiments, the method of producing a host cell comprises determining the level of CpD gene inactivation, wherein a level of CpD expression is determined after using a meganuclease system to inactivate the CpD gene in a host cell.

[0185] In some embodiments, the CpD expression level is determined at the DNA level. In some embodiments, the CpD expression level is determined at the RNA level. In some embodiments, the method of producing a host cell comprises

determining the CpD expression level using PCR. In some embodiments, the method of producing a host cell comprises determining the CpD expression level using PCR, wherein a variant sequence is detected. In some embodiments, the method of producing a host cell comprises determining the CpD expression level using qPCR. In some embodiments, the method of producing a host cell comprises determining the CpD expression level using qPCR.

[0186] In some embodiments, the CpD expression level is determined at the protein level. In some embodiments, the method of producing a host cell comprises determining the CpD expression level using immunohistochemistry. In some embodiments, the method of producing a host cell comprises determining the CpD expression level using Western blot. In some embodiments, the method of producing a host cell comprises determining the CpD expression level using flow cytometry.

[0187] In some embodiments, the level of CpD gene inactivation is determined by comparing the level of CpD expression after CpD gene inactivation to a control value. In some embodiments, the level of CpD gene inactivation is determined by comparing a level of CpD expression after after CpD gene inactivation to a level of CpD expression prior to CpD gene inactivation.

[0188] Also provided are methods of evaluating a host cell for suitability of expression of Fc-containing proteins comprising determining CpD expression, wherein a reduced level of CpD expression is indicative of suitability.

Methods of making a Fc-containing protein

[0189] The present application provides methods of making Fc-containing proteins described in the embodiments herein.

[0190] In some embodiments, the method of making a Fc-containing protein comprises: a) culturing the host cell; and b) obtaining the Fc-containing protein expressed by the host cell.

[0191] In some embodiments, the method of making a Fc-containing protein comprises: a) transforming a host cell with an expression vector comprising a nucleic acid encoding the Fc-containing protein; b) culturing the host cell; and c) obtaining the Fc-containing protein expressed by the host cell.

[0192] Methods for transforming a host cell using an expression vector are well known in the art. *See*, for example, Kim *et al.*, *Anal Bioanal Chem*, 397, 2010. Method for transfecting a host cell including, but are not limited to, transfection,

infection, calcium phosphate co-precipitation, electroporation, microinjection, lipofection, DEAE-dextran mediated transfection, or other known techniques. The method selected will in part be a function of the type of host cell to be used. These methods and other suitable methods are well known to the skilled artisan.

[0193] Expression systems and constructs in the form of plasmids, expression vectors, transcription or expression cassettes which comprise at least one Fc-containing protein described herein are also provided. In certain embodiments, a plasmid, expression vector, transcription or expression cassette provided herein comprises a polynucleotide encoding at least one Fc-containing protein.

[0194] In some embodiments, expression vectors used in the host cells will contain sequences for plasmid maintenance and for cloning and expression of exogenous nucleotide sequences. Such sequences, collectively referred to as "flanking sequences," in certain embodiments will typically include one or more of the following nucleotide sequences: a promoter, one or more enhancer sequences, an origin of replication, a transcriptional termination sequence, a complete intron sequence containing a donor and acceptor splice site, a sequence encoding a leader sequence for polypeptide secretion, a ribosome binding site, a polyadenylation sequence, a polylinker region for inserting the nucleic acid encoding the Fc-containing protein to be expressed, and a selectable marker element. Each of these sequences is discussed below.

[0195] Flanking sequences may be homologous (i.e., from the same species and/or strain as the host cell), heterologous (i.e., from a species other than the host cell species or strain), hybrid (i.e., a combination of flanking sequences from more than one source), synthetic or native. As such, the source of a flanking sequence may be any eukaryotic organism, any vertebrate or invertebrate organism, or any plant, provided that the flanking sequence is functional in, and can be activated by, the host cell machinery.

[0196] Flanking sequences useful in the vectors of this invention may be obtained by any of several methods well known in the art. Typically, flanking sequences useful herein will have been previously identified by mapping and/or by restriction endonuclease digestion and can thus be isolated from the proper tissue source using the appropriate restriction endonucleases. In some cases, the full nucleotide sequence of a flanking sequence may be known. Here, the flanking sequence may be synthesized using the methods described herein for nucleic acid synthesis or cloning.

[0197] Whether all or only a portion of the flanking sequence is known, it may be obtained using polymerase chain reaction (PCR) and/or by screening a genomic library with a suitable probe such as an oligonucleotide and/or flanking sequence fragment from the same or another species. Where the flanking sequence is not known, a fragment of DNA containing a flanking sequence may be isolated from a larger piece of DNA that may contain, for example, a coding sequence or even another gene or genes. Isolation may be accomplished by restriction endonuclease digestion to produce the proper DNA fragment followed by isolation using agarose gel purification, Qiagen® column chromatography (Chatsworth, CA), or other methods known to the skilled artisan. The selection of suitable enzymes to accomplish this purpose will be readily apparent to one of ordinary skill in the art.

[0198] An origin of replication is typically a part expression vectors purchased commercially, and the origin aids in the amplification of the vector in a host cell. If the vector of choice does not contain an origin of replication site, one may be chemically synthesized based on a known sequence, and ligated into the vector. For example, various viral origins (e.g., SV40, polyoma, adenovirus, vesicular stomatitis virus (VSV), or papillomaviruses such as HPV or BPV) are useful for cloning vectors in mammalian cells. Generally, the origin of replication component is not needed for mammalian expression vectors (for example, the SV40 origin is often used only because it also contains the virus early promoter).

[0199] A transcription termination sequence is typically located 3' to the end of a polypeptide coding region and serves to terminate transcription. While the sequence is easily cloned from a library or even purchased commercially as part of a vector, it can also be readily synthesized using methods for nucleic acid synthesis such as those described herein.

[0200] A selectable marker gene encodes a protein necessary for the survival and growth of a host cell grown in a selective culture medium. Typical selection marker genes encode proteins that (a) confer resistance to antibiotics or other toxins; (b) complement auxotrophic deficiencies of the cell; or (c) supply critical nutrients not available from complex or defined media. Specific selectable markers are the kanamycin resistance gene, the ampicillin resistance gene, and the tetracycline resistance gene. Advantageously, a neomycin resistance gene may also be used for selection in eukaryotic host cells.

[0201] Other selectable genes may be used to amplify the gene that will be expressed. Amplification is the process wherein genes that are required for production of a protein critical for growth or cell survival are reiterated in tandem within the chromosomes of successive generations of recombinant cells. Examples of suitable selectable markers for mammalian cells include dihydrofolate reductase (DHFR) and promoterless thymidine kinase genes. Mammalian cell transformants are placed under selection pressure wherein only the transformants are uniquely adapted to survive by virtue of the selectable gene present in the vector. Selection pressure is imposed by culturing the transformed cells under conditions in which the concentration of selection agent in the medium is successively increased, thereby leading to the amplification of both the selectable gene and the DNA that encodes another gene, such as an antibody light or heavy chain. As a result, increased quantities of a polypeptide are synthesized from the amplified DNA.

[0202] A ribosome-binding site is usually necessary for translation initiation of mRNA and is characterized by, for example, a Kozak sequence. The element is typically located 3' to the promoter and 5' to the coding sequence of the polypeptide to be expressed. In certain embodiments, one or more coding regions may be operably linked to an internal ribosome binding site (IRES), allowing translation of two open reading frames from a single RNA transcript.

[0203] Expression and cloning vectors of the invention will typically contain a promoter that is recognized by the host organism and operably linked to the molecule encoding the polypeptide. Promoters are untranscribed sequences located upstream (i.e., 5') to the start codon of a structural gene (generally within about 100 to 1000 bp) that control transcription of the structural gene. Promoters are conventionally grouped into one of two classes: inducible promoters and constitutive promoters. Inducible promoters initiate increased levels of transcription from DNA under their control in response to some change in culture conditions, such as the presence or absence of a nutrient or a change in temperature. Constitutive promoters, on the other hand, uniformly transcribe gene to which they are operably linked, that is, with little or no control over gene expression. A large number of promoters, recognized by a variety of potential host cells, are well known. A suitable promoter is operably linked to the DNA encoding e.g., heavy chain or light chain, by removing the promoter from the source DNA by restriction enzyme digestion and inserting the desired promoter sequence into the vector.

[0204] Suitable promoters for use with yeast hosts are also well known in the art. Yeast enhancers are advantageously used with yeast promoters. Suitable promoters for use with mammalian host cells are well known and include, but are not limited to, those obtained from the genomes of viruses such as polyoma virus, fowlpox virus, adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, retroviruses, hepatitis-B virus and most preferably Simian Virus 40 (SV40). Other suitable mammalian promoters include heterologous mammalian promoters, for example, heat-shock promoters and the actin promoter.

[0205] Additional promoters which may be of interest include, but are not limited to: SV40 early promoter (Benoist and Chambon, 1981, *Nature* 290:304-310); CMV promoter (Thomsen et al., 1984, *Proc. Natl. Acad. U.S.A.* 81:659-663); the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto et al., 1980, *Cell* 22:787-797); herpes thymidine kinase promoter (Wagner et al., 1981, *Proc. Natl. Acad. Sci. U.S.A.* 78:1444-1445); promoter and regulatory sequences from the metallothionein gene Prinster et al., 1982, *Nature* 296:39-42; or the tac promoter (DeBoer et al., 1983, *Proc. Natl. Acad. Sci. U.S.A.* 80:21-25). Also of interest are the following animal transcriptional control regions, which exhibit tissue specificity and have been utilized in transgenic animals: the elastase I gene control region that is active in pancreatic acinar cells (Swift et al., 1984, *Cell* 38:639-646; Omitz et al., 1986, *Cold Spring Harbor Symp. Quant. Biol.* 50:399-409; MacDonald, 1987, *Hepatology* 7:425-515); the insulin gene control region that is active in pancreatic beta cells (Hanahan, 1985, *Nature* 315: 115-122); the immunoglobulin gene control region that is active in lymphoid cells (Grosschedl et al., 1984, *Cell* 38:647-658; Adames et al., 1985, *Nature* 318:533-538; Alexander et al., 1987, *Mol. Cell. Biol.* 7:1436-1444); the mouse mammary tumor virus control region that is active in testicular, breast, lymphoid and mast cells (Leder et al., 1986, *Cell* 45:485-495); the albumin gene control region that is active in liver (Pinkert et al., 1987, *Genes and Devel.* 1:268-276); the alpha-feta-protein gene control region that is active in liver (Krumlauf et al., 1985, *Mol. Cell. Biol.* 5:1639-1648; Hammer et al., 1987, *Science* 253:53-58); the alpha 1-antitrypsin gene control region that is active in liver (Kelsey et al., 1987, *Genes and Devel.* 1: 161-171); the beta-globin gene control region that is active in myeloid cells (Mogam et al., 1985, *Nature* 315:338-340; Kollias et al., 1986, *Cell* 46:89-94); the myelin basic protein gene control region that is active in oligodendrocyte cells in the brain (Readhead et al., 1987, *Cell* 48:703-712); the

myosin light chain-2 gene control region that is active in skeletal muscle (Sani, 1985, Nature 314:283-286); and the gonadotropic releasing hormone gene control region that is active in the hypothalamus (Mason et al., 1986, Science 234: 1372-1378).

[0206] An enhancer sequence may be inserted into the vector to increase transcription of DNA encoding light chain or heavy chain of the invention by higher eukaryotes. Enhancers are cis-acting elements of DNA, usually about 10-300 bp in length, that act on the promoter to increase transcription. Enhancers are relatively orientation and position independent, having been found at positions both 5' and 3' to the transcription unit. Several enhancer sequences available from mammalian genes are known (e.g., globin, elastase, albumin, alphafeto-protein and insulin). Typically, however, an enhancer from a virus is used. The SV40 enhancer, the cytomegalovirus early promoter enhancer, the polyoma enhancer, and adenovirus enhancers known in the art are exemplary enhancing elements for the activation of eukaryotic promoters. While an enhancer may be positioned in the vector either 5' or 3' to a coding sequence, it is typically located at a site 5' from the promoter. A sequence encoding an appropriate native or heterologous signal sequence (leader sequence or signal peptide) can be incorporated into an expression vector, to promote extracellular secretion of the antibody. The choice of signal peptide or leader depends on the type of host cells in which the antibody is to be produced, and a heterologous signal sequence can replace the native signal sequence. Examples of signal peptides that are functional in mammalian host cells include the following: the signal sequence for interleukin-7 (IL-7) described in US Patent No. 4,965,195; the signal sequence for interleukin-2 receptor described in Cosman et al., 1984, Nature 312:768; the interleukin-4 receptor signal peptide described in EP Patent No. 0367 566; the type I interleukin-I receptor signal peptide described in U.S. Patent No. 4,968,607; the type II interleukin- I receptor signal peptide described in EP Patent No. 0 460 846.

[0207] The vector may contain one or more elements that facilitate expression when the vector is integrated into the host cell genome. Examples include an EASE element (Aldrich et al. 2003 Biotechnol Prog. 19: 1433-38) and a matrix attachment region (MAR). MARs mediate structural organization of the chromatin and may insulate the integrated vector from "position" effect. Thus, MARs are particularly useful when the vector is used to create stable transfectants. A number of natural and synthetic MAR-containing nucleic acids are known in the art, e.g., U.S. Pat. Nos. 6,239,328;

7,326,567; 6,177,612; 6,388,066; 6,245,974; 7,259,010; 6,037,525; 7,422,874; 7,129,062.

[0208] Expression vectors provided by the invention may be constructed from a starting vector such as a commercially available vector. Such vectors may or may not contain all of the desired flanking sequences. Where one or more of the flanking sequences described herein are not already present in the vector, they may be individually obtained and ligated into the vector. Methods used for obtaining each of the flanking sequences are well known to one skilled in the art.

[0209] After the vector has been constructed and a nucleic acid molecule encoding a Fc-containing protein sequence has been inserted into the proper site of the vector, the completed vector may be inserted into a suitable host cell for amplification and/or polypeptide expression.

[0210] Methods for making a vector comprising a nucleic acid encoding a Fc-containing protein are well known in the art. *See, e.g.,* U.S. Patent No. 7,923,221.

[0211] Construction of suitable vectors comprising a Fc-containing protein and the desired coding and control sequences employ standard ligation techniques. Isolated plasmids or DNA fragments are cleaved, tailored, and religated in the form desired to form the plasmids required. The methods employed are not dependent on the DNA source, or intended host.

[0212] In some embodiments, the method of making a Fc-containing protein further comprises determining an optimal ratio of the polynucleotide for introduction into a host cell. In some embodiments, mass spectrometry is used to determine Fc-containing protein yield, and the ratio is adjusted to maximize Fc-containing protein yield. In some embodiments, dual antigen ELISA is used to determine Fc-containing protein yield, and the ratio is adjusted to maximize Fc-containing protein yield.

[0213] Methods for culturing a host cell are well known to those in the art. *See, e.g.,* Li *et al.*, *MAbs*, 2, 2010. The host cells used to produce a desired Fc-containing protein of the embodiments herein may be cultured in a variety of media.

Commercially available media such as Ham's F10 (Sigma), Minimal Essential Medium ((MEM), (Sigma), RPMI-1640 (Sigma), and Dulbecco's Modified Eagle's Medium ((DMEM), Sigma) are suitable for culturing the host cells. In addition, any of the media described in Ham *et al.*, *Meth Enz*, 58, 1979, Barnes *et al.*, *Anal Biochem*, 102, 1980, U.S. Patent Nos. 4,767,704; 4,657,866; 4,927,762; 4,560,655; or 5,122,469; International Patent Application Nos. WO 90/03430 or WO 87/00195; or

U.S. Patent Reissue No. 30,985 may be used as culture media for the host cells. Any of these media may be supplemented as necessary with hormones and/or other growth factors (such as insulin, transferrin, or epidermal growth factor), salts (such as sodium chloride, calcium, magnesium, and phosphate), buffers (such as HEPES), nucleotides (such as adenosine and thymidine), antibiotics (such as GENTAMYCIN™ drug), trace elements (defined as inorganic compounds usually present at final concentrations in the micromolar range), and glucose or an equivalent energy source. Any other necessary supplements may also be included at appropriate concentrations that would be known to those skilled in the art. The culture conditions, such as temperature, pH, and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

[0214] Generally, the production of Fc-containing proteins is done on a large scale (such as a commercial scale). To achieve a population of a host cell suitable for commercial scale production, one of ordinary skill in the art will recognize the utility using a stepwise approach to expanding a host cell population. For example, the process involves growing a desired host cell on a smaller scale to allow for an increase in the host cell population, such as a seed train. To further increase the population of the host cell, methods generally involved using the seed train to inoculate a larger culture tank, such as an inoculum tank. Often, a series of inoculum tanks of increasing size are used to expand the population of a host cell, such as an inoculum train. This process will provide a suitable population of a host cell for culture in a production culture. In some embodiments, the production culture is a 1000L culture tank.

[0215] In some embodiments, the method of making a Fc-containing protein comprises culturing the host cell using a batch feed method. In some embodiments, the method of making a Fc-containing protein comprises culturing the host cell using a continuous feed method. In some embodiments, the method of making a Fc-containing protein comprises culturing the host cell using a feed method comprising a batch feed method and a continuous feed method.

[0216] Methods for obtaining a Fc-containing protein are well known in the art. See, e.g., Huse *et al.*, *J Biochem Bioph Meth*, 51, 2002. Fc-containing proteins can be produced intracellularly or directly secreted into the medium. If the Fc-containing proteins are produced intracellularly, as a first step, the particulate debris, either host cells or lysed fragments, are removed, for example, by centrifugation or ultrafiltration.

If the Fc-containing proteins are secreted into the medium, supernatants from such expression systems are generally first concentrated using a commercially available protein concentration filter, for example, an Amicon or Millipore Pellicon ultrafiltration unit. In some embodiments, a protease inhibitor may be included in any of the foregoing steps to inhibit proteolysis and antibiotics may be included to prevent the growth of adventitious contaminants.

[0217] The composition prepared from the cells can be purified using, for example, hydroxylapatite chromatography, gel electrophoresis, dialysis, and affinity chromatography, with affinity chromatography being the preferred purification technique. The suitability of protein A as an affinity ligand depends on the species and isotype of any immunoglobulin Fc domain that is present in the Fc-containing protein. Protein A can be used to purify Fc-containing proteins that are based on human immunoglobulins containing 1, 2, or 4 heavy chains (*See, e.g., Lindmark et al., J Immunol Meth, 62, 1983*). Protein G is recommended for all mouse isotypes and for human 3 (*See, e.g., Guss et al., EMBO, 5, 1986*). The matrix to which the affinity ligand is attached is often agarose, but other matrices are available. Mechanically stable matrices such as controlled pore glass or poly(styrene-divinyl)benzene allow for faster flow rates and shorter processing times than can be achieved with agarose. Other techniques for protein purification such as fractionation on an ion-exchange column, ethanol precipitation, Reverse Phase HPLC, chromatography on silica, chromatography on heparin SEPHAROSE™ chromatography on an anion or cation exchange resin (such as a polyaspartic acid column), chromatofocusing, SDS-PAGE, and ammonium sulfate precipitation are also available depending on the Fc-containing protein to be recovered.

[0218] In some embodiments, the method of purifying a Fc-containing protein comprises using a filter. In some embodiments, the filter is a diafiltration system. In some embodiments, the filter is an ultrafiltration system. In some embodiments, the filter is a viral filtration system. In some embodiments, the purifying a Fc-containing protein comprises using a series of filtration steps. In some embodiments, the series of filtration steps is selected from at least one of the following: diafiltration, ultrafiltration, and viral filtration.

[0219] In some embodiments, the method of purifying the Fc-containing protein comprises using a series of protein purification techniques selected from filtration, protein A purification, cation exchange purification, strong cation exchange

purification, anion exchange purification, reverse phase purification, and multimodal purification.

[0220] Also provided in the present application are methods of determining C-terminal lysine presence. In some embodiments, the method of determining C-terminal lysine presence comprises using isoelectric focusing, such as imaged capillary isoelectric focusing. In some embodiments, the method of determining C-terminal lysine presence comprises using mass spectrometry.

Methods of treatment

[0221] The present application provides methods of treating a disease in an individual in need thereof comprising administering to the individual a pharmaceutical composition described in the embodiments herein.

[0222] In some embodiments, an effective amount of a plurality of Fc-containing proteins is administered to the individual in need thereof, wherein substantially all of the plurality of Fc-containing proteins has a C-terminal lysine on each Fc domain.

[0223] The pharmaceutical composition described herein can be administered via various routes, such as parenterally, including intravenous, intra-arterial, intraperitoneal, intrapulmonary, oral, inhalation, intravesicular, intramuscular, intra-tracheal, subcutaneous, intraocular, intrathecal, or transdermal. In some embodiments, pharmaceutical composition described herein can be administered parenterally. In some embodiments, pharmaceutical composition described herein can be administered intravenously. In some embodiments, pharmaceutical composition described herein can be administered subcutaneously. In some embodiments, pharmaceutical composition described herein can be administered locally. In some embodiments, pharmaceutical composition described herein can be administered topically.

[0224] The diseases that can be treated by the methods herein are any disease that can be treated with a Fc-containing protein. In some embodiments, the disease is cancer. In some embodiments, the disease is an autoimmune disease. In some embodiments, the disease is an infection.

[0225] In some embodiments, the pharmaceutical composition described herein is used in combination with another administration modality or treatment.

[0226] Those skilled in the art will recognize that several embodiments are possible within the scope and spirit of this invention. The invention will now be described in

greater detail by reference to the following non-limiting examples. The following examples further illustrate the invention but, of course, should not be construed as in any way limiting its scope.

EXAMPLES

Example 1:

Carboxypeptidase D is responsible for C-terminal lysine cleavage in Chinese hamster ovary (CHO) cells

Methods:

[0227] Cell line A is an in-house developed antibody-producing (IgG1) cell line derived from DUXB-11-based DHFR deficient DP12 host. *See*, Hu et al., *Biotechnol Prog*, 29, 2013; Urlaub et al., *Proc Natl Acad Sci*, 77, 1980. Cell line B is an in-house developed antibody-producing (IgG1) cell line derived from CHOK1 host. *See*, Hu et al., *Biotechnol Prog*, 29, 2013. CHO cells were cultured in a proprietary DMEM/F12-based medium in 125 ml shake flask vessels at 150 rpm, 37 °C and 5% CO₂. Cells were passaged with a seeding density of 3x10⁵/mL every three to four days.

[0228] Total RNA was isolated using a RNeasy 96 kit (Cat# 74181, Qiagen, Valencia, CA) and treated with DNase digestion (Cat# 79254, RNase free DNase kit, Qiagen, Valencia, CA) to remove any residual DNA that may be present in the isolated RNA sample. Taqman was performed using a universal qRT-PCR master mixture according to the manufacturer's instruction (Cat# 4309169, Applied Biosystems, Foster City, CA). Expression levels of different carboxypeptidases were normalized to the housekeeping gene, β -2-microglobulin (β 2m).

[0229] Primer and probe sequences used for Taqman analysis were as follows:

CpD forward primer: CCC ACA CAT TAC AAA TCT TAC CA (SEQ ID NO. 1);

CpD reserve primer: GAG ATT TCG AGG GAC CAA AT (SEQ ID NO. 2);

CpD probe: TTG GGA CAG AGT GCT GAG TAT CGT CA (SEQ ID NO. 3);

CpN forward primer: GTG GGA TCA ATC ACG ATG TC (SEQ ID NO. 4);

CpN reserve primer: CCT TGG CAG TGA CAA TGT AAG TA (SEQ ID NO. 5);

CpN probe: ACA TGG GGA TTA CTT CCG TCT GCT G (SEQ ID NO. 6);

CpM forward primer: AAC TTG GAG AGT ACT ACC TGC TTC T (SEQ ID NO. 7);

CpM reserve primer: TCA TGC CCA GGG ACT GTA (SEQ ID NO. 8);

CpM probe: ATT GAT CAC GTA GGA CCC TGG CAA A (SEQ ID NO. 9);

CpB forward primer: TGA ATG CGC TGG TGA AAG G (SEQ ID NO. 10);

CpB reserve primer: TCC TGG GCC ATA TGT GTA CTT G (SEQ ID NO. 11);

CpB probe: CGG TCA AGG AAC TTG CCT CTC TGC A (SEQ ID NO. 12);

CpE forward primer: TGG CTA CCT GGC AAT AAC AA (SEQ ID NO. 13);

CpE reserve primer: CGA CTC CAG CTC AAA GTC AA (SEQ ID NO. 14);

CpE probe: AAG TGG CAG TTC CTT TCA AGC CTG C (SEQ ID NO. 15);

β -microglobulin forward primer: TCC TCT CAG TGG TCT GCT TGG (SEQ ID NO. 16);

β -microglobulin reserve primer: TGG CGT GTG TAG ACT TGC ACT T (SEQ ID NO. 17);

β -microglobulin probe: TGC CAT CCA GCG TCC CCC A (SEQ ID NO. 18).

[0230] All primers and probes were synthesized and purified at Genentech (Redwood City, CA).

[0231] Western blot analysis was performed to determine the relative carboxypeptidase D protein expression levels in CHO cells. β -actin, a cytosolic protein, was used as the sample loading control. Cell pellets were collected and lysed in 1X cell lysis buffer (Cat# 9803, Cell Signaling Technology, Danvers, MA). Total protein of the lysed supernatants was quantified by Bradford assay (Cat# 1856210,

Thermo Scientific, Rockford, IL), and heated at 95 °C for 3 minutes before loading onto a reducing SDS–PAGE gel. Separated proteins were transferred to a nitrocellulose membrane by electroblotting. Transferred membranes were then incubated with a rabbit polyclonal antibody against carboxypeptidase D (Cat# SAB2700486, Sigma, Saint Louis, MO) to detect CpD protein expression and a mouse antibody (Cat# A2228, Sigma, Saint Louis, MO) against β -actin as a loading control for equal sample loading.

[0232] Designer of Small Interfering RNA (DSIR) program (<http://www.biomedcentral.com>) was used to design CpD and CpN-specific small interfering RNA (siRNA) sequences:

CpD siRNA target sequence: GGAAGAGAACTGCTACTAA (SEQ ID NO. 19);

CpN siRNA target sequence: GAATGGTGCTTGATGAGAA (SEQ ID NO. 20).

[0233] The above siRNA sequences were synthesized and individually cloned into the pSilencer 3.1-H1vector (Cat# AM5766, Ambion, Austin, TX) to make siRNA expression constructs. Each construct was introduced to antibody-expressing cell line A and antibody-expressing cell line B using lipofectamine 2000 CD according to the manufacturer's recommendation (Invitrogen, Carlsbad, CA) in transient transfection. 24 hours post transfection, cell pellets were collected to assess each carboxypeptidase mRNA expression by Taqman. Additionally, Cell Culture Fluid (CCF) was collected and concentrated by Amicon Ultra-15 (30,000 MWCO, Millipore, Bedford, MA) for charge variant product quality analysis by Imaged Capillary Isoelectric Focusing (IcIEF).

[0234] Charge variant distribution was assessed by using an iCE280 analyzer (ProteinSimple) with a fluorocarbon coated capillary cartridge (100 μ m $\times$ 5cm). The ampholyte solution consisted of a mixture of 0.35% methyl cellulose (MC), 1.34% 3–10 carrier ampholytes, 1.34% 6.7–7.7 carrier ampholytes, pI markers (pI 6.61 and pI 9.22), and 10 mM L-arginine free base in purified water. The anolyte was 80 mM phosphoric acid, and the catholyte was 100 mM sodium hydroxide, both in 0.1% MC. Samples were diluted, mixed with the ampholyte solution, and then focused by introducing a potential of 1500 V for 1 minute, followed by a potential of 3000 V for 8 minutes. An image of the focused charge variants was obtained by passing 280 nm ultraviolet light through the capillary and into the lens of a charge coupled device

digital camera. In order to remove heavy chain C-terminal Lys residue, carboxypeptidase B was added to each sample at the dilution step at an enzyme-to-substrate ratio of 1:100 (w/w) followed by incubation at 37 °C for 20 minutes.

[0235] The following guide RNA (gRNA) sequences below were designed to sit on CpD exon 2 and exon 21, respectively:

CpD guide RNA1 sequence (gRNA 1) for exon 2: GAA GAC GAG ACT
TTC AAA GAC GG (SEQ ID NO. 21);

CpD guide RNA2 sequence (gRNA 2) for exon 21: TCA GTT AGG TGG
CTT AGG TTC GG (SEQ ID NO. 22).

[0236] When co-transfected along with caspase 9 (Cas9), the gRNA sequences are expected to enable a ~46 kb deletion encompassing a majority of the annotated CpD locus.

[0237] Individual gRNAs were cloned and transcribed under the control of the human U6 promoter of the pLKO.5 vector (Cat# SHC-201, Sigma, St. Louis, MO). The two gRNAs and Cas9 constructs were co-transfected at an equimolar ratio. 72 hours post transfection, cells were seeded into 384-well plates by limiting dilution. Three weeks later, colonies were picked and transferred to 96-well plates and screened by Polymerase Chain Reaction (PCR). PCR primers for screening CpD knockout (~46 kb deletion) and wild type alleles were as follows:

Forward primer for CpD knockout allele: AGT TCA TTT ATG AAA
GAT CCT GTG G (SEQ ID NO. 23);

Reverse primer for CpD knockout allele: GGA AAG GAG TCC TTC
AGT GAA CAC (SEQ ID NO. 24);

Forward primer for CpD wild type allele: CCA GTT CTG CTG TTA CAC
TTT GAG (SEQ ID NO. 25);

Reverse primer for CpD wild type allele: AAT GTT TCC TCT TTC CTG
GAC CTT (SEQ ID NO. 26);

[0238] Diluted protein samples (1 mg/mL) were mixed with a 20 mg/mL TCEP solution at 1:1 ratio and incubated for 10 minutes at 60 °C to reduce and dissociate antibodies into light and heavy chains. Each sample was then analyzed by liquid chromatography electrospray ionization mass spectrometry (LC-ESI-MS) using an Agilent 6210 time-of-flight mass spectrometer coupled with a nano-Chip LC-ESI source. Briefly, about 5 ng of protein samples was injected onto a custom chip with a 40 nL trap column for desalting and an equivalent analytical column dimensions of 43

mm x 75 μ m, Zorbax 300SB-C8 (5 μ m, Agilent Technologies, Santa Clara, CA) for separation at 400 nL/min. *See*, Lu et al., *Mabs*, 5, 2013. Mobile phase A was 0.1% formic acid in water and mobile phase B was 0.1% formic acid in acetonitrile. MS data was extracted and deconvoluted using Agilent MassHunter Quantitative Analysis Workstation Software B.04.00. To estimate the percentage of C-terminal lysine, the intensity of deconvoluted mass corresponding to antibody heavy chains with or without C-terminal lysine was used.

[0239] Fed-batch production was performed in shake flask vessels in proprietary chemically defined medium with bolus feeds on days 3, 7, and 10. A temperature shift from 37 °C to 35 °C was carried out on day 3. Day 14 titers were determined using protein A affinity chromatography with UV detection.

Results:

[0240] qPCR analysis was performed to measure mRNA levels of possible endogenous carboxypeptidases in DUXB-11 based DHFR deficient DP12, DHFR positive CHOK1 hosts, and antibody-producing cell lines from each host (cell line A and cell line B). Cell line A was derived from DHFR deficient DP12 host and cell line B was derived from DHFR positive CHOK1 host.

[0241] 5 carboxypeptidases, CpD, CpN, CpM, CpB, and CpE, were chosen for mRNA expression analysis by qPCR. CpD had the highest mRNA expression levels in both CHO hosts and their respective antibody-expressing lines (**FIG. 1A** and **FIG. 1B**). CpD and CpN mRNA levels were higher than those for CpM, CpE, and CpB in cell line A and cell line B (**FIG. 1B**).

[0242] siRNA-mediated knockdown (*See, e.g.*, Rao et al., *Adv Drug Deliv Rev*, 61, 2009) experiments were performed using the two antibody-producing cell lines with the two most abundant carboxypeptidases, namely, CpD and CpN. Each cell line was transiently transfected with a siRNA construct specifically targeting CpD (CpDi construct) or CpN (CpNi construct), respectively. A scrambled vector control construct was also transfected as a negative control. As shown in **FIG. 2A**, CpD mRNA levels were knocked down to 39% in cell line A and 28% in cell line B. Similar results were observed with regard to CpN mRNA inhibition by CpNi construct, where CpN mRNA level was reduced to 43% in cell line A and 35% in cell line B (**FIG. 2B**).

[0243] The percentage of C-terminal lysine present on the antibody in cell culture was calculated. To calculate the percentage of C-terminal lysine present on the

antibody in cell culture, supernatants from CpDi and CpNi transfected cell lines, with and without CpB treatment, were purified by protein-A affinity chromatography and analyzed by Imaged Capillary Isoelectric Focusing (ICIEF). The differential value between CpB treated and untreated samples represents the percentage of antibody that has intact C-terminal lysines. Significantly increased C-terminal lysine levels were observed in both cell lines when CpD mRNA was knocked down to approximately 30-40% (**FIG. 2A** and **FIG. 3**). Minimal C-terminal lysine level changes were observed in the two cell lines transfected with either the scrambled RNAi construct (negative control) or the CpNi construct which knocked down CpN mRNA to approximately 40% (**FIG. 2B** and **FIG. 3**). These results demonstrated that reduction in CpD mRNA levels led to increased C-terminal lysine levels, while changes in CpN mRNA levels had no effect on C-terminal lysine in the two tested antibody-producing lines.

[0244] A knockout experiment using CRISPR technology (*see, e.g.,* He et al., *Biotechnol Bioeng*, 2014; Jinek et al., *Elife*, 2, 2013) was performed. The known CHOK1 CpD genomic sequence in NCBI database is approximately 55 kb including most of the coding region sequences for CHO CpD. Two guide RNA sequences were designed to simultaneously target exon 2 and 21 of the Chinese hamster CpD gene aiming to delete the 46 kb sequence between the two exons as illustrated in **FIG. 4A**. When the two gRNA constructs and Cas9 endonuclease construct were co-transfected into the antibody-expressing cell line B, Cas9 endonuclease cleaves specific DNA sequences that are complementary to the gRNAs. This cleavage leads to double-strand DNA breaks and subsequent non-homologous end-joining (NHEJ) DNA repair with deletions or insertions that can abolish expression of a target gene. Using PCR primers adjacent and upstream of the gRNA sequences, CpD knockout cells should yield a PCR product of approximately 600 base pairs when the intervening sequences between the two gRNAs are deleted as illustrated in **FIG. 4B**. A total of 180 clones were screened for the CpD sequence deletion, and based on PCR product sizes and sequence analysis, two clones with both CpD alleles deleted were obtained and used for CpD mRNA and protein expression analysis (data not shown). As shown in **FIG. 5A**, CpD mRNA was undetectable in the two CpD KO clones by qPCR analysis. Furthermore, Western blot analysis using the anti-CpD antibody showed that a protein with the molecular weight of approximately 180 kDa was absent in two CpD KO clones, but present in CpD WT clones as shown in **FIG. 5B**.

[0245] CpD KO and WT clones were evaluated for productivity and product quality attributes. No obvious differences were observed in integrated viable cell count cell (IVCC), specific productivities (Qp), and titers among two KO and two WT clones (FIG. 6), indicating that absence of CpD expression in CHO cells did not impact cell productivity and growth as assessed by the fed-batch shake flask production assay. Importantly, C-terminal lysine level was measured by mass spectrometry-based analysis of reduced antibody showing that approximately 4-6% C-terminal lysine was observed in the two WT clones, whereas 100% of C-terminal lysine was present in the two KO clones (FIG. 7). Thus, it was demonstrated that CpD is the only carboxypeptidase which is responsible for removing C-terminal lysine of an antibody heavy chain in CHO cells. The same result was also observed by an IeIEF assay (data not shown). Other product quality attributes and cell culture metabolite profiles were comparable between CpD KO and WT clones (data not shown).

CLAIMS

What is claimed is:

1. A host cell for expression of a Fc-containing protein, wherein the host cell has at least a 60% reduction of a carboxypeptidase D (CpD) expression level as compared to a host cell comprising a wild type CpD gene without CpD gene inactivation.
2. The host cell of claim 1, wherein the host cell is a eukaryotic cell.
3. The host cell of claim 2, wherein the host cell is a mammalian cell.
4. The host cell of claim 3, wherein the host cell is a Chinese hamster ovary (CHO) cell.
5. The host cell of claim 1, wherein the Fc-containing protein is an antibody.
6. The host cell of claim 5, wherein the antibody is a monoclonal antibody.
7. The host cell of claim 1, wherein the Fc-containing protein is a Fc-containing fusion protein.
8. The host cell of any one of claims 1-7, wherein the CpD gene in the host cell is inactivated.
9. The host cell of claim 8, wherein the CpD gene is inactivated by siRNA.
10. The host cell of claim 8, wherein the CpD gene is inactivated by shRNA.
11. The host cell of claim 8, wherein the CpD gene is inactivated by gene deletion.
12. The host cell of claim 8, wherein the CpD gene is inactivated by gene addition or substitution.
13. The host cell of claim 11 or 12, wherein the CpD gene is inactivated using a clustered, regularly interspaced, short palindromic repeats (CRISPR) system.
14. The host cell of claim 11 or 12, wherein the CpD gene is inactivated using a transcription activator-like effector nuclease (TALEN) system.

15. The host cell of claim 11 or 12, wherein the CpD gene is inactivated using a zinc-finger nuclease (ZFN) system.
16. The host cell of claim 11 or 12, wherein the CpD gene is inactivated using a meganuclease system.
17. The host cell of any one of claims 1-16, comprising an expression vector comprising a nucleic acid encoding the Fc-containing protein.
18. A cell culture system comprising the host cell of any one of claims 1-17.
19. A method of producing a host cell of any one of claims 1-17, wherein the method comprises inactivating the CpD gene in the host cell, thereby producing the host cell.
20. The method of claim 19, wherein the CpD gene is inactivated using a siRNA system.
21. The method of claim 19, wherein the CpD gene is inactivated using a shRNA system.
22. The method of claim 19, wherein the CpD gene is inactivated by gene deletion.
23. The method of claim 19, wherein the CpD gene is inactivated by gene addition or substitution.
24. The method of claim 22 or 23, wherein the CpD gene is inactivated using a CRISPR system.
25. The method of claim 22 or 23, wherein the CpD gene is inactivated using a TALEN system.
26. The method of claim 22 or 23, wherein the CpD gene is inactivated using a ZFN system.
27. The method of claim 22 or 23, wherein the CpD gene is inactivated using a meganuclease system.

28. A method of making a Fc-containing protein comprising:
 - a) culturing the host cell of any one of claims 1-17;
 - b) obtaining the Fc-containing protein expressed by the host cell.
29. A composition comprising a plurality of Fc-containing proteins, wherein substantially all Fc-containing proteins in the composition have a C-terminal lysine on each Fc domain.
30. The composition of claim 29, wherein substantially all Fc-containing proteins of the composition have the same amino acid sequence.
31. The composition of claim 29 or 30, wherein the plurality of Fc-containing proteins are substantially homogeneous in charge state.
32. The composition of any one of claims 29-31, wherein the Fc-containing protein is a Fc-containing fusion protein.
33. The composition of any one of claims 29-31, wherein the Fc-containing protein is an antibody.
34. The composition of claim 33, wherein the antibody is a human antibody.
35. The composition of claim 33, wherein the antibody is a humanized antibody.
36. The composition of any one of claims 33-35, wherein the antibody comprises two heavy chains, and wherein each heavy chain comprises a C-terminal lysine.
37. The composition of any one of claims 32-36, wherein the Fc-containing protein is conjugated to a drug.
38. The composition of any one of claims 32-37, wherein Fc-containing proteins in the composition comprise an IgG1 Fc domain.
39. The composition of any one of claims 32-37, wherein the Fc-containing proteins in the composition comprise an IgG2 Fc domain.

40. The composition of any one of claims 32-37, wherein the Fc-containing proteins in the composition comprise an IgG4 Fc domain.
41. The composition of any one of claims 29-40, wherein the composition is a pharmaceutical composition.
42. The composition of claim 41, wherein the pharmaceutical composition is a sterile pharmaceutical composition.
43. The composition of any one of claims 29-40, wherein the composition is a cell culture medium.
44. The composition of any one of claims 29-40, wherein the composition is a cell lysate.
45. The composition of any one of claims 29-40, wherein the composition is an eluate from a protein purification column.
46. A method of treating a disease in an individual in need thereof comprising: administering to the individual the composition of claim 41 or 42.
47. The method of claim 46, wherein the composition is administered parenterally.
48. The method of claim 47, wherein the composition is administered intravenously or subcutaneously.
49. The method of claim 46, wherein the composition is administered locally.
50. The method of claim 49, wherein the composition is administered topically.

1/7

FIG. 1A

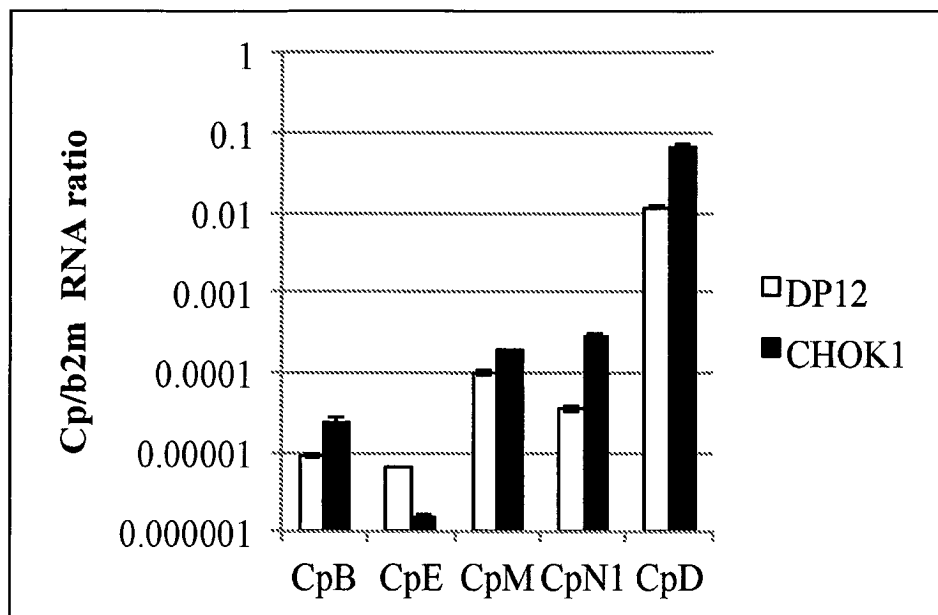
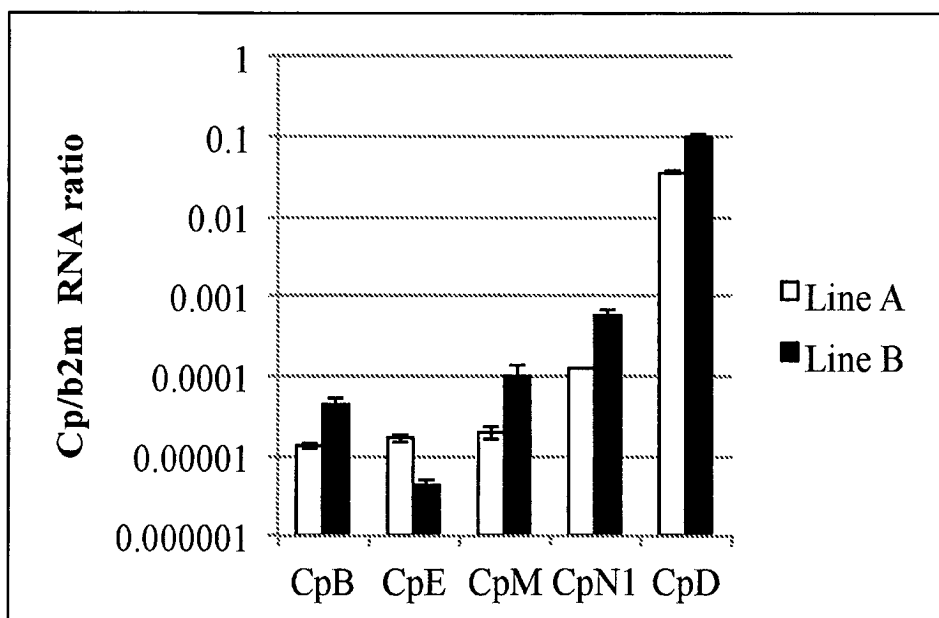


FIG. 1B



2/7

FIG. 2A

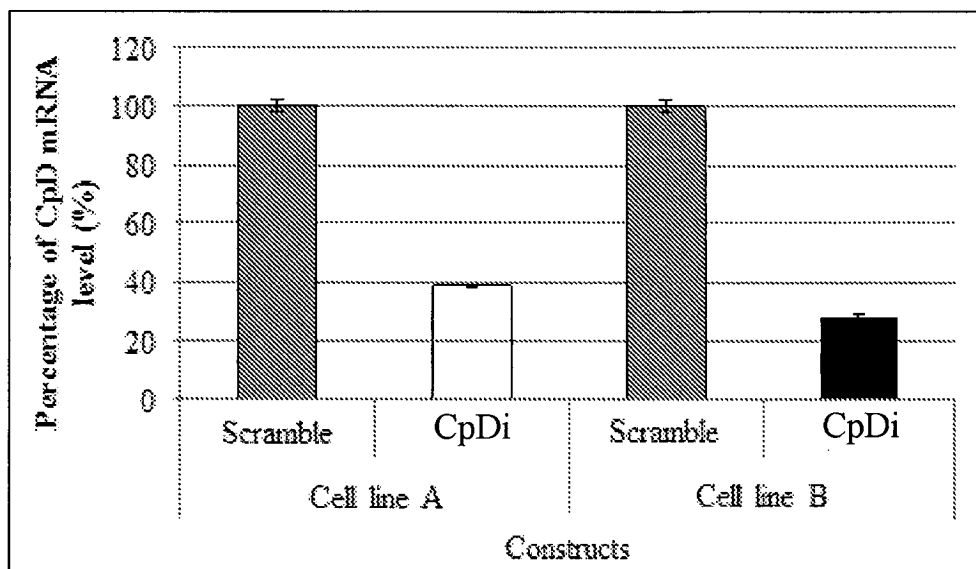
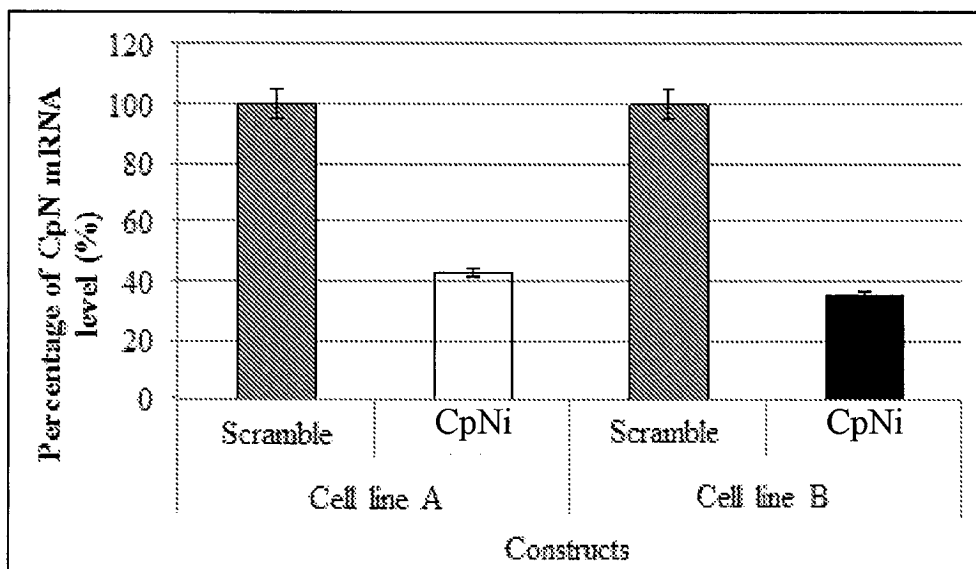


FIG. 2B



3/7

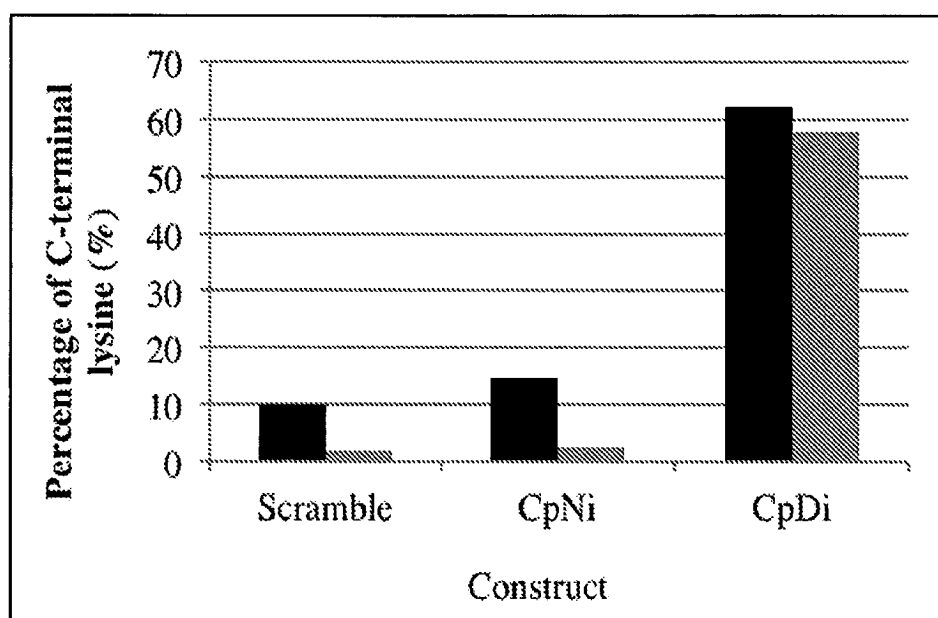


FIG. 3

FIG. 4A

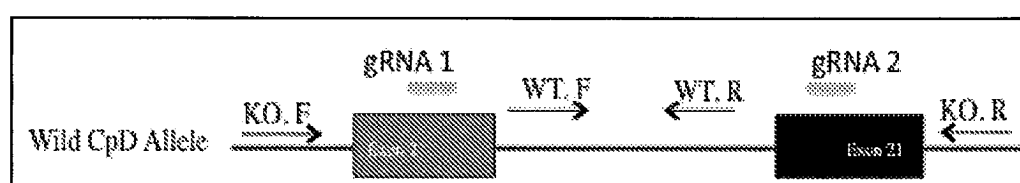


FIG. 4B

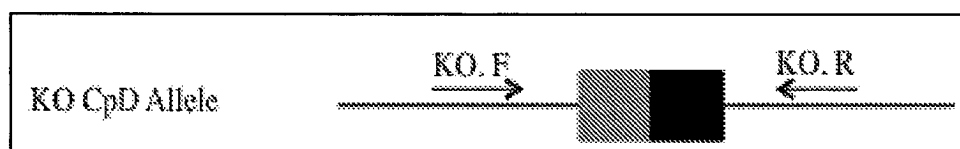


FIG. 5A

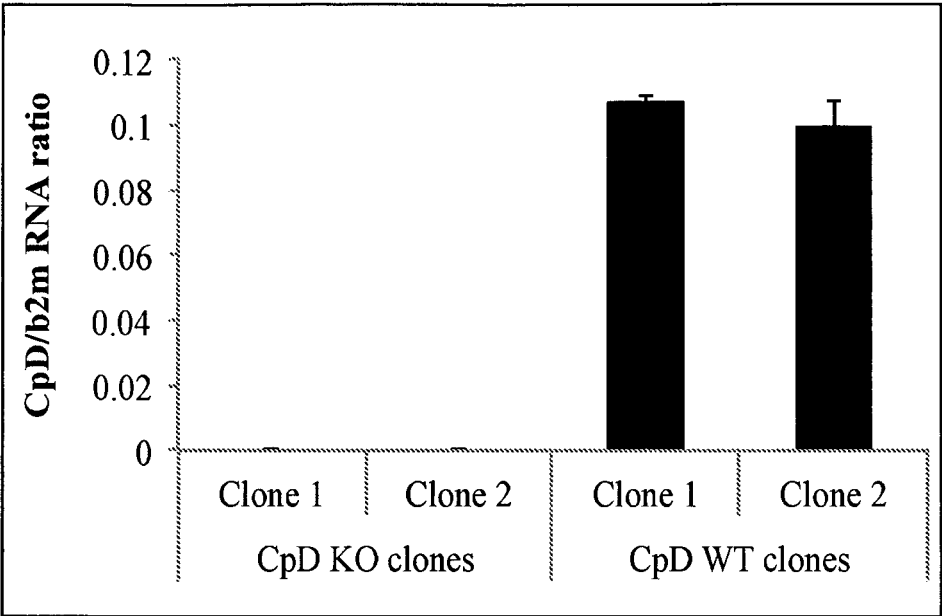
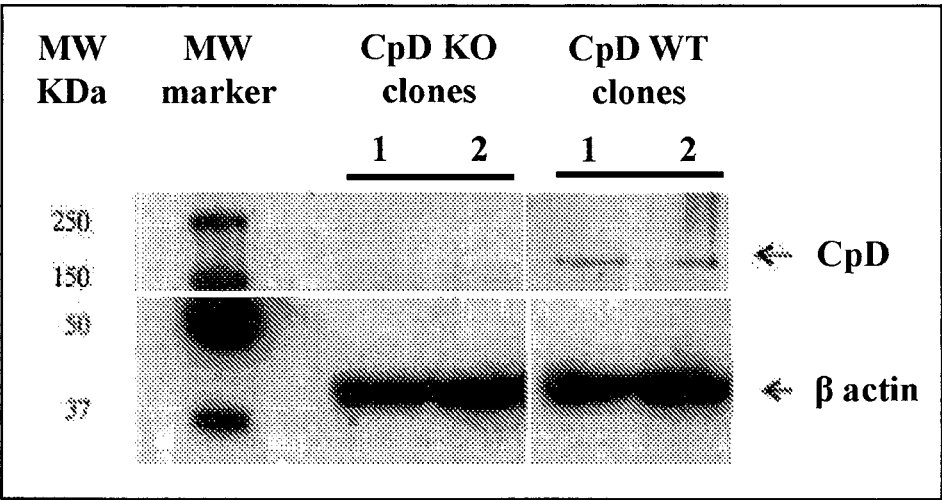


FIG. 5B



6/7

FIG. 6A

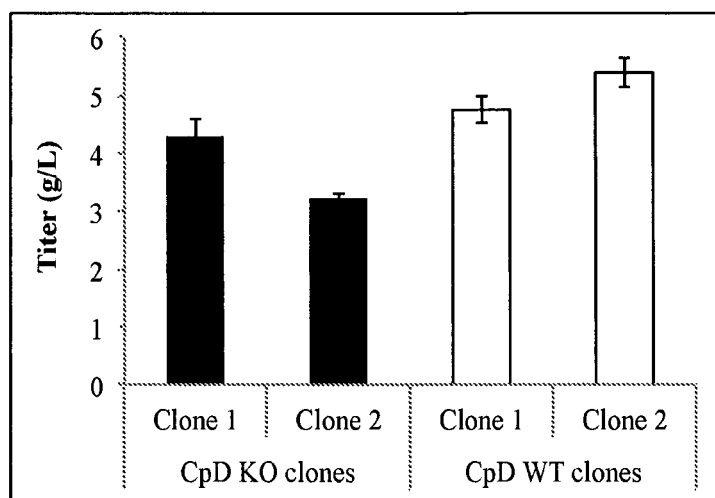


FIG. 6B

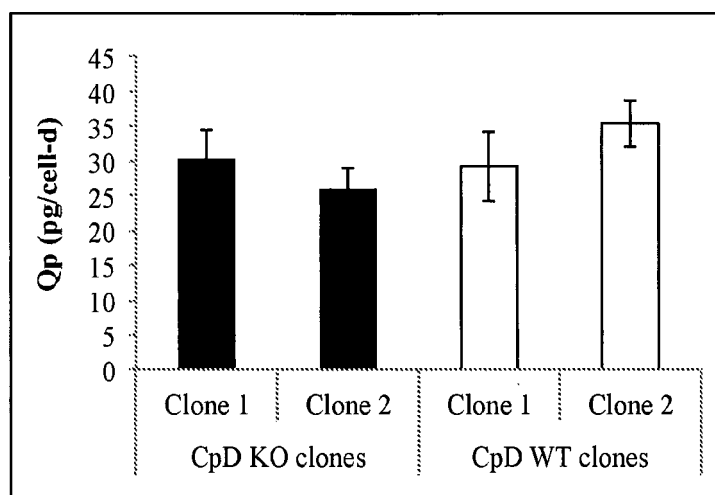
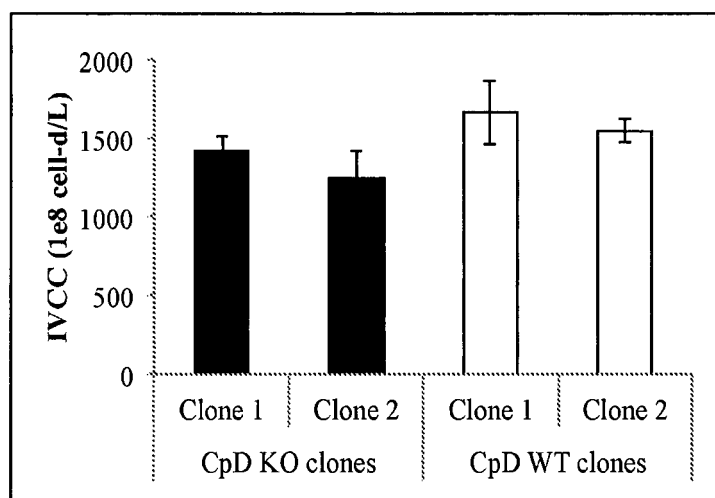


FIG. 6C



7/7

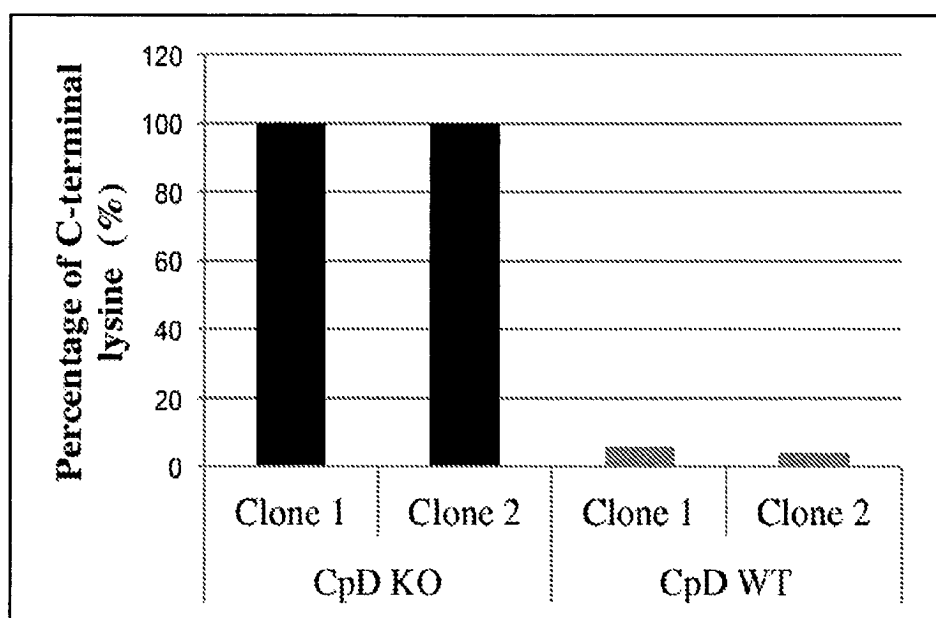


FIG. 7

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2016/052962

A. CLASSIFICATION OF SUBJECT MATTER

INV. C12N5/10 A61K39/00 C07K16/00 C12N15/79 C12N15/85
ADD.

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)

C12N A61K C07K

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

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

EPO-Internal, BIOSIS, EMBASE, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	TAO JIN ET AL: "SiRNA-targeted carboxypeptidase D inhibits hepatocellular carcinoma growth", CELL BIOLOGY INTERNATIONAL., vol. 37, no. 9, 30 April 2013 (2013-04-30) , pages 929-939, XP055323242, GB	1-28
A	ISSN: 1065-6995, DOI: 10.1002/cbin.10113 abstract page 933, left-hand column, paragraph 2 page 935; figure 3 ----- -/--	29-50

☒ 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

"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

28 November 2016

Date of mailing of the international search report

13/12/2016

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Grötzing, Thilo

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2016/052962

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2011/064282 A1 (NOVO NORDISK AS [DK]; ANDERSEN ASSER SLOTH [DK]; LAUTRUP-LARSEN INGER) 3 June 2011 (2011-06-03)	1-28
A	abstract page 2, line 9 - line 13 page 3, line 27 - line 33 page 4, line 23 - page 5, line 8 -----	29-50
X	US 2013/323785 A1 (SOHN JUNG HOON [KR] ET AL) 5 December 2013 (2013-12-05)	1-28
A	abstract page 6; example 4 page 7; example 5 figure 6 -----	29-50
X	WO 2014/159579 A1 (ABBVIE INC [US]) 2 October 2014 (2014-10-02)	29-50
A	abstract page 7, paragraph 7 - page 8, paragraph 1 -----	1-28
X	WO 2013/009648 A2 (MOMENTA PHARMACEUTICALS INC [US]; PRENTICE HOLLY [US]) 17 January 2013 (2013-01-17)	29-50
A	abstract page 31, paragraph 128 page 33, paragraphs 135,136 page 34, paragraph 137 page 35, paragraph 140 -----	1-28
X,P	SUMA SREENIVAS ET AL: "Disruption of KEX1 gene reduces the proteolytic degradation of secreted two-chain Insulin glargine in <i>Pichia pastoris</i> ", PROTEIN EXPRESSION AND PURIFICATION., vol. 118, 1 February 2016 (2016-02-01), pages 1-9, XP055259294, SAN DIEGO, CA. ISSN: 1046-5928, DOI: 10.1016/j.pep.2015.10.002 -----	1-28
A,P	abstract page 2, right-hand column, paragraph 3 page 6; figure 5b -----	29-50
X,P	ZHILAN HU ET AL: "Carboxypeptidase D is the only enzyme responsible for antibody C-terminal lysine cleavage in Chinese hamster ovary (CHO) cells", BIOTECHNOLOGY AND BIOENGINEERING., vol. 113, no. 10, 1 October 2016 (2016-10-01), pages 2100-2106, XP055323077, US ISSN: 0006-3592, DOI: 10.1002/bit.25977 the whole document -----	1-50

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2016/052962

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2011064282 A1	03-06-2011	CN 102639559 A EP 2504355 A1 ES 2550761 T3 JP 5850848 B2 JP 2013511967 A US 2012282653 A1 WO 2011064282 A1	15-08-2012 03-10-2012 12-11-2015 03-02-2016 11-04-2013 08-11-2012 03-06-2011
US 2013323785 A1	05-12-2013	KR 20120047838 A US 2013323785 A1 WO 2012060666 A2	14-05-2012 05-12-2013 10-05-2012
WO 2014159579 A1	02-10-2014	TW 201522371 A US 2014271626 A1 US 2015132320 A1 UY 35444 A WO 2014159579 A1	16-06-2015 18-09-2014 14-05-2015 31-10-2014 02-10-2014
WO 2013009648 A2	17-01-2013	CN 103717729 A EP 2729561 A2 US 2014120583 A1 US 2014142286 A1 WO 2013009648 A2	09-04-2014 14-05-2014 01-05-2014 22-05-2014 17-01-2013



(12)发明专利申请

(10)申请公布号 CN 108138147 A

(43)申请公布日 2018.06.08

(21)申请号 201680054785.8

(74)专利代理机构 北京市中咨律师事务所

(22)申请日 2016.09.21

11247

代理人 张莉 黄革生

(30)优先权数据

62/222,187 2015.09.22 US

(51)Int.Cl.

G12N 5/10(2006.01)

(85)PCT国际申请进入国家阶段日

A61K 39/00(2006.01)

2018.03.21

C07K 16/00(2006.01)

(86)PCT国际申请的申请数据

G12N 15/79(2006.01)

PCT/US2016/052962 2016.09.21

G12N 15/85(2006.01)

(87)PCT国际申请的公布数据

W02017/053482 EN 2017.03.30

(71)申请人 豪夫迈·罗氏有限公司

地址 瑞士巴塞尔

(72)发明人 B·哈利 Z·胡 J·C·乔利

A·Y·沈 B·R·斯奈德克尔

权利要求书2页 说明书36页

序列表5页 附图6页

(54)发明名称

含Fc蛋白的表达

(57)摘要

本发明提供包含含Fc蛋白的组合物,其中基本上所有的Fc结构域都具有C-末端赖氨酸。本发明还提供用于生产所述组合物的宿主细胞、制备所述宿主细胞和组合物的方法、以及其使用方法。

1. 一种用于表达含Fc蛋白的宿主细胞,其中与包含不具有CpD基因失活的野生型CpD基因的宿主细胞相比,所述宿主细胞具有至少60%的羧肽酶D (CpD) 表达水平降低。

2. 根据权利要求1的宿主细胞,其中所述宿主细胞是真核细胞。

3. 根据权利要求2的宿主细胞,其中所述宿主细胞是哺乳动物细胞。

4. 根据权利要求3的宿主细胞,其中所述宿主细胞是中国仓鼠卵巢 (CHO) 细胞。

5. 根据权利要求1的宿主细胞,其中所述含Fc蛋白是抗体。

6. 根据权利要求5的宿主细胞,其中所述抗体是单克隆抗体。

7. 根据权利要求1的宿主细胞,其中所述含Fc蛋白是含Fc的融合蛋白。

8. 根据权利要求1-7中任一项的宿主细胞,其中所述宿主细胞中的CpD基因被失活。

9. 根据权利要求8的宿主细胞,其中所述CpD基因通过siRNA失活。

10. 根据权利要求8的宿主细胞,其中所述CpD基因通过shRNA失活。

11. 根据权利要求8的宿主细胞,其中所述CpD基因通过基因缺失而失活。

12. 根据权利要求8的宿主细胞,其中所述CpD基因通过基因添加或取代而失活。

13. 根据权利要求11或12的宿主细胞,其中所述CpD基因使用成簇的规律间隔的短回文重复序列 (CRISPR) 系统失活。

14. 根据权利要求11或12的宿主细胞,其中使用转录激活因子样效应物核酸酶 (TALEN) 系统使CpD基因失活。

15. 根据权利要求11或12的宿主细胞,其中使用锌指核酸酶 (ZFN) 系统使CpD基因失活。

16. 根据权利要求11或12的宿主细胞,其中使用大范围核酸酶系统使CpD基因失活。

17. 根据权利要求1至16中任一项的宿主细胞,其包含表达载体,所述表达载体包含编码所述含Fc蛋白的核酸。

18. 一种细胞培养系统,其包含权利要求1-17中任一项所述的宿主细胞。

19. 生产权利要求1-17中任一项的宿主细胞的方法,其中所述方法包括使宿主细胞中的CpD基因失活,从而产生该宿主细胞。

20. 根据权利要求19的方法,其中使用siRNA系统使CpD基因失活。

21. 根据权利要求19的方法,其中使用shRNA系统使CpD基因失活。

22. 根据权利要求19的方法,其中通过基因缺失使CpD基因失活。

23. 根据权利要求19的方法,其中通过基因添加或取代使CpD基因失活。

24. 根据权利要求22或23的方法,其中使用CRISPR系统使CpD基因失活。

25. 根据权利要求22或23的方法,其中使用TALEN系统使CpD基因失活。

26. 根据权利要求22或23的方法,其中使用ZFN系统使CpD基因失活。

27. 根据权利要求22或23的方法,其中使用大范围核酸酶系统使CpD基因失活。

28. 一种制备含Fc蛋白的方法,其包括:

a) 培养权利要求1-17中任一项的宿主细胞;

b) 获得宿主细胞表达的含Fc蛋白。

29. 一种包含多个含Fc蛋白的组合物,其中所述组合物中的基本上所有含Fc蛋白在每个Fc结构域上具有C-末端赖氨酸。

30. 根据权利要求29的组合物,其中所述组合物的基本上所有含Fc蛋白具有相同的氨基酸序列。

31. 根据权利要求29或30的组合物,其中所述多个含Fc蛋白在电荷状态上是基本上均质的。
32. 根据权利要求29-31中任一项的组合物,其中所述含Fc蛋白是含Fc的融合蛋白。
33. 根据权利要求29-31中任一项的组合物,其中所述含Fc蛋白是抗体。
34. 根据权利要求33的组合物,其中所述抗体是人抗体。
35. 根据权利要求33的组合物,其中所述抗体是人源化抗体。
36. 根据权利要求33-35中任一项的组合物,其中所述抗体包含两条重链,并且其中每条重链包含C-端赖氨酸。
37. 根据权利要求32-36中任一项的组合物,其中所述含Fc蛋白缀合至药物。
38. 根据权利要求32-37中任一项的组合物,其中所述组合物中的含Fc蛋白包含IgG1Fc结构域。
39. 根据权利要求32-37中任一项的组合物,其中所述组合物中的含Fc蛋白包含IgG2Fc结构域。
40. 根据权利要求32-37中任一项的组合物,其中所述组合物中的含Fc蛋白包含IgG4Fc结构域。
41. 根据权利要求29-40中任一项的组合物,其中所述组合物是药物组合物。
42. 根据权利要求41的组合物,其中所述药物组合物是无菌药物组合物。
43. 根据权利要求29-40中任一项的组合物,其中所述组合物是细胞培养基。
44. 根据权利要求29-40中任一项的组合物,其中所述组合物是细胞裂解物。
45. 根据权利要求29-40中任一项的组合物,其中所述组合物是来自蛋白质纯化柱的洗脱物。
46. 治疗有需要的个体的疾病的方法,其包括:向个体施用权利要求41或42的组合物。
47. 根据权利要求46的方法,其中所述组合物肠胃外施用。
48. 根据权利要求47的方法,其中所述组合物通过静脉内或皮下施用。
49. 如权利要求46的方法,其中所述组合物局部施用。
50. 如权利要求49的方法,其中所述组合物外部施用。

含Fc蛋白的表达

[0001] 相关申请的交叉引用

[0002] 本申请要求于2015年9月22日提交的美国临时专利申请NO.62/222,187的优先权，其公开内容通过引用整体并入本文。

[0003] 在ASCII文本文件中提交序列表

[0004] 以下提交的ASCII文本文件的内容通过引用整体并入本文：序列表的计算机可读形式(CRF)(文件名:146392032640SEQLIST.TXT,记录日期:2016年8月18日,尺寸:5KB)。

技术领域

[0005] 本发明涉及包含含Fc蛋白的组合物，其中基本上所有含Fc的蛋白在每个Fc结构域上都具有C-末端赖氨酸。本发明进一步涉及用于产生所述组合物的宿主细胞，制备所述宿主细胞和组合物的方法以及其使用方法。

背景技术

[0006] 含有Fc结构域的治疗性蛋白质，例如单克隆抗体和含有Fc的融合蛋白，已经成为治疗疾病，如癌症、自身免疫疾病和感染的重要药物，并且目前是增长最快的一类治疗剂。参见例如Beck等，*Nat Rev Immunol*,10,2010。临床上，基于多肽的治疗剂由于其高度的靶特异性和参与免疫应答的能力而表现出有效的活性。参见例如Sidhu，*Nat Biotechnol*,25,2007。然而，基于多肽的均质治疗剂的商业生产是具有挑战性的，因为多肽会经历许多修饰，例如不同的二硫化物配对，脱酰胺化，氧化，N-末端谷氨酰胺环化，和不同的聚糖结构。参见例如Carson，*Nat Biotechnol*,23,2005。此外，多肽修饰的存在及结构一致性对生产条件具有高度敏感性。在生物活性多肽上包含免疫球蛋白Fc结构域，可以为该生物活性剂提供有益的生物学和药理学性质，包括增加的血浆半衰期以允许延长的治疗活性和较低的给药频率。参见Czajkowsky等，*Mol Med*,2012年4月。然而，Fc结构域的存在进一步增加了含Fc蛋白产物异质性的可能，例如跨每个Fc结构域上的C端赖氨酸存在的差异。

[0007] 用于生物药如含Fc蛋白生产的工业标准要求证明产品在所有临床和商业批次上的一致性。必须进行多肽产品的可比性研究，以确保生产参数发生变化后的一致性，例如，在I期至III期生产优化、商业规模化、和商品化后更改期间发生的生产参数变化。生产参数变化包括但不限于，细胞培养工艺参数、细胞培养基、宿主细胞、纯化、储存和配制的变化。此外，对于多地点生产，有必要确保所有生产设施的产品一致性。多肽属性(例如Fc结构域C-末端赖氨酸存在)水平的变异性和异质性，被认为是缺乏生产过程控制的指标。参见，Chirino等，*Nat Biotechnol*,22,2004;和Kozlowski等人，*J Pharm Sci*,97,2006。细胞系的差异及其对生产工艺(例如培养条件)的敏感性，影响最终产品的特性。例如，在单克隆抗体(由两条重链组成)的生产过程中，Fc结构域C末端赖氨酸的加工(即切割)可以导致带有零个、一个或两个Fc结构域C末端赖氨酸残基的抗体同种型的混合物。在抗体生产中可观察到广泛的Fc结构域C-末端赖氨酸加工，导致整个抗体群体的大百分比带有零或一个Fc结构域C-末端赖氨酸。参见Dick等，*Biotechnol Bioeng*,100,2008。因此，Fc结构域C-末端赖氨酸

存在的异质性是含Fc的治疗性蛋白的生产中的一个独特挑战。

[0008] 已经进行了广泛的细胞培养条件优化研究,以降低在含Fc蛋白生产期间产生的Fc结构域C-末端赖氨酸存在的异质性。例如,美国专利申请号20130280274已经描述了,控制细胞培养基的重金属如锌浓度、细胞培养基的氨基酸浓度,pH和温度等,对抗体Fc结构域C末端赖氨酸存在的影响。然而,如美国专利申请号20130280274所示,细胞培养参数的优化并没有消除Fc结构域C-末端赖氨酸存在的异质性。

[0009] IgG1,IgG2和IgG4同种型免疫球蛋白的Fc结构域在各Fc结构域的C末端含有保守的碱性氨基酸赖氨酸。这些免疫球蛋白同种型的Fc结构域C-末端赖氨酸存在的异质性,被认为是由细胞培养生产过程中内源性羧肽酶的蛋白水解引起的。参见Harris,J Chromatogr A,705,1995;和Luo等人,Biotechnol Bioeng,109,2012。羧肽酶是从蛋白质和肽的C-末端水解切割氨基酸的酶。有13种已知的羧肽酶成员存在于大多数哺乳动物物种中。参见Reznik等人,Cell Mol Life Sci,58,2001。虽然已经在CHO细胞(一种常用的生产宿主细胞)中检测到不同的羧肽酶蛋白质及其活性,但不知道是哪种羧肽酶负责Fc结构域C-末端赖氨酸的去除。参见Dick等,Biotechnol Bioeng,100,2008。先前推测,一种或多种羧肽酶,包括CpB和CpM,可能参与Fc结构域C末端赖氨酸加工。参见Dick等,Biotechnol Bioeng,100,2008;Harris,J Chromatogr A,705,1995;和Luo等,Biotechnol Bioeng,109,2012。

[0010] 本文提到的所有出版物、专利、专利申请和公布的专利申请的公开内容通过引用整体并入本文。

[0011] 发明概述

[0012] 本文提供用于表达含有Fc的蛋白质的宿主细胞,其中与包含没有CpD基因失活的野生型CpD基因的宿主细胞相比,该宿主细胞具有至少60%的羧肽酶D(CpD)表达水平降低。

[0013] 在根据(或应用于)上述任何实施方案的一些实施方案中,宿主细胞是真核细胞。在一些实施方案中,宿主细胞是哺乳动物细胞。在一些实施方案中,宿主细胞是中国仓鼠卵巢(CHO)细胞。

[0014] 在根据(或应用于)上述任何实施方案的一些实施方案中,含Fc的蛋白质是抗体。在一些实施方案中,抗体是单克隆抗体。

[0015] 在根据(或应用于)上述任何实施方案的一些实施方案中,含Fc的蛋白质是含Fc的融合蛋白。

[0016] 在根据(或应用于)上述任何实施方案的一些实施方案中,宿主细胞中的CpD基因失活。在一些实施方案中,CpD基因通过siRNA失活。在一些实施方案中,CpD基因通过shRNA失活。在一些实施方案中,通过基因缺失,使CpD基因失活。在一些实施方案中,通过基因添加或取代,使CpD基因失活。在一些实施方案中,使用成簇的规律间隔的短回文重复序列(CRISPR)系统,使CpD基因失活。在一些实施方案中,使用转录激活因子样效应物核酸酶(TALEN)系统,使CpD基因失活。在一些实施方案中,使用锌指核酸酶(ZFN)系统,使CpD基因失活。在一些实施方案中,使用大范围核酸酶(meganuclease)系统,使CpD基因失活。

[0017] 在根据(或应用于)上述任何实施方案的一些实施方案中,宿主细胞包含含有编码含Fc蛋白的核酸的表达载体。

[0018] 本文提供了包含上述任何实施方案的宿主细胞的细胞培养系统。

[0019] 本文提供了产生宿主细胞的方法,其中所述方法包括使宿主细胞中的CpD基因失活,由此产生上述任何实施方案的宿主细胞。

[0020] 在一些实施方案中,产生宿主细胞的方法包括使用siRNA系统使CpD基因失活。在一些实施方案中,产生宿主细胞的方法包括使用shRNA系统使CpD基因失活。在一些实施方案中,产生宿主细胞的方法包括通过基因缺失使CpD基因失活。在一些实施方案中,产生宿主细胞的方法包括通过基因添加或取代使CpD基因失活。在一些实施方案中,产生宿主细胞的方法包括使用CRISPR系统使CpD基因失活。在一些实施方案中,产生宿主细胞的方法包括使用TALEN系统使CpD基因失活。在一些实施方案中,产生宿主细胞的方法包括使用ZFN系统使CpD基因失活。在一些实施方案中,产生宿主细胞的方法包括使用大范围核酸酶系统使CpD基因失活。

[0021] 本文提供制备含Fc蛋白的方法,其包括:a) 培养宿主细胞,和b) 获得宿主细胞表达的含Fc蛋白。

[0022] 本文提供了包含多个含Fc蛋白的组合物,其中组合物中基本上所有的含Fc蛋白在每个Fc结构域上都具有C-末端赖氨酸。在一些实施方案中,组合物中基本上所有的含Fc蛋白都具有相同的氨基酸序列。在一些实施方案中,该多个含Fc蛋白在电荷状态上是基本上均质的。在一些实施方案中,含Fc蛋白是含Fc的融合蛋白。在一些实施方案中,含Fc蛋白是抗体。在一些实施方案中,抗体是人抗体。在一些实施方案中,抗体是人源化抗体。在一些实施方案中,抗体包含两条重链,并且其中每条重链包含C末端赖氨酸。在一些实施方案中,含Fc蛋白与药物缀合。在一些实施方案中,组合物中的含Fc蛋白包含IgG1 Fc结构域。在一些实施方案中,组合物中的含Fc蛋白包含IgG2 Fc结构域。在一些实施方案中,组合物中的含Fc蛋白包含IgG4 Fc结构域。

[0023] 在根据(或应用于)上述任何实施方案的一些实施方案中,包含多个含Fc蛋白的组合物是药物组合物。在一些实施方案中,包含多个含Fc蛋白的药物组合物是无菌药物组合物。

[0024] 在根据(或应用于)任何上述实施方案的一些实施方案中,包含多个含Fc蛋白的组合物是细胞培养基。

[0025] 在根据(或应用于)上述任何实施方案的一些实施方案中,包含多个含Fc蛋白的组合物是细胞裂解物。

[0026] 在根据(或应用于)任何上述实施方案的一些实施方案中,包含多个含Fc蛋白的组合物是来自蛋白纯化柱的洗脱物。

[0027] 本文还提供治疗有需要的个体的疾病的方法,其包括向个体施用上文实施方案中描述的药物组合物或无菌药物组合物。在一些实施方案中,治疗有需要的个体的疾病的方法包括肠胃外施用组合物。在一些实施方案中,治疗有需要的个体的疾病的方法包括静脉内或皮下施用组合物。在一些实施方案中,治疗有需要的个体的疾病的方法包括局部(locally)施用组合物。在一些实施方案中,治疗有需要的个体的疾病的方法包括外用(topically)施用组合物。

[0028] 根据随后的详细描述和所附权利要求,本发明的这些和其它方面和优点将是清楚的。应该理解,可以将本文描述的各个实施方案的一个,一些或全部特性组合,以形成本发明的其他实施方案。

[0029] 附图简述

[0030] 图1A-1B显示细胞系中内源性羧肽酶的qPCR mRNA分析。每个误差条表示1个标准差。图1A显示DP12和CHOK1宿主细胞中相对的羧肽酶B (CpB)、羧肽酶E (CpE)、羧肽酶M (CpM)、羧肽酶N1 (CpN1) 和羧肽酶D (CpD) mRNA表达水平。将每种羧肽酶的mRNA表达水平相对于管家基因 β -2-微球蛋白 (b2m) 进行归一化。白色条表示来自DP12宿主的测量值, 黑色条表示来自CHOK1宿主的测量值。图1B显示抗体表达细胞系中的相对羧肽酶mRNA表达水平。将每种羧肽酶的mRNA表达水平相对于管家基因b2m归一化。白色条表示来自抗体表达细胞系A的测量值, 黑色条表示来自抗体表达细胞系B的测量值。

[0031] 图2A-2B显示在抗体表达细胞系A (白色条) 和抗体表达细胞系B (黑色条) 中用羧肽酶特异性siRNA构建体转染后羧肽酶mRNA的表达水平。每个误差条表示1个标准差。图2A显示抗体表达细胞系A和抗体表达细胞系B中的相对CpD mRNA表达水平。灰色条表示以乱码构建体 (scramble construct) 转染细胞系A和细胞系B后在细胞系中的CpD mRNA水平; 白色条表示CpDi构建体转染后细胞系A中的CpD mRNA水平; 黑色条表示CpDi构建体转染后细胞系B中的CpD mRNA。将每个细胞系中CpD mRNA的表达水平与用乱码构建体转染后的相应细胞系中的CpD表达水平进行比较。图2B显示抗体表达细胞系A和抗体表达细胞系B中的相对CpN mRNA表达水平。灰色条表示在乱码构建体转染后细胞系中的mRNA水平; 白色条表示在CpNi构建体转染后细胞系A中的mRNA水平; 黑色条表示在CpNi构建体转染后细胞系B中的mRNA水平。将每个细胞系中CpN mRNA的表达水平与用乱码构建体转染后的相应细胞系中的CpN表达水平进行比较。

[0032] 图3显示在抗体表达细胞系A (黑色条) 和抗体表达细胞系B (灰色条) 中用羧肽酶特异性siRNA构建体转染后表达抗体的C-末端赖氨酸水平。

[0033] 图4A-4B显示野生型 (WT) 和敲除 (KO) CpD等位基因的示意图。图4A显示野生型CpD等位基因的示意图。两个引导RNA (gRNA) 序列 (gRNA 1和gRNA 2) 分别用于靶向CHO CpD基因的外显子2和外显子21, 以通过CRISPR敲除基因。两个gRNA之间的距离约为46kb。用于检测旨在通过CRISPR去除的WT CpD序列的正向和反向PCR引物分别表示为WT.F和WT.R。图4B显示KO CpD等位基因的示意图。用于检测KO CpD等位基因的正向和反向PCR引物分别表示为KO.F和KO.R。

[0034] 图5A-5B显示来自抗体表达细胞系B的两个CpD敲除克隆中的CpD表达水平。图5A显示两个CpD KO克隆和两个CpD WT克隆的CpD RNA表达水平。图5B显示两个CpD KO克隆和两个CpD WT克隆的Western印迹分析。

[0035] 图6A-6C显示抗体生产细胞系B的两个CpD KO克隆和两个CpD WT克隆的补料分批摇瓶评估数据。每个柱代表两次重复的平均值。每个误差条表示1个标准差。图6A显示在第14天表达抗体的滴度。图6B显示每细胞每天的皮克 (pg) 比生产力 (Qp)。图6C显示在培养14天后在1亿个细胞-天/升中的积分活细胞计数。

[0036] 图7显示抗体生产细胞系B的两个CpD KO克隆和两个CpD WT克隆中的C-末端赖氨酸水平。

[0037] 发明详述

[0038] 本申请基于令人惊讶的发现, 即羧肽酶D (CpD) (一种水解蛋白质的单个C-末端氨基酸的锌结合酶) 是负责Fc结构域C-末端赖氨酸加工 (即切割) 的唯一羧肽酶。具体而言, 使

用质谱分析,我们显示在CpD敲除细胞中Fc结构域C末端赖氨酸切割被完全消除,从而证明CpD是CHO细胞中切割抗体Fc结构域C末端赖氨酸的唯一内源性羧肽酶。

[0039] 因此,本申请提供了一种生产含Fc蛋白的方法,其中通过使用具有降低的CpD表达水平的宿主细胞,来产生在每个Fc结构域上具有C-末端赖氨酸的含Fc蛋白。通过这种方法产生的产物可以导致Fc结构域C-末端赖氨酸存在的100%同质性。这种方法使得无需考虑生产条件改变对Fc结构域C末端赖氨酸存在状态的影响,并且因此极大地促进了含Fc蛋白的制造过程。例如,使用这种方法产生含Fc蛋白,可以导致无需优化细胞培养条件来最小化多个生产批次之间(例如,比较多个临床和商业批次时)Fc结构域C-末端赖氨酸存在的变异性。此外,例如,使用这种方法产生含Fc蛋白,可以导致无需优化细胞培养条件来平衡Fc结构域C-末端赖氨酸存在的变异性最小化与含Fc蛋白的其他特征的变异性和状态(例如含Fc蛋白的滴度和翻译后修饰的存在)。此外,这种产生含Fc蛋白的方法可以大大促进含Fc蛋白的随后分析。例如,使用这种方法产生含Fc蛋白,可以导致无需引入额外步骤来测定具有异质性Fc结构域C-末端赖氨酸存在的多个含Fc蛋白的电荷变体(例如,涉及与CpB温育以除去所有Fc结构域C末端赖氨酸的步骤)。

[0040] 此外,Fc结构域C-末端赖氨酸的保留可抑制在含Fc蛋白生产期间Fc结构域C-末端氨基酸的随后改变。例如,IgG1、IgG2和IgG3免疫球蛋白的保守Fc结构域C端氨基酸序列是脯氨酸-甘氨酸-赖氨酸。Fc结构域C-末端赖氨酸的切割可产生末端甘氨酸,从而允许肽酰甘氨酸 α -酰胺化单加氧酶催化两步骤酰胺化反应,以除去该甘氨酸残基并将酰胺部分添加至IgG1、IgG2和IgG3中的该脯氨酸(该酶的活性也取决于培养基中的铜浓度)。Fc结构域C-末端脯氨酸的酰胺化及其程度,可以进一步在含Fc蛋白群体内引起电荷变体。

[0041] 因此,本申请一方面提供用于表达含Fc蛋白的宿主细胞,其中宿主细胞具有降低的羧肽酶D(CpD)表达水平。另一方面,本申请提供包含该宿主细胞的细胞培养系统。本申请在另一方面提供生产宿主细胞的方法,其中所述方法包括使宿主细胞中的CpD基因失活,从而产生宿主细胞。本申请再一方面提供制备含Fc蛋白的方法,包括培养宿主细胞并获得宿主细胞表达的含Fc蛋白。本申请的又一方面提供了包含多个含Fc蛋白的组合物,其中组合物中基本上所有含Fc蛋白在每个Fc结构域上具有C-末端赖氨酸。本申请再一方面提供了治疗有需要的个体的疾病的方法,其包括向个体施用组合物,其中所述组合物是包含多个在每个Fc结构域上具有C-末端赖氨酸的含Fc蛋白的药物组合物。

[0042] 定义

[0043] 如本文所用,“Fc结构域”是指免疫球蛋白重链的Fc区或其C端片段。该术语包括野生型Fc结构域和变体Fc结构域。在一些实施方案中,人IgG重链Fc结构域从Cys226或从Pro230延伸至重链的羧基末端(氨基酸编号根据EU编号系统,也称为EU索引,如Kabat等, *Sequences of Proteins of Immunological Interest*, 第5版, Public Health Service, National Institutes of Health, Bethesda, MD, 1991中的描述)。在一些实施方案中,该术语包括免疫球蛋白重链的C端片段和一个或多个恒定区。在一些实施方案中,对于IgG, Fc结构域可以包含免疫球蛋白结构域CH2和CH3以及CH1和CH2之间的铰链。

[0044] 如本文所用,“含Fc蛋白”是指包含Fc结构域的蛋白质(例如抗体或含Fc的融合蛋白质)。在一些实施方案中,含Fc蛋白包含一个或多个蛋白质亚基。在一些实施方案中,含Fc蛋白包含一或多个多肽。

[0045] 如本文所用,“含Fc的融合蛋白”是指包含与至少一个其他异源蛋白单位或多肽融合的Fc结构域的蛋白。

[0046] 本文使用的术语“多肽”是指单一多肽链。

[0047] 本文使用的术语“重链”是指免疫球蛋白重链。

[0048] 本文中的术语“抗体”以最广泛的含义使用并且涵盖各种抗体结构,包括但不限于单克隆抗体,多克隆抗体,多特异性抗体(例如双特异性抗体)和抗体片段,只要它们包含Fc结构域。

[0049] 术语“嵌合”抗体是指这样的抗体,其中重链和/或轻链的一部分来自特定来源或物种,而重链和/或轻链的其余部分来自不同的来源或物种。

[0050] “人抗体”是指具有这样的氨基酸序列的抗体,所述氨基酸序列对应于人或人细胞产生的抗体的氨基酸序列、或对应于来自非人来源的抗体的氨基酸序列,其中所述非人来源利用人抗体库或其他的人抗体编码序列。人抗体的该定义特别排除包含非人抗原结合残基的人源化抗体。

[0051] “人源化”抗体是指包含来自非人高变区(HVR)的氨基酸残基和来自人构架区(FR)的氨基酸残基的嵌合抗体。在某些实施方案中,人源化抗体将包含至少一个,并且通常两个可变结构域的基本上全部,其中全部或基本上全部HVR(例如CDR)对应于非人抗体的HVR,而全部或基本上全部的FR对应于人抗体的FR。人源化抗体任选地可以包含源自人抗体的抗体恒定区的至少一部分。抗体(例如非人抗体)的“人源化形式”是指已经历人源化的抗体。

[0052] 如本文所用的术语“单克隆抗体”是指从基本上同质的抗体群体获得的抗体,即除了通常以小量存在的可能变体抗体以外,群体中包含的各抗体是相同的和/或结合相同的表位,其中所述可能的变体抗体可以例如含有天然存在的突变或在单克隆抗体制备物生产期间出现。与通常包含针对不同决定簇(表位)的不同抗体的多克隆抗体制备物相反,单克隆抗体制备物中的所有单克隆抗体均针对抗原上的一个决定簇。因此,修饰语“单克隆”表示从基本上同质的抗体群获得的抗体的特征,并且不应被解释为需要通过任何特定方法产生该抗体。

[0053] 如本文所用,术语“免疫粘附素”是指,将异源蛋白质(“粘附素”)的结合特异性与免疫球蛋白恒定结构域的效应子功能组合在一起的分子。在结构上,免疫粘附素包含具有所需结合特异性的氨基酸序列和免疫球蛋白恒定结构域序列(例如IgG的CH2和/或CH3序列)的融合,其中所述具有所需结合特异性的氨基酸序列不同于抗体的抗原识别和结合位点(即,与抗体的恒定区相比是“异源的”)。示例性的粘附素序列包括连续的氨基酸序列,其包含与感兴趣的蛋白质结合的受体或配体的一部分。粘附素序列也可以是非受体或配体序列、但与感兴趣蛋白质结合的序列(例如,肽体(peptibody)中的粘附素序列)。这样的多肽序列可以通过各种方法选择或鉴定,包括噬菌体展示技术和高通量分选方法。免疫粘附素中的免疫球蛋白恒定结构域序列可以从任何免疫球蛋白例如IgG1,IgG2,IgG3或IgG4亚型,IgA(包括IgA1和IgA2),IgE,IgD或IgM获得。

[0054] 如本文所用,“宿主细胞”是指能够产生蛋白质或多肽产物的细胞。例如,宿主细胞可以产生含Fc蛋白。

[0055] 术语“个体”是指哺乳动物,并且包括但不限于人,牛,马,猫科动物,犬科动物,啮齿类动物或灵长类动物。

[0056] 如本文所用,“治疗”是用于获得包括临床结果在内的有益或期望结果的方法。为了本发明的目的,有益的或期望的临床结果包括但不限于以下一种或多种:缓解由疾病引起的一种或多种症状,减轻疾病的程度,稳定疾病(例如预防或延迟疾病恶化),预防或延迟疾病的传播(例如转移),预防或延迟疾病的复发,延迟或减缓疾病进展,改善疾病状态,提供疾病(部分或全部)缓解,减少治疗疾病所需的一种或多种其他药物的剂量,延缓疾病的进展,提高生活质量和/或延长生存期。本发明的方法涵盖治疗的这些方面中的任何一个或多个。

[0057] “治疗有效量”是至少实现特定病症的可测量改善所需的最小浓度。本文中的治疗有效量可以根据诸如患者的疾病状态,年龄,性别和体重以及抗体在个体中引起期望反应的能力等因素而变化。治疗有效量也是其中治疗有益效果超过抗体的任何毒性或有害作用的量。“预防有效量”是指在必需的剂量和时间段内有效达到所需预防结果的量。典型地但非必需地,由于预防剂量在疾病之前或疾病的早期阶段用于受试者,所以预防有效量可以小于治疗有效量。

[0058] 应该理解,这里描述的本发明的方面和实施方案包括“由.....组成”和/或“基本上由.....组成”的方面和实施方案。

[0059] 在此提及的“约”值或参数包括(并描述)涉及该值或参数本身的变化。例如,提及“大约X”的描述中包括“X”的描述。

[0060] 如本文和所附权利要求书中所使用的,除非上下文另有明确规定,否则单数形式“一”,“或”和“该”包括复数指代物。

[0061] 宿主细胞

[0062] 本申请提供用于表达含Fc蛋白的宿主细胞,其中宿主细胞具有降低的羧肽酶D(CpD)表达水平。

[0063] 可以使用的宿主细胞包括真核细胞,例如酵母或高等真核细胞。高等真核细胞包括昆虫细胞和建立的哺乳动物来源的细胞系。

[0064] 合适的哺乳动物宿主细胞的实例包括COS-7系的猴肾细胞(ATCC CRL 1651)(Gluzman等,Cell,23,1981),L细胞,293细胞,C127细胞,3T3细胞(ATCC CCL 163),中国仓鼠卵巢(CHO)细胞或其衍生物,如Veggie CHO和在无血清培养基中生长的相关细胞系(Rasmussen等,Cytotechnology,28,1998),HeLa细胞,BHK(ATCC CRL10)细胞系,以及McMahan等人,EMBO J,10,1991所述的源自非洲绿猴肾细胞系CVI(ATCC CCL 70)的CVI/EBNA细胞系,人胚胎肾细胞如293,293EBNA或MSR 293,人表皮A431细胞,人Colo205细胞,其他转化的灵长类细胞系,正常二倍体细胞,源自原代组织的体外培养的细胞株,原代外植体,HL-60,U937,HaK或Jurkat细胞。任选地,例如,哺乳动物细胞系如HepG2/3B,KB,NIH 3T3或S49可用作宿主细胞。

[0065] 在一些实施方案中,宿主细胞是CHO细胞。CHO细胞在本领域中是熟知的。参见例如Xu等,Nat Biotechnol,29,2011。在一些实施方案中,宿主细胞是DP12宿主细胞。在一些实施方案中,宿主细胞是DUXB-11衍生的DHFR缺陷型DP12宿主细胞。在一些实施方案中,宿主细胞是CHOK1宿主细胞。在一些实施方案中,宿主细胞是DHFR阳性CHOK1宿主细胞。在一些实施方案中,宿主细胞是CHOK1M宿主细胞。

[0066] 在一些实施方案中,宿主细胞是小鼠宿主细胞。在一些实施方案中,宿主细胞是

Sp2/0宿主细胞。在一些实施方案中,宿主细胞是NS0宿主细胞。

[0067] 在一些实施方案中,宿主细胞是杂交瘤。在一些实施方案中,杂交瘤是抗体产生细胞,其中在用抗原免疫宿主之后从宿主收集该抗体产生细胞。在一些实施方案中,抗体产生细胞与骨髓瘤细胞融合。在一些实施方案中,宿主细胞是小鼠骨髓瘤来源的细胞系。

[0068] 备选地,宿主细胞可以是低等真核生物如酵母。合适的酵母包括酿酒酵母(*Saccharomyces cerevisiae*),粟酒裂殖酵母(*Schizosaccharomyces pombe*),克鲁维酵母(*Kluyveromyces*)菌株,假丝酵母(*Candida*)或任何能够表达异源多肽的酵母菌株。

[0069] 如本文所用,“CpD表达水平降低”是指,与包含野生型CpD基因且CpD基因未失活的宿主细胞相比,CpD表达水平降低至少60%。在一些实施方案中,与失活前的宿主细胞相比,宿主细胞具有至少60%的CpD表达水平降低。在一些实施方案中,CpD表达水平降低至少约60%,65%,70%,75%,80%,85%,90%,95%或100%。在一些实施方案中,CpD表达水平降低至少95%。在一些实施方案中,CpD表达水平降低100%。在一些实施方案中,CpD表达水平降低约60%至100%,约70%至100%,约80%至100%,约90%至100%,或约95%至100%。

[0070] 在一些实施方案中,宿主细胞中的CpD基因失活,其中CpD表达水平的降低基于DNA水平。在一些实施方案中,宿主细胞中的CpD基因失活,其中CpD表达水平的降低基于RNA水平。在一些实施方案中,宿主细胞中的CpD基因失活,其中CpD表达水平的降低基于多肽水平。

[0071] 在一些实施方案中,宿主细胞中的CpD基因失活,其中降低的表达水平基于基因失活后的CpD表达水平。在一些实施方案中,宿主细胞中的CpD基因失活,其中降低的表达水平基于DNA水平。在一些实施方案中,宿主细胞中的CpD基因失活,其中降低的表达水平基于RNA水平。在一些实施方案中,宿主细胞中的CpD基因失活,其中降低的表达水平基于多肽水平。

[0072] 在一些实施方案中,宿主细胞中的CpD基因失活,其中CpD表达水平降低至少60%。在一些实施方案中,宿主细胞中的CpD基因失活,其中CpD表达水平降低至少90%。在一些实施方案中,宿主细胞中的CpD基因失活,其中CpD表达水平降低至少95%。在一些实施方案中,宿主细胞中的CpD基因失活,其中CpD表达水平降低100%。

[0073] 在一些实施方案中,宿主细胞中的CpD基因失活。如本文所用,“失活”是指抑制基因的翻译或潜在的未来翻译(即蛋白质的表达)。失活可发生在基因表达(包括但不限于转录,翻译和蛋白质表达)的任何阶段或过程中,并且失活可影响任何基因或基因产物,包括但不限于DNA, RNA, 如mRNA和多肽。

[0074] 用于使宿主细胞中的CpD基因失活的方法和技术包括但不限于:小干扰RNA(siRNA),小发夹RNA(shRNA;也称为短发夹RNA),成簇的规律间隔的短回文重复序列(CRISPR),转录激活因子样效应物核酸酶(TALEN),锌指核酸酶(ZFN),同源重组,非同源末端连接,和大范围核酸酶。参见例如O'Keefe, *Mater Methods*, 3, 2013; Doench等, *Nat Biotechnol*, 32, 2014; Gaj等, *Trends Biotechnol*, 31, 2014; 和Silva等, *Curr Gene Ther*, 11, 2011。

[0075] 在一些实施方案中,通过小干扰RNA(siRNA)系统,使CpD基因失活。鉴定适合CpD基因失活的siRNA序列的方法在本领域中是公知的。例如,开发和鉴定靶向CpD基因的siRNA的一般考虑因素包括:a)长度优选为21-23个核苷酸的第一搜索序列(随后根据需要减少序列

长度), b) 避免起始密码子和终止密码子的50-100碱基对之内的区域, c) 避免内含子区域, d) 避免4个或更多碱基的延伸, 例如AAAA, e) 避免GC含量小于30%或大于60%的区域, f) 避免重复和低序列复杂性, g) 避免单核苷酸多态性位点, 和h) 避免与其他脱靶基因中的序列互补的序列。参见例如Rules of siRNA design for RNA interference, Protocol Online, 2004年5月29日; 和Reynolds等, Nat Biotechnol, 22, 2004。

[0076] 在一些实施方案中, siRNA系统包含长度约10至200个核苷酸, 或约10至100个核苷酸, 或约15至100个核苷酸, 或约10至60个核苷酸长度, 或约15至60个核苷酸, 或约10至50个核苷酸, 或约15至50个核苷酸; 或约10至30个核苷酸, 或约15至30个核苷酸的siRNA核苷酸序列。在一些实施方案中, siRNA核苷酸序列的长度大约为10-25个核苷酸。在一些实施方案中, siRNA核苷酸序列的长度为约15-25个核苷酸。在一些实施方案中, siRNA核苷酸序列的长度为至少约10, 至少约15, 至少约20或至少约25个核苷酸。在一些实施方案中, siRNA系统包含与CpD mRNA分子的区域至少约80%, 至少约85%, 至少约90%, 至少约95%或100%互补的核苷酸序列。在一些实施方案中, siRNA系统包含与CpD pro-mRNA分子的区域至少约80%, 至少约85%, 至少约90%, 至少约95%或100%互补的核苷酸序列。在一些实施方案中, siRNA系统包含双链RNA分子。在一些实施方案中, siRNA系统包含单链RNA分子。在一些实施方案中, 宿主细胞包含如本文任何实施方案中所述的siRNA系统。在一些实施方案中, 宿主细胞包含如本文任何实施方案中所述加工成活性siRNA分子的前-siRNA核苷酸序列。在一些实施方案中, 宿主细胞包含与CpD mRNA分子的区域至少约80%, 至少约85%, 至少约90%, 至少约95%或100%互补的siRNA核苷酸序列。在一些实施方案中, 宿主细胞包含编码如本文任何实施方案中所述的siRNA分子的表达载体。在一些实施方案中, 宿主细胞包含编码如本文任何实施方案中所述的前-siRNA分子的表达载体。

[0077] 在一些实施方案中, siRNA系统包含递送载体。在一些实施方案中, 宿主细胞包含递送载体。在一些实施方案中, 递送载体包含前-siRNA和/或siRNA分子。

[0078] 示例性的CpD siRNA靶序列列于表1中。

[0079] 表1. 示例性的CpD siRNA靶序列siRNA.

[0080]

SEQ ID NO.:	CpD siRNA核苷酸序列:
19	GGA AGA GAA CTG CTA CTA A

[0081] 在一些实施方案中, 通过小发夹RNA (shRNA; 也称为短发夹RNA) 系统, 使CpD基因失活。通过shRNA系统的基因失活在本领域中是众所周知的。在一些实施方案中, shRNA系统包含长度约10至200个核苷酸, 或约10至100个核苷酸, 或约15至100个核苷酸, 或约10至60个核苷酸, 或约15至60个核苷酸, 或约10至50个核苷酸, 或约15至50个核苷酸, 或约10至30个核苷酸, 或约15至30个核苷酸的核苷酸序列。在一些实施方案中, shRNA核苷酸序列长约10-25个核苷酸。在一些实施方案中, shRNA核苷酸序列长度为约15-25个核苷酸。在一些实施方案中, shRNA核苷酸序列的长度为至少约10, 至少约15, 至少约20或至少约25个核苷酸。在一些实施方案中, shRNA系统包含与CpD mRNA分子的区域至少约80%, 至少约85%, 至少约90%, 至少约95%或100%互补的核苷酸序列。在一些实施方案中, shRNA系统包含与CpD pro-mRNA分子的区域至少约80%, 至少约85%, 至少约90%, 至少约95%或100%互补的核苷酸序列。在一些实施方案中, shRNA系统包含双链RNA分子。在一些实施方案中, shRNA系统

包含单链RNA分子。在一些实施方案中,宿主细胞包含如本文任何实施方案中所述的shRNA系统。在一些实施方案中,宿主细胞包含在本文任何实施方案中描述的加工为活性shRNA核苷酸序列的前shRNA核苷酸序列。在一些实施方案中,由DNA组成的前shRNA分子。在一些实施方案中,前shRNA分子是DNA构建体。在一些实施方案中,宿主细胞包含与CpD mRNA分子的区域至少约80%,至少约85%,至少约90%,至少约95%或100%互补的shRNA核苷酸序列。在一些实施方案中,宿主细胞包含编码如本文任何实施方案中所述的shRNA分子的表达载体。在一些实施方案中,宿主细胞包含编码如本文任何实施方案中所述的前shRNA分子的表达载体。

[0082] 在一些实施方案中,shRNA系统包含递送载体。在一些实施方案中,宿主包含递送载体。在一些实施方案中,递送载体包含前shRNA和/或shRNA分子。

[0083] 在一些实施方案中,通过基因缺失使CpD基因失活。如本文所用,“基因缺失”是指从基因或在基因附近除去至少一部分DNA序列。在一些实施方案中,经历基因缺失的序列包含基因的外显子序列。在一些实施方案中,经历基因缺失的序列包含基因的启动子序列。在一些实施方案中,经历基因缺失的序列包含基因的侧翼序列。在一些实施方案中,从基因去除基因序列的一部分。在一些实施方案中,从CpD基因或在CpD基因附近移除CpD基因序列的一部分。在一些实施方案中,从染色体中去除该整个基因序列。在一些实施方案中,完整的CpD序列从染色体中去除。在一些实施方案中,宿主细胞包含如本文任何实施方案中所述的基因缺失。在一些实施方案中,宿主细胞在CpD基因中包含基因缺失。在一些实施方案中,宿主细胞在CpD基因附近包含基因缺失。

[0084] 在一些实施方案中,通过基因添加或取代,使CpD基因失活。如本文所用,“基因添加”或“基因取代”是指基因序列的改变,包括插入或取代一个或多个核苷酸或核苷酸碱基对。在一些实施方案中,基因的内含子序列被改变。在一些实施方案中,基因的外显子序列被改变。在一些实施方案中,基因的启动子序列被改变。在一些实施方案中,基因的侧翼序列被改变。在一些实施方案中,将一个核苷酸或核苷酸碱基对添加至基因序列。在一些实施方案中,将至少一个连续的核苷酸或核苷酸碱基对添加至基因序列。在一些实施方案中,宿主细胞包含如本文任何实施方案中所述的基因添加或取代。在一些实施方案中,宿主细胞在CpD基因中包含基因添加或取代。

[0085] 在一些实施方案中,通过基因缺失,使CpD基因失活,其中基因序列中至少一个核苷酸或核苷酸碱基对的缺失导致非功能性基因产物。在一些实施方案中,通过基因缺失使CpD基因失活,其中基因序列中至少一个核苷酸的缺失导致不再具有原始基因产物功能或活性的基因产物。在一些实施方案中,通过基因缺失使CpD基因失活,其中基因序列中至少一个核苷酸缺失导致功能失调的基因产物。

[0086] 在一些实施方案中,通过基因添加或取代,使CpD基因失活,其中将至少一个核苷酸或核苷酸碱基对添加或取代到CpD基因序列中导致非功能性基因产物。在一些实施方案中,通过基因失活使CpD基因失活,其中将至少一个核苷酸并入或置换至CpD基因序列中导致不再具有原始基因产物功能或活性的基因产物。在一些实施方案中,通过基因添加或取代使CpD基因失活,其中将至少一个核苷酸并入或置换至CpD基因序列中导致功能失调的基因产物。

[0087] 在一些实施方案中,宿主细胞包含非功能性CpD基因产物。在一些实施方案中,宿

主细胞包含不具有原始CpD基因产物功能或活性的CpD基因产物。在一些实施方案中,宿主细胞包含功能失调的CpD基因产物。

[0088] 在一些实施方案中,宿主细胞包含失活的CpD基因,其中失活的CpD基因不表达全长和功能性CpD基因产物(例如全长CpD多肽序列)。在一些实施方案中,宿主细胞包含失活的CpD基因,其中失活的CpD基因不表达内源CpD基因产物序列。在一些实施方案中,宿主细胞包含失活的CpD基因,其中失活的CpD基因表达变体CpD基因产物。在一些实施方案中,宿主细胞包含变体CpD基因产物。

[0089] 在一些实施方案中,宿主细胞包含递送载体。在一些实施方案中,递送载体是病毒载体。在一些实施方案中,递送载体是慢病毒。在一些实施方案中,递送载体是腺病毒。在一些实施方案中,载体包含启动子。

[0090] 在一些实施方案中,宿主细胞是稳定的敲低(knockdown)宿主细胞。在一些实施方案中,宿主细胞是稳定的CpD敲低细胞系。在一些实施方案中,宿主细胞是瞬时敲低细胞系。在一些实施方案中,宿主细胞是瞬时CpD敲低细胞系。

[0091] 在一些实施方案中,宿主细胞还包含除CpD以外的失活基因。

[0092] 通常,用包含编码期望的含Fc蛋白的DNA的重组表达载体转化宿主细胞。另外,本申请的宿主细胞可以是空白宿主细胞。如本文所用,“空白宿主”是指不含编码含Fc蛋白的表达载体的细胞。在一些实施方案中,空白宿主细胞是CHO细胞。在一些实施方案中,空白宿主细胞是小鼠细胞。

[0093] 本申请还提供包含编码本文所述含Fc蛋白的核酸的宿主细胞。本发明提供的核酸分子包括单链和双链形式的DNA和RNA以及相应的互补序列。DNA包括例如cDNA,基因组DNA,化学合成的DNA,通过PCR扩增的DNA,及其组合。本发明的核酸分子包括全长基因或cDNA分子以及其片段的组合。本文提供的核酸优选来自人类来源。

[0094] 如本文其他地方所述,对于从天然来源分离的核酸,“分离的”核酸是指,该核酸与存在于该核酸所分离自的生物的基因组中的相邻基因序列分离。对于由模板酶促合成或通过化学方法合成的核酸(例如PCR产物,cDNA分子或寡核苷酸),可以理解,由这些方法产生的核酸是分离的核酸。分离的核酸分子可以指单独片段形式或作为较大核酸构建体的组分的核酸分子。

[0095] 在某些实施方案中,核酸基本上不含污染性内源物质。核酸分子优选源自至少一次以基本上纯的形式和一定的量或浓度分离的DNA或RNA,其中所述量或浓度使得能够通过标准生物化学方法(例如Sambrook等人Molecular Cloning:A Laboratory Manual,第2版,Cold Spring Harbor Laboratory,Cold Spring Harbor,NY(1989)描述的方法)鉴定、操作和回收其组分核苷酸序列。这样的序列优选以未被通常存在于真核基因中的内部非翻译序列或intran间断的开放阅读框形式提供和/或构建。非翻译DNA的序列可以存在于开放阅读框的5'或3',位于其不干扰编码区的操作或表达的位置。

[0096] 在一些实施方案中,宿主细胞能够表达含Fc蛋白。在一些实施方案中,宿主细胞包含含Fc蛋白。在一些实施方案中,宿主细胞能够分泌含Fc蛋白。

[0097] 在一些实施方案中,宿主细胞能够以与CpD基因失活之前宿主细胞类似的输出速率表达含Fc蛋白。在一些实施方案中,宿主细胞能够以与CpD基因失活之前的宿主细胞相同的输出速率表达含Fc蛋白。在一些实施方案中,宿主细胞能够以CpD基因失活之前宿主细胞

的输出速率的约70%、75%、80%、85%、90%、95%或100%，表达含Fc蛋白。

[0098] 在一些实施方案中，宿主细胞还包含基因修饰。为了产生含Fc蛋白质的目的，通过基因修饰对宿主细胞的优化是本领域众所周知的，并且包括涉及例如以下的考虑因素：糖基化模式、用于整合方法的载体选择性质、以及对于产生含Fc蛋白而言期望操作的任何其他细胞性质。在一些实施方案中，基因修饰是靶向的基因修饰。在一些实施方案中，基因修饰是敲除基因修饰。在一些实施方案中，基因修饰是敲入基因修饰。

[0099] 含Fc蛋白

[0100] 本申请提供含Fc蛋白，其中该含Fc蛋白质的每个Fc结构域具有C-末端赖氨酸。

[0101] 在一些实施方案中，含Fc蛋白包含Fc结构域。在一些实施方案中，含Fc蛋白包含一个或多个Fc结构域。在一些实施方案中，含Fc蛋白包含两个Fc结构域。

[0102] 在一些实施方案中，含Fc蛋白包含重链。在一些实施方案中，含Fc蛋白包含至少一个重链。在一些实施方案中，含Fc蛋白包含一个或多个重链。在一些实施方案中，含Fc蛋白包含两条重链。

[0103] 在一些实施方案中，含Fc蛋白包含Fc结构域，其中Fc结构域包含C末端赖氨酸。在一些实施方案中，含Fc蛋白包含至少一个Fc结构域，其中每个Fc结构域包含C-末端赖氨酸。在一些实施方案中，含Fc蛋白包含一个或多个Fc结构域，其中每个Fc结构域包含C-末端赖氨酸。在一些实施方案中，含Fc蛋白包含两个Fc结构域，其中每个Fc结构域包含C-末端赖氨酸。

[0104] 在一些实施方案中，含Fc蛋白包含重链，其中重链包含C末端赖氨酸。在一些实施方案中，含Fc蛋白包含至少一条重链，其中每条重链包含C末端赖氨酸。在一些实施方案中，含Fc蛋白包含一条或多条重链，其中每条重链包含C末端赖氨酸。在一些实施方案中，含Fc蛋白包含两条重链，其中每条重链包含C末端赖氨酸。

[0105] 在一些实施方案中，含Fc蛋白是多聚体蛋白质。

[0106] 在一些实施方案中，含Fc蛋白是抗体。在一些实施方案中，抗体是人抗体。在一些实施方案中，抗体是人源化抗体。在一些实施方案中，抗体是单克隆抗体。在一些实施方案中，抗体是嵌合抗体。在一些实施方案中，抗体是双特异性抗体。在一些实施方案中，抗体是多特异性抗体。

[0107] 在一些实施方案中，含Fc蛋白是含Fc的融合蛋白。在一些实施方案中，含Fc的融合蛋白包含一个或多个Fc结构域。在一些实施方案中，含Fc的融合蛋白包含一个或多个Fc结构域，每个Fc结构域包含C末端赖氨酸。

[0108] 在一些实施方案中，含Fc的融合蛋白的Fc结构域延长了含Fc的融合蛋白的血浆半衰期。在一些实施方案中，含Fc融合蛋白的Fc结构域延长了含Fc融合蛋白的生物学活性。在一些实施方案中，含Fc融合蛋白的Fc结构域降低含Fc融合蛋白的肾清除率。在一些实施方案中，含Fc融合蛋白的Fc结构域增加含Fc融合蛋白的溶解度。在一些实施方案中，含Fc融合蛋白的Fc结构域增加含Fc融合蛋白的稳定性。

[0109] 含Fc的融合蛋白在本领域中是公知的。参见例如Czajkowsky等, EMBO Mol Med, 4, 2012。在一些实施方案中，含Fc的融合蛋白是免疫粘附素。在一些实施方案中，含Fc的融合蛋白是细胞因子-Fc融合蛋白。

[0110] 在一些实施方案中，含Fc蛋白与药剂缀合。在一些实施方案中，含Fc蛋白与至少约

2、3、4、5、6、7、8、9、或10个分子的药剂缀合。在一些实施方案中，含Fc蛋白与约2-10，约4-10，约6-10或约8-10个分子的药剂缀合。在一些实施方案中，含Fc蛋白与药剂缀合，其中药剂缀合至含Fc蛋白的Fc结构域。在一些实施方案中，药剂是治疗剂。在一些实施方案中，治疗剂是小分子治疗剂。在一些实施方案中，治疗剂是化学治疗剂。在一些实施方案中，药剂是检测剂。在一些实施方案中，检测剂是放射性标记。在一些实施方案中，检测剂是荧光标记。在一些实施方案中，检测剂是免疫标记。在一些实施方案中，含Fc蛋白是伴随诊断(companion diagnostic, CD)。

[0111] 在一些实施方案中，含Fc蛋白包含IgG1 Fc结构域。在一些实施方案中，含Fc蛋白包含含C末端赖氨酸的IgG1 Fc结构域。在一些实施方案中，含Fc蛋白包含一个或多个IgG1 Fc结构域。在一些实施方案中，含Fc蛋白包含一个或多个IgG1 Fc结构域，每个Fc结构域包含C末端赖氨酸。在一些实施方案中，含Fc蛋白包含IgG1重链。在一些实施方案中，含Fc蛋白包含含C末端赖氨酸的IgG1重链。在一些实施方案中，含Fc蛋白包含一个或多个IgG1重链。在一些实施方案中，含Fc蛋白包含一个或多个IgG1重链，每条重链包含C末端赖氨酸。

[0112] 在一些实施方案中，含Fc蛋白包含IgG2 Fc结构域。在一些实施方案中，含Fc蛋白包含含C末端赖氨酸的IgG2 Fc结构域。在一些实施方案中，含Fc蛋白包含一个或多个IgG2 Fc结构域。在一些实施方案中，含Fc蛋白包含一个或多个IgG2 Fc结构域，每个Fc结构域包含C-末端赖氨酸。在一些实施方案中，含Fc蛋白包含IgG2重链。在一些实施方案中，含Fc蛋白包含含C末端赖氨酸的IgG2重链。在一些实施方案中，含Fc蛋白包含一个或多个IgG2重链。在一些实施方案中，含Fc蛋白包含一个或多个IgG2重链，每条重链包含C末端赖氨酸。

[0113] 在一些实施方案中，含Fc蛋白包含IgG4 Fc结构域。在一些实施方案中，含Fc蛋白包含含C末端赖氨酸的IgG4 Fc结构域。在一些实施方案中，含Fc蛋白包含一个或多个IgG4 Fc结构域。在一些实施方案中，含Fc蛋白包含一个或多个IgG4 Fc结构域，每个Fc结构域包含C-末端赖氨酸。在一些实施方案中，含Fc蛋白包含IgG4重链。在一些实施方案中，含Fc蛋白包含含C末端赖氨酸的IgG4重链。在一些实施方案中，含Fc蛋白包含一个或多个IgG4重链。在一些实施方案中，含Fc蛋白包含一个或多个IgG4重链，每条重链包含C-末端赖氨酸。

[0114] 在一些实施方案中，含Fc蛋白包含翻译后修饰。在一些实施方案中，翻译后修饰是非酶促产生的。在一些实施方案中，翻译后修饰是酶促产生的。在一些实施方案中，翻译后修饰选自二硫化物配对，脱酰胺化，氧化，N-末端谷氨酰胺环化和糖基化。

[0115] 包含多个含Fc蛋白的组合物

[0116] 另一方面，本申请提供包含多个含Fc蛋白的组合物，其中所述多个含Fc蛋白中基本上所有的含Fc蛋白在每个Fc结构域上具有C-末端赖氨酸。

[0117] 如本文所用，术语“基本上所有”是指至少约90%，包括例如95%或100%。因此，例如，组合物可以包含多个含Fc蛋白，其中所述多个含Fc蛋白中至少约90%，91%，92%，93%，94%，95%，96%，97%，98%，99%或100%的含Fc蛋白在每个Fc结构域上具有C-末端赖氨酸。在一些实施方案中，组合物包含多个含Fc蛋白，其中所述多个含Fc蛋白中100%的含Fc蛋白在每个Fc结构域上具有C-末端赖氨酸。

[0118] 在一些实施方案中，组合物包含多个含Fc蛋白，其中所述多个含Fc蛋白中约90%至100%的含Fc蛋白在每个Fc结构域上具有C-末端赖氨酸。

[0119] 在一些实施方案中，组合物包含多个含Fc蛋白，其中所述多个含Fc蛋白中100%的

含Fc蛋白在每个Fc结构域上具有C-末端赖氨酸。

[0120] 在一些实施方案中,组合物包含多个含Fc蛋白,其中所述多个含Fc蛋白中基本上所有的含Fc蛋白具有相同的氨基酸序列。在一些实施方案中,组合物包含多个含Fc蛋白,其中所述多个含Fc蛋白中基本上所有的含Fc蛋白具有相同的Fc结构域氨基酸序列。在一些实施方案中,组合物包含多个含Fc蛋白,其中所述多个含Fc蛋白中基本上所有的含Fc蛋白具有相同的重链氨基酸序列。

[0121] 在一些实施方案中,组合物包含多个含Fc蛋白,其中所述多个含Fc蛋白中至少约80%,85%,90%,95%或100%的含Fc蛋白具有相同的氨基酸序列。在一些实施方案中,组合物包含多个含Fc蛋白,其中所述多个含Fc蛋白中至少约80%,85%,90%,95%或100%的含Fc蛋白具有相同的Fc结构域氨基酸序列。在一些实施方案中,组合物包含多个含Fc蛋白,其中所述多个含Fc蛋白中至少约80%,85%,90%,95%或100%的含Fc蛋白具有相同的重链氨基酸序列。

[0122] 在一些实施方案中,组合物包含多个含Fc蛋白,其中所述多个含Fc蛋白中至少约90%至100%的含Fc蛋白具有相同的氨基酸序列。在一些实施方案中,组合物包含多个含Fc蛋白,其中所述多个含Fc蛋白中至少约90%至100%的含Fc蛋白具有相同的Fc结构域氨基酸序列。在一些实施方案中,组合物包含多个含Fc蛋白,其中所述多个含Fc蛋白中至少约90%至100%的含Fc蛋白具有相同的重链氨基酸序列。

[0123] 在一些实施方案中,组合物包含多个含Fc蛋白,其中所述多个含Fc蛋白中100%的含Fc蛋白具有相同的氨基酸序列。在一些实施方案中,组合物包含多个含Fc蛋白,其中所述多个含Fc蛋白中100%的含Fc蛋白具有相同的Fc结构域氨基酸序列。在一些实施方案中,组合物包含多个含Fc蛋白,其中所述多个含Fc蛋白中100%的含Fc蛋白具有相同的重链氨基酸序列。

[0124] 在一些实施方案中,组合物包含多个含Fc蛋白,其中所述多个含Fc蛋白在电荷状态方面是基本上均质的。如本文所用,“基本上均质的电荷状态”是指多个含Fc蛋白中的至少约80%,85%,90%,95%或100%具有相同的电荷。本领域充分明了,含Fc蛋白的电荷(例如表面电荷和净电荷)高度依赖于所述含Fc蛋白的分子组成。在一些实施方案中,组合物包含具有基本上均质的电荷状态的多个含Fc蛋白,其中所述多个含Fc蛋白中的至少约80%,85%,90%,95%或100%具有相同的表面电荷。在一些实施方案中,组合物包含具有基本上均质的电荷状态的多个含Fc蛋白,其中所述多个含Fc蛋白中的至少约80%,85%,90%,95%或100%具有相同的净电荷。在一些实施方案中,组合物包含具有基本上均质的电荷状态的多个含Fc蛋白,其中对所述多个含Fc蛋白中的每一个进行比较,所述多个含Fc蛋白中的每个含Fc蛋白在氨基酸序列的一个位置的氨基酸残基上具有相同的电荷(例如,-1,0,+1)。在一些实施方案中,组合物包含具有基本上均质的电荷状态的多个含Fc蛋白,其中所述多个含Fc蛋白中的每个含Fc蛋白具有相同的表面电荷。在一些实施方案中,组合物包含具有基本上均质的电荷状态的多个含Fc蛋白,其中所述多个含Fc蛋白中的每个含Fc蛋白具有相同的净电荷。在一些实施方案中,组合物包含具有基本上均质的电荷状态的多个含Fc蛋白,其中所述多个含Fc蛋白中的至少约80%,85%,90%,95%或100%具有相同的pI。在一些实施方案中,电荷与含Fc蛋白(例如氨基酸残基)相关。在一些实施方案中,电荷与含Fc蛋白的翻译后修饰相关。

[0125] 在一些实施方案中,组合物是药物组合物。在一些实施方案中,药物组合物为用于储存的形式。在一些实施方案中,药物组合物呈用于产品运输的形式。在一些实施方案中,药物组合物是冷冻的。在一些实施方案中,药物组合物被冻干。在一些实施方案中,药物组合物被重构。在一些实施方案中,药物组合物是施用组合物。在一些实施方案中,药物组合物为用于施用于有需要的个体的形式。

[0126] 在一些实施方案中,药物组合物是无菌药物组合物。无菌药物制剂根据本领域技术人员已知的药物级灭菌标准(例如,美国药典797,1072和1211章;加州商业和职业条例4127.7款;加州行政规则第16章1751条,联邦行政规则第21章21条,或这些规定的前美国对应规定),进行混合或制备。

[0127] 在一些实施方案中,药物组合物是稳定的制剂。如本文所用,“稳定”制剂是指其中的含Fc蛋白在储存时基本上保持物理和化学稳定性和完整性。用于测量蛋白质稳定性的各种分析技术在本领域中是可得,并且在例如Jones,Adv Drug Delivery Rev,10,1993中进行了综述。可以在选定的温度下评估一段选定时间的稳定性。例如,储存期间的聚集程度可以用作蛋白质稳定性的指标。因此,“稳定”制剂可以是这样的制剂,其中小于约10%且优选小于约5%的含Fc蛋白在制剂中以聚集物存在。

[0128] 在一些实施方案中,药物组合物是重构制剂。如本文所用,“重构”制剂是通过将冻干的含Fc蛋白制剂溶解在稀释剂中以使得含Fc蛋白分散在整个制剂中而制备的制剂。重构制剂适合于施用(例如静脉内或皮下施用)至有需要的个体。

[0129] 在一些实施方案中,药物组合物是等渗制剂。如本文所用,“等渗”制剂是具有与人血液基本相同的渗透压的制剂。等渗制剂通常具有约250至350mOsm的渗透压。术语“低渗”描述渗透压低于人血液的制剂。相应地,术语“高渗”用于描述渗透压高于人血液的制剂。

[0130] 在一些实施方案中,药物组合物具有特定的pH。在一些实施方案中,药物组合物的pH为约5-7,约5-6或约5-5.5。在一些实施方案中,药物组合物的pH为约5.3。在一些实施方案中,药物组合物的pH为约5.4。在一些实施方案中,药物组合物是经pH调节的。

[0131] 在一些实施方案中,药物组合物还包含药学上可接受的载体。如本文所用,“药学上可接受的载体”是指除活性成分之外的药物组合物中的成分。药学上可接受的载体包括但不限于缓冲剂,赋形剂,稳定剂或防腐剂。药学上可接受的载体通常在所采用的剂量和浓度下对接受者是无毒的,并且包括但不限于:缓冲剂,例如磷酸盐,柠檬酸盐和其他有机酸;抗氧化剂,包括抗坏血酸和甲硫氨酸;防腐剂(例如十八烷基二甲基苄基氯化铵;氯化六甲铵;苯扎氯铵;苄索氯铵;苯酚丁基或苄基醇;对羟基苯甲酸烷基酯,如对羟基苯甲酸甲酯或对羟基苯甲酸丙酯;儿茶酚;间苯二酚;环己醇;3-戊醇;和间甲酚);低分子量(少于约10个残基)多肽;蛋白质,如血清白蛋白,明胶或免疫球蛋白;亲水聚合物如聚乙烯吡咯烷酮;氨基酸如甘氨酸,谷氨酰胺,天冬酰胺,组氨酸,精氨酸或赖氨酸;单糖,二糖和其他碳水化合物,包括葡萄糖,甘露糖或糊精;螯合剂如EDTA;糖如蔗糖,甘露糖醇,海藻糖或山梨糖醇;成盐抗衡离子如钠;金属络合物(例如Zn-蛋白质复合物);和/或非离子表面活性剂如聚乙二醇(PEG)。本文中示例性的药学上可接受的载体进一步包括间质药物分散剂,例如可溶性中性活性透明质酸酶糖蛋白(sHASEGP),例如人可溶性PH-20透明质酸酶糖蛋白,例如rHuPH20(**HYLENEX**[®],Baxter International,Inc.)。在一些实施方案中,药学上可接受的载体选自乙酸钠,蔗糖,聚山梨酸酯(例如聚山梨酸酯20),琥珀酸钠,组氨酸HCl和氯化钠。

[0132] 在一些实施方案中,药物组合物还包含药学上可接受的酸。如本文所用,“药学上可接受的酸”包括在配制它们的浓度和方式下无毒的无机酸和有机酸。例如,合适的无机酸包括盐酸,高氯酸,氢溴酸,氢碘酸,硝酸,硫酸,磺酸,亚磺酸,对氨基苯磺酸,磷酸,碳酸等。合适的有机酸包括直链和支链烷基、芳族、环状、脂环族、芳基脂肪族、杂环、饱和、不饱和、单和二羧酸,包括,例如甲酸,乙酸,2-羟基乙酸,三氟乙酸,苯乙酸,三甲基乙酸,叔丁基乙酸,邻氨基苯甲酸,丙酸,2-羟基丙酸,2-氧代丙酸,丙二酸,环戊烷丙酸,环戊烷丙酸,3-苯基丙酸,丁酸,丁二酸,苯甲酸,3-(4-羟基苯甲酰)苯甲酸,2-乙酰氧基苯甲酸,抗坏血酸,肉桂酸,月桂基硫酸,硬脂酸,粘康酸,扁桃酸,琥珀酸,双羧萘酸,富马酸,苹果酸,马来酸,羟基马来酸,丙二酸,乳酸,柠檬酸,酒石酸,乙醇酸,葡糖酸(glyconic),葡糖酸(gluconic),丙酮酸,乙醛酸,草酸,甲磺酸,琥珀酸,水杨酸,邻苯二甲酸,palmoic,palmeic,硫氰酸,甲磺酸,乙磺酸,1,2-乙烷二磺酸,2-羟基乙烷磺酸,苯磺酸,4-氯苯磺酸,萘-2-磺酸,对甲苯磺酸,樟脑磺酸,4-甲基双环[2,2,2]-辛-2-烯-1-羧酸,葡庚糖酸,4,4'-亚甲基双-3-(羟基-2-烯-1-羧酸),羟基萘酸。

[0133] 在一些实施方案中,药物组合物还包含药学上可接受的碱。如本文所用,“药学上可接受的碱”包括在配制它们的浓度和方式下无毒的无机和有机碱。例如,合适的碱包括由无机碱形成金属例如锂,钠,钾,镁,钙,铵,铁,锌,铜,锰,铝,N-甲基葡糖胺,吗啉,哌啶形成的那些碱和有机无毒碱,包括伯胺,仲胺和叔胺,取代的胺,环胺和碱性离子交换树脂,例如异丙胺,三甲胺,二乙胺,三乙胺,三丙胺,乙醇胺,2-二乙基氨基乙醇,三甲胺,二环己胺,赖氨酸,精氨酸,组氨酸,咖啡因,普鲁卡因,hydrabamine,胆碱,甜菜碱,乙二胺,葡糖胺,甲基葡糖胺,可可碱,嘌呤,哌嗪,哌啶,N-乙基哌啶,聚胺树脂等。特别优选的有机无毒碱是异丙胺,二乙胺,乙醇胺,三甲胺,二环己胺,胆碱和咖啡因。

[0134] 可用于本发明的其它药学上可接受的酸和碱包括衍生自氨基酸例如组氨酸,甘氨酸,苯丙氨酸,天冬氨酸,谷氨酸,赖氨酸和天冬酰胺的那些。

[0135] 在一些实施方案中,药物组合物还包含药学上可接受的缓冲剂或盐,例如衍生自上述酸和碱的酸和碱加成盐的那些。特定的缓冲液和/或盐包括组氨酸,琥珀酸盐和乙酸盐。

[0136] 在一些实施方案中,药物组合物还包含药学上可接受的糖。如本文所用,“药学上可接受的糖”是与含Fc蛋白组合时显著防止或减少含Fc蛋白在储存时的化学和/或物理不稳定性的分子。当打算将制剂冻干然后重构时,“药学上可接受的糖”也可称为“冻干保护剂”。示例性的糖及其相应的糖醇包括:氨基酸如谷氨酸一钠或组氨酸;甲胺如甜菜碱;离液盐如硫酸镁;多元醇例如三元醇或更高分子量的糖醇,甘油,葡聚糖,赤藓糖醇,甘油,阿拉伯糖醇,木糖醇,山梨醇和甘露糖醇;丙二醇;聚乙二醇;**PLURONICS®**;以及它们的组合。另外的示例性冻干保护剂包括甘油和明胶,以及糖类蜜二糖,松三糖,棉子糖,甘露三糖和水苏糖。还原糖的实例包括葡萄糖,麦芽糖,乳糖,麦芽酮糖,异麦芽酮糖和乳果糖。非还原糖的实例包括选自糖醇和其它直链多元醇的多羟基化合物的非还原糖苷。优选的糖醇是单糖苷,特别是通过还原二糖如乳糖,麦芽糖,乳果糖和麦芽酮糖而获得的那些化合物。糖苷侧基可以是葡糖苷或半乳糖苷。糖醇的其他例子是葡萄糖醇,麦芽糖醇,乳糖醇和异麦芽酮糖。优选的药学上可接受的糖是非还原糖海藻糖或蔗糖。药学上可接受的糖以“保护量”(例如冻干前)加入制剂中,这意味着蛋白质在储存期间(例如在重构和储存之后)基本上保

持其物理和化学稳定性和完整性。

[0137] 在一些实施方案中,药物组合物还包含药学上可接受的防腐剂。如本文所用,“药学上可接受的防腐剂”是可以加入本文制剂中以降低细菌活性的化合物。潜在的防腐剂的实例包括但不限于十八烷基二甲基苄基氯化铵,氯化六甲双铵,苯扎氯铵(其中烷基为长链化合物的烷基苄基二甲基氯化铵的混合物)和苄索氯铵。其他类型的防腐剂包括芳族醇如苯酚,丁基和苯甲醇,对羟基苯甲酸烷基酯如对羟基苯甲酸甲酯或对羟基苯甲酸丙酯,儿茶酚,间苯二酚,环己醇,3-戊醇和间甲酚。

[0138] 在一些实施方案中,包含多个含Fc蛋白的组合物是细胞培养基,其中基本上所有含Fc蛋白在每个Fc结构域上具有C-末端赖氨酸。在一些实施方案中,细胞培养基是营养培养基。营养培养基含有宿主细胞生长所需的所有元素。在一些实施方案中,细胞培养基是基本培养基。基本培养基含有可用于宿主细胞生长的最少营养素,例如,通常不存在氨基酸。在一些实施方案中,细胞培养基是选择性培养基。选择性培养基包含抑制选择生物体生长的试剂。

[0139] 在一些实施方案中,细胞培养基还包含用于细胞支持和/或生长的细胞培养基营养。在一些实施方案中,细胞培养基营养物选自例如:蛋白质;肽;氨基酸;碳水化合物;金属和矿物质,例如钙,镁,铁;痕量金属,例如磷酸盐和硫酸盐;缓冲剂;pH指示剂,例如酚红,溴甲酚紫;和抗微生物剂。

[0140] 在一些实施方案中,细胞培养基还包含宿主细胞。在一些实施方案中,培养基基本上没有宿主细胞。

[0141] 在一些实施方案中,组合物是细胞裂解物。在一些实施方案中,细胞裂解物包含多个含Fc蛋白和宿主细胞组分。在一些实施方案中,细胞裂解物是离心后的细胞裂解物。在一些实施方案中,细胞裂解物包含细胞裂解物的沉淀部分和细胞裂解物的上清部分。在一些实施方案中,细胞裂解物包含细胞裂解物的沉降部分和细胞裂解物的上清液部分。

[0142] 在一些实施方案中,组合物是来自蛋白质纯化柱的洗脱液。如本文所用,“洗脱物”是指通过蛋白质纯化柱的任何流体。在一些实施方案中,洗脱物包含从流通液分离的流体。在一些实施方案中,洗脱物包含从洗涤液分离的流体。在一些实施方案中,洗脱物包含从一种或多种洗涤液分离的流体。在一些实施方案中,洗脱物包含从洗脱液体分离的流体。

[0143] 用于富集含Fc蛋白的蛋白质纯化柱是本领域已知的。以下是蛋白质纯化柱的示例性类型:免疫亲和,蛋白A,蛋白G,离子交换,反相,阳离子交换,强阳离子交换,阴离子交换,疏水,混合模式,羟基磷灰石和凝胶过滤。

[0144] 在一些实施方案中,组合物是含Fc蛋白的文库,其中所述多个含Fc蛋白中的至少两个含Fc蛋白是不同的。在一些实施方案中,文库包含至少两个结合不同抗原的含Fc蛋白。在一些实施方案中,文库包含至少两个结合不同表位的含Fc蛋白。在一些实施方案中,不同的含Fc蛋白包含在不同的容器中。

[0145] 在一些实施方案中,本发明提供了包含至少2,3,4,5,10,30,100,250,500,750,1000,2500,5000,7500,10000,25000,50000,75000,100000,250000,500000,750000,1000000,2500000,5000000,7500000,10000000,或10000000个以上的不同含Fc蛋白的文库。

[0146] 批料(batch)

[0147] 本申请提供本文实施方案中描述的任何组合物的大规模批料(例如,商业批料或生产规模批料)。

[0148] 在一些实施方案中,批料包含至少约5g,10g,50g,100g,200g,300g,400g,500g,600g,700g,800g,900g,1,000g,1,500g,2,000g,2,500g,3,000g,3,500g,4,000g,4,500g或5,000g的多个含Fc蛋白,其中所述多个含Fc蛋白中的每个含Fc蛋白在每个Fc结构域上具有C-末端赖氨酸。在一些实施方案中,批料包含至少约5-5,000g,50-4,000g或约100-1,000g的多个含Fc蛋白,其中所述多个含Fc蛋白中的每个含Fc蛋白在每个Fc结构域上具有C-末端赖氨酸。

[0149] 在一些实施方案中,批料为用于药物储存的形式。在一些实施方案中,批料为用于产品运输的形式。在一些实施方案中,批料为用于向有需要的个体施用的形式。在一些实施方案中,批料是冻干的。在一些实施方案中,批料不与药剂缀合。在一些实施方案中,批料与药剂缀合。在一些实施方案中,批料进一步包含制剂组分。

[0150] 在一些实施方案中,批料是细胞培养基。在一些实施方案中,批料是细胞裂解物。

[0151] 在一些实施方案中,批料或其一部分在容器中。在一些实施方案中,批料或其部分在小瓶中。在一些实施例中,批料或其部分在多个小瓶中。在一些实施方案中,批料或其部分在注射器中。

[0152] 在一些实施方案中,批料或其部分在多个小瓶中,其中每个小瓶包含多个含Fc蛋白,其中每个含Fc蛋白在每个Fc结构域上具有C-末端赖氨酸。在一些实施方案中,批料或其部分在多个小瓶中,其中每个小瓶包含多个含Fc蛋白,其中每个小瓶中的多个含Fc蛋白中的至少约90%,91%,92%,93%,94%,95%,96%,97%,98%,99%或100%在每个Fc结构域上具有C-末端赖氨酸。在一些实施方案中,批料或其部分在多个小瓶中,其中至少约50%,55%,60%,65%,70%,75%,80%,85%,90%95%或100%的小瓶包含的多个含Fc蛋白有至少约90%,91%,92%,93%,94%,95%,96%,97%,98%,99%或100%在每个Fc结构域上包含C-末端赖氨酸。

[0153] 在一些实施方案中,批料等分成单位剂量。如本文所用,“单位剂量”是意图作为单个单位剂量施用的含Fc蛋白的量。在一些实施方案中,单个单位剂量是约1至约500mg的含Fc蛋白。在一些实施例中,单位剂量被包装在容器中。在一些实施方案中,单位剂量被包装在小瓶中。

[0154] 在一些实施方案中,商业批料的大小不大于临床批料大小的约1、2、3、4、5、6、7、8、9或10倍。如本文所用,“商业批料”是指在了为了商业生产和/或分配的目的而完成的一个或多个生产过程中产生的含Fc蛋白的量。如本文所用,“临床批料”是指在了为了临床测试目的而完成的一个或多个生产过程中产生的含Fc蛋白的量。在一些实施方案中,商业批料的大小不大于临床批料大小的10倍。

[0155] 在一些实施方案中,容器包含如本文所述的商业批料的等分样,其中所述商业批料包含含有含Fc蛋白的组合物,其中所述含Fc蛋白的每个Fc结构域具有C-末端赖氨酸。在一些实施方案中,小瓶包含如本文所述的商业批料的等分样,其中所述商业批料包含含有含Fc蛋白的组合物,其中所述含Fc蛋白的每个Fc结构域具有C-末端赖氨酸。在一些实施方案中,注射器包含如本文所述的商业批料的等分样,其中商业批料包含含有含Fc蛋白的组合物,其中含Fc蛋白的每个Fc结构域具有C-末端赖氨酸。

[0156] 细胞培养系统

[0157] 本发明提供包含本文实施方案中描述的任何宿主细胞的细胞培养系统。

[0158] 在一些实施方案中,细胞培养系统将宿主细胞维持在用于生产含Fc蛋白的环境中。在一些实施方案中,细胞培养系统将宿主细胞维持在生长环境中。

[0159] 在一些实施方案中,细胞培养系统包含种子培养物。在一些实施方案中,细胞培养系统包含接种物培养物。在一些实施方案中,接种物培养物是一级接种物培养物。在一些实施方案中,接种物培养物是二级接种物。在一些实施方案中,细胞培养系统包含生产培养物。

[0160] 在一些实施方案中,细胞培养系统包含用于在指定的环境中维持宿主细胞的机制。在一些实施方案中,指定的环境是指定的温度。在一些实施方案中,指定温度为约15℃至约45℃。在一些实施方案中,指定温度为约30℃。在一些实施方案中,指定的环境是指定的pH。在一些实施例中,指定的环境是指定的溶氧浓度。在一些实施方案中,指定的环境是指定的营养物水平。

[0161] 在一些实施方案中,细胞培养物还包含培养基。在一些实施方案中,细胞培养系统包含种子培养物。在一些实施方案中,培养基是接种物培养物。在一些实施方案中,接种物培养基是一级接种物培养基。在一些实施方案中,接种物培养基是二级接种物培养基。在一些实施方案中,培养基是生产培养基。

[0162] 在一些实施方案中,培养系统包含如本文实施方案中所述的含Fc蛋白。在一些实施方案中,培养系统包含如本文实施方案中所述的多个含Fc蛋白。在一些实施方案中,培养系统包含如本文实施方案中所述的包含含Fc蛋白的组合物。

[0163] 在一些实施方案中,培养系统包含含Fc蛋白,其中含Fc蛋白的每个Fc结构域具有C-末端赖氨酸。在一些实施方案中,培养系统包含多个含Fc蛋白,其中所述多个含Fc蛋白中的每个含Fc蛋白具有C-末端赖氨酸Fc结构域。在一些实施方案中,培养系统包含含有多个含Fc蛋白的组合物,其中所述多个含Fc蛋白中的每个含Fc蛋白在每个Fc结构域上具有C-末端赖氨酸。

[0164] 产生具有降低的CpD表达水平的宿主细胞的方法

[0165] 本申请提供了生产本文实施方案中描述的任何宿主细胞的方法,其中所述方法包括使宿主细胞的CpD基因失活。

[0166] 通常,生产宿主细胞的方法包括使宿主细胞的CpD基因失活。在一些实施方案中,生产宿主细胞的方法(其中宿主细胞具有CpD表达水平的降低)包括使用siRNA系统使CpD基因失活。在一些实施方案中,生产宿主细胞的方法包括使用siRNA系统使CpD基因失活,所述siRNA系统包含长度约10至200个核苷酸、或约10至100个核苷酸、或约15至100个核苷酸、或约10至60个核苷酸、或约15至60个核苷酸、或约10至50个核苷酸、或约15至50个核苷酸、或约10至30个核苷酸、或约15至30个核苷酸的siRNA核苷酸序列。在一些实施方案中,生产宿主细胞的方法包括使用包含约10-25个核苷酸长度的siRNA核苷酸序列的siRNA系统,使CpD基因失活。在一些实施方案中,生产宿主细胞的方法包括使用包含约15-25个核苷酸长度的siRNA核苷酸序列的siRNA系统,使CpD基因失活。在一些实施方案中,产生宿主细胞的方法包括使用siRNA系统使CpD基因失活,所述siRNA系统包含长度为至少约10,至少约15,至少约20或至少约25个核苷酸的siRNA核苷酸序列。在一些实施方案中,生产宿主细胞的方法包

括使用siRNA系统使CpD基因失活,所述siRNA系统包含与CpD mRNA分子的区域至少约80%,至少约85%,至少约90%,至少约95%或100%互补的核苷酸序列。在一些实施方案中,生产宿主细胞的方法包括使用siRNA系统使CpD基因失活,所述siRNA系统包含与CpD前mRNA分子的区域至少约80%,至少约85%,至少约90%,至少约95%或100%互补的siRNA核苷酸序列。在一些实施方案中,生产宿主细胞的方法包括使用包含双链RNA分子的siRNA系统,使CpD基因失活。在一些实施方案中,生产宿主细胞的方法包括使用包含单链RNA分子的siRNA系统,使CpD基因失活。示例性的CpD siRNA核苷酸序列列于表1中。

[0167] 在一些实施方案中,生产宿主细胞的方法包括将siRNA系统递送至宿主细胞。在一些实施方案中,生产宿主细胞的方法包括将siRNA系统递送至宿主细胞,其中所述siRNA系统包含siRNA核苷酸序列。在一些实施方案中,生产宿主细胞的方法包括使用电穿孔将siRNA系统递送至宿主细胞。在一些实施方案中,生产宿主细胞的方法包括使用转染技术将siRNA系统递送至宿主细胞。在一些实施方案中,生产宿主细胞的方法包括使用病毒将siRNA系统递送至宿主细胞。

[0168] 在一些实施方案中,生产宿主细胞的方法(其中宿主细胞具有降低水平的CpD表达)还包括确定宿主细胞中CpD基因失活的水平。在一些实施方案中,确定CpD基因失活水平包括在将siRNA核苷酸序列递送至宿主细胞之前确定CpD表达水平。在一些实施方案中,确定CpD基因失活水平包括在将siRNA核苷酸序列递送至宿主细胞后确定CpD表达水平。在一些实施方案中,在RNA水平确定CpD表达水平。在一些实施方案中,确定CpD基因失活水平的方法包括使用PCR确定CpD表达水平。在一些实施方案中,在蛋白质水平确定CpD表达水平。在一些实施方案中,确定CpD基因失活水平的方法包括使用免疫组织化学确定CpD表达水平。在一些实施方案中,确定CpD基因失活水平的方法包括使用Western印迹确定CpD表达水平。在一些实施方案中,确定CpD基因失活水平的方法包括使用流式细胞术确定CpD表达水平。在一些实施方案中,CpD基因失活水平通过将siRNA递送后的CpD表达水平与对照值进行比较来确定。在一些实施方案中,通过将siRNA递送后的CpD表达水平与野生型值进行比较来确定CpD基因失活水平。在一些实施方案中,CpD基因失活水平通过比较递送siRNA后CpD表达水平与递送CpD siRNA前CpD表达水平来确定。

[0169] 表2中提供了用于确定RNA水平的CpD水平的示例性引物和探针。

[0170] 表2. 用于确定CpD的RNA表达水平的示例性核苷酸序列。

[0171]

	SEQ ID NO.:	核苷酸序列:
CpD 正向引物	1	CCC ACA CAT TAC AAA TCT TAC CA
CpD 反向引物	2	GAG ATT TCG AGG GAC CAA AT
CpD 探针	3	TTG GGA CAG AGT GCT GAG TAT CGT CA

[0172] 在一些实施方案中,生产宿主细胞的方法(其中宿主细胞具有降低水平的CpD表达)包括使用shRNA系统使CpD基因失活。在一些实施方案中,生产宿主细胞的方法包括使用shRNA系统使CpD基因失活,所述shRNA系统包含长度约10至200个核苷酸,或约10至100个核

核苷酸、或约15至100个核苷酸、或约10至60个核苷酸、或约15至60个核苷酸、或约10至50个核苷酸、或约15至50个核苷酸、或约10至30个核苷酸、或约15至30个核苷酸的shRNA核苷酸序列。在一些实施方案中,生产宿主细胞的方法包括使用包含约10-25个核苷酸长度的shRNA核苷酸序列的shRNA系统,使CpD基因失活。在一些实施方案中,生产宿主细胞的方法包括使用包含约15-25个核苷酸长度的shRNA核苷酸序列的shRNA系统,使CpD基因失活。在一些实施方案中,生产宿主细胞的方法包括使用包含长度为至少约10,至少约15,至少约20或至少约25个核苷酸的shRNA核苷酸序列的shRNA系统,使CpD基因失活。在一些实施方案中,生产宿主细胞的方法包括使用shRNA系统使CpD基因失活,所述shRNA系统包含与CpD mRNA分子的区域至少约80%,至少约85%,至少约90%,至少约95%或100%互补的核苷酸序列。在一些实施方案中,生产宿主细胞的方法包括使用shRNA系统使CpD基因失活,所述shRNA系统包含与CpD前mRNA分子的区域至少约80%,至少约85%,至少约90%,至少约95%或100%互补的shRNA核苷酸序列。在一些实施方案中,生产宿主细胞的方法包括使用包含双链RNA分子的shRNA系统使CpD基因失活。在一些实施方案中,生产宿主细胞的方法包括使用包含单链RNA分子的shRNA系统使CpD基因失活。示例性的CpD shRNA核苷酸序列列于表1中。

[0173] 在一些实施方案中,生产宿主细胞的方法包括将shRNA系统递送至宿主细胞。在一些实施方案中,生产宿主细胞的方法包括将shRNA系统递送至宿主细胞,其中shRNA系统包含shRNA核苷酸序列。在一些实施方案中,生产宿主细胞的方法包括使用电穿孔将shRNA系统递送至宿主细胞。在一些实施方案中,生产宿主细胞的方法包括使用转染技术将shRNA系统递送至宿主细胞。在一些实施方案中,生产宿主细胞的方法包括使用病毒将shRNA系统递送至宿主细胞。

[0174] 在一些实施方案中,生产宿主细胞的方法(其中宿主细胞具有降低的CpD表达水平)还包括确定宿主细胞中CpD基因失活的水平。在一些实施方案中,确定CpD基因失活水平包括在将shRNA核苷酸序列递送至宿主细胞之前确定CpD表达水平。在一些实施方案中,确定CpD基因失活水平包括在将shRNA核苷酸序列递送至宿主细胞后确定CpD表达水平。在一些实施方案中,在RNA水平确定CpD表达水平。在一些实施方案中,确定CpD基因失活水平的方法包括使用PCR确定CpD表达水平。在一些实施方案中,在蛋白质水平确定CpD表达水平。在一些实施方案中,确定CpD基因失活水平的方法包括使用免疫组织化学确定CpD表达水平。在一些实施方案中,确定CpD基因失活水平的方法包括使用Western印迹确定CpD表达水平。在一些实施方案中,确定CpD基因失活水平的方法包括使用流式细胞术确定CpD表达水平。在一些实施方案中,通过比较递送shRNA后的CpD表达水平与对照值来确定CpD基因失活水平。在一些实施方案中,通过将shRNA递送后的CpD表达水平与野生型值进行比较来确定CpD基因失活水平。在一些实施方案中,通过比较递送shRNA之后的CpD表达水平与递送CpD shRNA之前的CpD表达水平来确定CpD基因失活水平。

[0175] 表2中提供了用于在RNA水平确定CpD水平的示例性引物和探针。

[0176] 本申请还提供,通过基因缺失或基因添加或取代,产生宿主细胞的方法。例如,方法包括但不限于使用CRISPR,TALEN,ZFN和大范围核酸酶(meganuclease)系统。

[0177] 通常,CRISPR系统包含caspase蛋白如Cas9、和包含与目的序列互补的核苷酸序列(称为引导序列)的RNA序列。caspase和RNA序列形成识别宿主细胞DNA序列的复合体,随后caspase的核酸酶活性允许切割DNA链。caspase同种型具有单链DNA或双链DNA核酸酶活性。

设计在CRISPR系统中使用的引导RNA序列和引导RNA序列的数量,可以允许除去特定的基因片段和/或添加DNA序列。

[0178] 在一些实施方案中,生产宿主细胞的方法(其中宿主细胞具有降低的CpD表达水平)包括使用CRISPR系统使CpD基因失活。在一些实施方案中,生产宿主细胞的方法包括使用包含编码载体的CRISPR系统,使CpD基因失活。在一些实施方案中,生产宿主细胞的方法包括使用包含编码载体的CRISPR系统使CpD基因失活,所述载体含有DNA内切核酸酶基因。在一些实施方案中,生产宿主细胞的方法包括使用包含编码载体的CRISPR系统使CpD基因失活,所述载体包含CAS基因。在一些实施方案中,生产宿主细胞的方法包括使用包含编码载体的CRISPR系统使CpD基因失活,所述载体包含CAS9基因。在一些实施方案中,生产宿主细胞的方法包括使用包含编码载体的CRISPR系统使CpD基因失活,所述载体编码CAS9基因。在一些实施方案中,生产宿主细胞的方法包括使用包含Cas蛋白的CRISPR系统使CpD基因失活。在一些实施方案中,生产宿主细胞的方法包括使用包含Cas9蛋白的CRISPR系统,使CpD基因失活。在一些实施方案中,生产宿主细胞的方法包括使用包含编码载体的CRISPR系统使CpD基因失活,其中所述载体编码能够与Cas9蛋白相互作用的RNA分子。在一些实施方案中,生产宿主细胞的方法包括使用包含编码载体的CRISPR系统使CpD基因失活,其中所述载体编码包含引导RNA(gRNA)单位的RNA分子,其中所述gRNA单位包含与一部分CpD基因序列互补的核苷酸序列。在一些实施方案中,生产宿主细胞的方法包括使用包含RNA分子的CRISPR系统使CpD基因失活,其中所述RNA分子包含gRNA单位,其中所述gRNA单位包含与一部分CpD基因序列互补的核苷酸序列。在一些实施方案中,生产宿主细胞的方法包括使用包含编码载体的CRISPR系统使CpD基因失活,其中所述载体编码包含反式激活crRNA(tracrRNA)单位的RNA分子。在一些实施方案中,生产宿主细胞的方法包括使用包含RNA分子的CRISPR系统使CpD基因失活,其中所述RNA分子包含tracrRNA单位。

[0179] 在一些实施方案中,生产宿主细胞的方法包括使用包含编码载体的CRISPR系统使CpD基因失活,其中所述载体编码包含gRNA单位和tracrRNA单位的RNA分子,其中所述gRNA单位包含与部分基因序列互补的核苷酸序列。在一些实施方案中,生产宿主细胞的方法包括使用包含RNA分子的CRISPR系统使CpD基因失活,其中所述RNA分子包含gRNA单位和tracrRNA单位,其中所述gRNA单位包含与部分基因序列互补的核苷酸序列。在一些实施方案中,生产宿主细胞的方法包括使用CRISPR系统使CpD基因失活,所述CRISPR系统包含:a)包含gRNA单位的第一RNA分子,其中所述gRNA单位包含与一部分CpD基因序列互补的第一核苷酸序列;和b)包含gRNA单位的第二RNA分子,其中所述gRNA单位包含与一部分CpD基因序列互补的第二核苷酸序列。在一些实施方案中,第一核苷酸序列和第二核苷酸序列是不同的。在一些实施方案中,与第一核苷酸序列互补的CpD基因部分,和与第二核苷酸序列互补的CpD基因部分,位于不同的区域位置。

[0180] 包含在gRNA序列中的示例性CpD靶序列列于表3中。

[0181] 表3. 包含在gRNA中的示例性CpD靶序列。

[0182]

SEQ ID NO.:	CpD靶序列:
21	GAA GAC GAG ACT TTC AAA GAC GG
22	TCA GTT AGG TGG CTT AGG TTC GG

[0183] 在一些实施方案中,生产宿主细胞的方法包括将CRISPR系统递送至宿主细胞。在一些实施方案中,生产宿主细胞的方法包括使用递送载体递送包含CRISPR系统的载体。在一些实施方案中,递送载体是病毒载体。在一些实施方案中,递送载体是慢病毒。在一些实施方案中,递送载体是腺病毒。在一些实施方案中,载体包含启动子。

[0184] 通常,TALEN系统包含一个或多个限制性核酸酶和两个或更多个蛋白的复合物,所述复合物允许识别DNA序列和随后切割双链DNA。TALEN系统的蛋白质复合物包含多个转录激活因子样效应物 (TALE) 和限制性核酸酶的结构域,其中每个TALE都识别特定的核苷酸。通常,设计TALEN系统,使得两个蛋白质复合物(各自包含TALE和限制性核酸酶的结构域)可以独自结合DNA序列以允许限制性核酸酶的两个结构域(从每个蛋白质复合物各得一个)形成活性核酸酶并切割特定的DNA序列。通过设计TALEN系统中的蛋白复合物的数量和待切割的序列,允许除去特定基因区段和/或添加DNA序列。

[0185] 在一些实施方案中,生产宿主细胞的方法(其中宿主细胞具有降低水平的CpD表达)包括使用TALEN系统使CpD基因失活。在一些实施方案中,生产宿主细胞的方法包括使用包含第一TALEN单位的TALEN系统使CpD基因失活。在一些实施例中,第一TALEN单位包括第一TALEN结合单位。在一些实施方案中,第一TALEN结合单位包含至少一个转录激活因子样效应物 (TALE) 和第一核酸酶结构域。在一些实施方案中,第一TALEN结合单位包含至少3、4、5、6、7、8、9、10、11、12、13、14或15个TALE和第一核酸酶结构域,其中所述TALE连接在一起,并且其中连接的TALE识别CpD核苷酸序列的一部分。在一些实施方案中,所述第一TALEN结合单位包含至少3、4、5、6、7、8、9、10、11、12、13、14或15个TALE和第一核酸酶结构域,其中所述TALE连接在一起,其中连接的TALE识别CpD核苷酸序列的一部分,并且其中连接的TALE进一步连接至第一核酸酶结构域。在一些实施例中,第一TALEN单位还包含第二TALEN结合单位。在一些实施方案中,第二TALEN结合单位包含至少一个转录激活因子样效应物 (TALE) 和第二核酸酶结构域。在一些实施方案中,第二TALEN结合单位包含至少3、4、5、6、7、8、9、10、11、12、13、14或15个TALE和第二核酸酶结构域,其中所述TALE连接在一起,并且其中连接的TALE识别CpD核苷酸序列的一部分。在一些实施方案中,第二TALEN结合单位包含至少3、4、5、6、7、8、9、10、11、12、13、14或15个TALE和第二核酸酶结构域,其中所述TALE连接在一起,其中连接的TALE识别CpD核苷酸序列的一部分,并且其中连接的TALE进一步连接至第二核酸酶结构域。在一些实施方案中,第一TALEN结合单位和第二TALEN结合单位结合CpD基因的不同序列。在一些实施方案中,第一核酸酶结构域是内切核酸酶的结构域。在一些实施方案中,第一核酸酶结构域是限制性内切核酸酶的结构域。在一些实施方案中,第一核酸酶结构域是FokI的结构域。在一些实施方案中,第二核酸酶结构域是内切核酸酶的结构域。在一些实施方案中,第二核酸酶结构域是限制性内切核酸酶的结构域。在一些实施方案中,第二核酸酶结构域是FokI的结构域。在一些实施方案中,第一核酸酶结构域和第二核酸酶结构域缔合以包含活性限制性内切核酸酶。在一些实施方案中,第一核酸酶结构域和第二核酸酶结构域缔合以包含活性FokI酶。

[0186] 在一些实施方案中,生产宿主细胞的方法包括使用还包含第二TALEN单位的TALEN系统使CpD基因失活。在一些实施例中,第二TALEN单位包括第三TALEN结合单位。在一些实施方案中,第三TALEN结合单位包含至少一个转录激活因子样效应物 (TALE) 和第三核酸酶结构域。在一些实施方案中,第三TALEN结合单位包含至少3、4、5、6、7、8、9、10、11、12、13、14

或15个TALE和第三核酸酶结构域,其中所述TALE连接在一起,并且其中连接的TALE识别CpD核苷酸序列的一部分。在一些实施方案中,第三TALEN结合单位包含至少3、4、5、6、7、8、9、10、11、12、13、14或15个TALE和第三核酸酶结构域,其中所述TALE连接在一起,其中连接的TALE识别CpD核苷酸序列的一部分,并且其中连接的TALE进一步连接至第三核酸酶结构域。在一些实施例中,第二TALEN单位还包含第四TALEN结合单位。在一些实施方案中,第四TALEN结合单位包含至少一个转录激活因子样效应物(TALE)和第四核酸酶结构域。在一些实施方案中,第四TALEN结合单位包含至少3、4、5、6、7、8、9、10、11、12、13、14或15个TALE和第四核酸酶结构域,其中所述TALE连接在一起,并且其中连接的TALE识别CpD核苷酸序列的一部分。在一些实施方案中,第四TALEN结合单位包含至少3、4、5、6、7、8、9、10、11、12、13、14或15个TALE和第四核酸酶结构域,其中所述TALE连接在一起,其中连接的TALE识别CpD核苷酸序列的一部分,并且其中连接的TALE进一步连接至第四核酸酶结构域。在一些实施方案中,第三TALEN结合单位和第四TALEN结合单位结合CpD基因的不同序列。在一些实施方案中,第三核酸酶结构域是内切核酸酶的结构域。在一些实施方案中,第三核酸酶结构域是限制性内切核酸酶的结构域。在一些实施方案中,第三核酸酶结构域是FokI的结构域。在一些实施方案中,第四核酸酶结构域是内切核酸酶的结构域。在一些实施方案中,第四核酸酶结构域是限制性内切核酸酶的结构域。在一些实施方案中,第四核酸酶结构域是FokI的结构域。在一些实施方案中,第三核酸酶结构域和第四核酸酶结构域缔合以包含活性限制性内切核酸酶。在一些实施方案中,第三核酸酶结构域和第四核酸酶结构域缔合以包含活性FokI酶。

[0187] 在一些实施方案中,生产宿主细胞的方法包括使用TALEN系统使CpD基因失活,其中所述宿主细胞具有降低水平的CpD表达,其中所述TALEN系统包含第一TALEN单位和第二TALEN单位,其结合CpD基因序列的不同的不重叠部分。

[0188] 在一些实施方案中,生产宿主细胞的方法(其中宿主细胞具有降低水平的CpD表达)包括使用TALEN系统使CpD基因失活,其中所述TALEN系统包含编码第一TALEN单位的编码载体。在一些实施方案中,生产宿主细胞的方法(其中宿主细胞具有降低水平的CpD表达)包括使用TALEN系统使CpD基因失活,其中所述TALEN系统包含编码第一TALEN单位的编码载体,其中所述TALEN系统包含编码第二TALEN单位的编码载体。在一些实施方案中,生产宿主细胞的方法(其中宿主细胞具有降低水平的CpD表达)包括使用TALEN系统使CpD基因失活,其中所述TALEN系统包含编码第一TALEN单位的编码载体,其中所述TALEN系统包括编码第一TALEN单位和第二TALEN单位的编码载体。

[0189] 在一些实施方案中,生产宿主细胞的方法(其中宿主细胞具有降低水平的CpD表达)包括使用TALEN系统使CpD基因失活,其中TALEN系统包含第一TALEN单位。在一些实施方案中,生产宿主细胞的方法(其中宿主细胞具有降低水平的CpD表达)包括使用TALEN系统使CpD基因失活,其中TALEN系统包含第二TALEN单位。在一些实施方案中,生产宿主细胞的方法(其中宿主细胞具有降低水平的CpD表达)包括使用TALEN系统使CpD基因失活,其中TALEN系统包含第一TALEN单位和第二TALEN单位。

[0190] 在一些实施方案中,第一TALEN结合单位包含一组连接的TALE,其中该组TALE识别核苷酸序列。在一些实施方案中,核苷酸序列是包含CpD基因的一部分的序列。在一些实施方案中,核苷酸序列是包含CpD基因启动子的一部分的序列。在一些实施方案中,核苷酸序

列是包含CpD基因侧翼序列的一部分的序列。在一些实施方案中,该序列与CpD基因的一部分至少约80%,至少约85%,至少约90%,至少约95%或100%同源。在一些实施方案中,所述序列与CpD基因启动子的一部分至少约80%,至少约85%,至少约90%,至少约95%或100%同源。在一些实施方案中,该序列与CpD基因侧翼序列的一部分至少约80%,至少约85%,至少约90%,至少约95%或100%同源。

[0191] 在一些实施方案中,生产宿主细胞的方法包括将TALEN系统递送至宿主细胞。在一些实施方案中,生产宿主细胞的方法包括使用递送载体递送包含TALEN系统的载体。在一些实施方案中,递送载体是病毒载体。在一些实施方案中,递送载体是慢病毒。在一些实施方案中,递送载体是腺病毒。

[0192] 通常,ZFN系统包含一个或多个限制性核酸酶和两个或两个以上蛋白质复合物,所述复合物允许识别DNA序列和随后切割双链DNA。ZFN系统的蛋白质复合物包含多个锌指(每个锌指识别特定的核苷酸密码子)和限制性核酸酶的结构域。通常,设计ZFN系统,使得两个蛋白质复合物(各自包含锌指和限制性核酸酶结构域)可以独自结合DNA序列从而允许限制性核酸酶的两个结构域(各来自一个蛋白质复合物)形成活性核酸酶并切割特定的DNA序列。设计ZFN系统中蛋白复合物的数量和待切割序列,可以允许去除基因的特定区段和/或添加DNA序列。

[0193] 在一些实施方案中,生产宿主细胞的方法(其中宿主细胞具有降低水平的CpD表达)包括使用ZFN系统使CpD基因失活。在一些实施方案中,产生宿主细胞的方法包括使用包含第一ZFN单位的ZFN系统使CpD基因失活。在一些实施例中,第一ZFN单位包含第一ZFN结合单位。在一些实施方案中,第一ZFN结合单位包含至少一个锌指和第一核酸酶结构域。在一些实施方案中,第一ZFN结合单位包含至少2、3、4、5、6、7、8、9或10个锌指和第一核酸酶结构域,其中锌指连接在一起,并且其中连接的锌指识别CpD核苷酸序列的一部分。在一些实施方案中,第一ZFN结合单位包含至少2、3、4、5、6、7、8、9或10个锌指和第一核酸酶结构域,其中锌指连接在一起,其中连接的锌指识别CpD核苷酸序列的一部分,并且其中连接的锌指还与第一核酸酶结构域连接。在一些实施例中,第一ZFN单位还包括第二ZFN结合单位。在一些实施方案中,第二ZFN结合单位包含至少一个锌指和第二核酸酶结构域。在一些实施方案中,第二ZFN结合单位包含至少2、3、4、5、6、7、8、9或10个锌指和第二核酸酶结构域,其中锌指连接在一起,并且其中连接的锌指识别CpD核苷酸序列的一部分。在一些实施方案中,第二ZFN结合单位包含至少2、3、4、5、6、7、8、9或10个锌指和第二核酸酶结构域,其中锌指连接在一起,其中连接的锌指识别CpD核苷酸序列的一部分,并且其中连接的锌指还与第二核酸酶结构域连接。在一些实施方案中,第一ZFN结合单位和第二ZFN结合单位结合CpD基因的不同序列。在一些实施方案中,第一核酸酶结构域是内切核酸酶的结构域。在一些实施方案中,第一核酸酶结构域是限制性内切核酸酶的结构域。在一些实施方案中,第一核酸酶结构域是FokI的结构域。在一些实施方案中,第二核酸酶结构域是内切核酸酶的结构域。在一些实施方案中,第二核酸酶结构域是限制性内切核酸酶的结构域。在一些实施方案中,第二核酸酶结构域是FokI的结构域。在一些实施方案中,第一核酸酶结构域和第二核酸酶结构域缔合以包含活性限制性内切核酸酶。在一些实施方案中,第一核酸酶结构域和第二核酸酶结构域缔合以包含活性FokI酶。

[0194] 在一些实施方案中,生产宿主细胞的方法包括使用还包含第二ZFN单位的ZFN系统

使CpD基因失活。在一些实施例中,第二ZFN单位包括第三ZFN结合单位。在一些实施方案中,第三ZFN结合单位包含至少一个锌指和第三核酸酶结构域。在一些实施方案中,第三ZFN结合单位包含至少2、3、4、5、6、7、8、9或10个锌指和第三核酸酶结构域,其中锌指连接在一起,并且其中连接的锌指识别CpD核苷酸序列的一部分。在一些实施方案中,第三ZFN结合单位包含至少2、3、4、5、6、7、8、9或10个锌指和第三核酸酶结构域,其中锌指连接在一起,其中连接的锌指识别CpD核苷酸序列的一部分,并且其中连接的锌指还与第三核酸酶结构域连接。在一些实施方案中,第二ZFN单位还包括第四ZFN结合单位。在一些实施方案中,第四ZFN结合单位包含至少一个锌指和第四核酸酶结构域。在一些实施方案中,所述第四ZFN结合单位包含至少2、3、4、5、6、7、8、9或10个锌指和第四核酸酶结构域,其中所述锌指连接在一起,并且其中所述连接的锌指识别CpD核苷酸序列的一部分。在一些实施方案中,第四ZFN结合单位包含至少2、3、4、5、6、7、8、9或10个锌指和第四核酸酶结构域,其中锌指连接在一起,其中连接的锌指识别CpD核苷酸序列的一部分,并且其中连接的锌指进一步连接至第四核酸酶结构域。在一些实施方案中,第三ZFN结合单位和第四ZFN结合单位结合CpD基因的不同序列。在一些实施方案中,第三核酸酶结构域是内切核酸酶的结构域。在一些实施方案中,第三核酸酶结构域是限制性内切核酸酶的结构域。在一些实施方案中,第三核酸酶结构域是FokI的结构域。在一些实施方案中,第四核酸酶结构域是内切核酸酶的结构域。在一些实施方案中,第四核酸酶结构域是限制性内切核酸酶的结构域。在一些实施方案中,第四核酸酶结构域是FokI的结构域。在一些实施方案中,第三核酸酶结构域和第四核酸酶结构域缔合以包含活性限制性内切核酸酶。在一些实施方案中,第三核酸酶结构域和第四核酸酶结构域缔合以包含活性FokI酶。

[0195] 在一些实施方案中,生产宿主细胞的方法(其中宿主细胞具有降低的CpD表达水平)包括使用ZFN系统使CpD基因失活,其中所述ZFN系统包含第一ZFN单位和第二TALEN单位,其结合CpD基因序列的不同且不重叠的部分。

[0196] 在一些实施方案中,生产宿主细胞的方法(其中宿主细胞具有降低的CpD表达水平)包括使用ZFN系统使CpD基因失活,其中ZFN系统包含编码第一ZFN单位的编码载体。在一些实施方案中,生产宿主细胞的方法(其中宿主细胞具有降低水平的CpD表达)包括使用ZFN系统使CpD基因失活,其中所述ZFN系统包含编码第一ZFN单位的编码载体,其中所述ZFN系统包含编码第二ZFN单位的编码载体。在一些实施方案中,生产宿主细胞的方法(其中宿主细胞具有降低水平的CpD表达)包括使用ZFN系统使CpD基因失活,其中所述ZFN系统包含编码第一ZFN单位的编码载体,其中所述ZFN系统包含编码第一ZFN单位和第二ZFN单位的编码载体。

[0197] 在一些实施方案中,生产宿主细胞的方法(其中宿主细胞具有降低水平的CpD表达)包括使用ZFN系统使CpD基因失活,其中所述ZFN系统包含第一ZFN单位。在一些实施方案中,生产宿主细胞的方法(其中宿主细胞具有降低水平的CpD表达)包括使用ZFN系统使CpD基因失活,其中ZFN系统包含第二ZFN单位。在一些实施方案中,生产宿主细胞(其中宿主细胞具有降低水平的CpD表达)的方法包括使用ZFN系统使CpD基因失活,其中所述ZFN系统包含第一ZFN单位和第二ZFN单位。

[0198] 在一些实施方案中,第一ZFN结合单位包含一组连接的锌指,其中该组锌指识别核苷酸序列。在一些实施方案中,核苷酸序列是包含CpD基因的一部分的序列。在一些实施方

案中,核苷酸序列是包含CpD基因启动子的一部分的序列。在一些实施方案中,核苷酸序列是包含CpD基因侧翼序列的一部分的序列。在一些实施方案中,该序列与CpD基因的一部分至少约80%,至少约85%,至少约90%,至少约95%或100%同源。在一些实施方案中,所述序列与CpD基因启动子的一部分至少约80%,至少约85%,至少约90%,至少约95%或100%同源。在一些实施方案中,该序列与CpD基因侧翼序列的一部分至少约80%,至少约85%,至少约90%,至少约95%或100%同源。

[0199] 在一些实施方案中,生产宿主细胞的方法包括将ZFN系统递送至宿主细胞。在一些实施方案中,生产宿主细胞的方法包括使用递送载体递送包含ZFN系统的载体。在一些实施方案中,递送载体是病毒载体。在一些实施方案中,递送载体是慢病毒。在一些实施方案中,递送载体是腺病毒。

[0200] 通常,大范围核酸酶系统包含一个或多个允许识别DNA序列和随后切割双链DNA的大范围核酸酶。

[0201] 在一些实施方案中,生产宿主细胞的方法(其中宿主细胞具有降低水平的CpD表达)包括使用大范围核酸酶系统使CpD基因失活。在一些实施方案中,大范围核酸酶具有长度为约8至约35个核苷酸碱基对的DNA识别序列。在一些实施方案中,大范围核酸酶具有长度为约12至约30个核苷酸碱基对的DNA识别序列。在一些实施方案中,DNA识别序列是包含CpD基因的一部分的序列。在一些实施方案中,DNA识别序列是包含CpD基因启动子的一部分的序列。在一些实施方案中,DNA识别序列是包含CpD基因侧翼序列的一部分的序列。

[0202] 在一些实施方案中,生产宿主细胞的方法包括将大范围核酸酶系统递送至宿主细胞。在一些实施方案中,生产宿主细胞的方法包括使用递送载体递送包含大范围核酸酶系统的载体。在一些实施方案中,递送载体是病毒载体。在一些实施方案中,递送载体是慢病毒。在一些实施方案中,递送载体是腺病毒。

[0203] 在一些实施方案中,生产宿主细胞的方法(其中宿主细胞具有降低水平的CpD表达)还包括确定CpD基因失活水平,例如基因缺失或基因添加或取代。在一些实施方案中,生产宿主细胞的方法包括确定CpD基因失活的水平,其中检测CpD基因缺失。在一些实施方案中,生产宿主细胞的方法包括确定CpD基因失活的水平,其中检测CpD基因添加或取代。在一些实施方案中,生产宿主细胞的方法包括确定CpD基因失活的水平,其中在失活宿主细胞中的基因之前确定CpD表达水平。在一些实施方案中,生产宿主细胞的方法包括确定CpD基因失活水平,其中在使用CRISPR系统使宿主细胞中的CpD基因失活后测定CpD表达水平。在一些实施方案中,生产宿主细胞的方法包括确定CpD基因失活水平,其中在使用TALEN系统使宿主细胞中的CpD基因失活后测定CpD表达水平。在一些实施方案中,生产宿主细胞的方法包括确定CpD基因失活水平,其中在使用ZFN系统使宿主细胞中的CpD基因失活后测定CpD表达水平。在一些实施方案中,生产宿主细胞的方法包括确定CpD基因失活水平,其中在使用大范围核酸酶系统使宿主细胞中的CpD基因失活后测定CpD表达水平。

[0204] 在一些实施方案中,在DNA水平确定CpD表达水平。在一些实施方案中,在RNA水平确定CpD表达水平。在一些实施方案中,生产宿主细胞的方法包括使用PCR确定CpD表达水平。在一些实施方案中,生产宿主细胞的方法包括使用PCR确定CpD表达水平,其中检测变体序列。在一些实施方案中,生产宿主细胞的方法包括使用qPCR确定CpD表达水平。在一些实施方案中,生产宿主细胞的方法包括使用qPCR确定CpD表达水平。

[0205] 在一些实施方案中,在蛋白质水平确定CpD表达水平。在一些实施方案中,生产宿主细胞的方法包括使用免疫组织化学确定CpD表达水平。在一些实施方案中,生产宿主细胞的方法包括使用Western印迹确定CpD表达水平。在一些实施方案中,生产宿主细胞的方法包括使用流式细胞术确定CpD表达水平。

[0206] 在一些实施方案中,通过将CpD基因失活后CpD表达水平与对照值进行比较来确定CpD基因失活水平。在一些实施方案中,通过比较CpD基因失活后的CpD表达水平与CpD基因失活前的CpD表达水平来确定CpD基因失活水平。

[0207] 还提供了评估宿主细胞是否合适表达含Fc蛋白的方法,其包括测定CpD表达,其中降低的CpD表达水平指示适合性。

[0208] 制备含Fc蛋白的方法

[0209] 本申请提供制备本文实施方案中描述的含Fc蛋白的方法。

[0210] 在一些实施方案中,制备含Fc蛋白的方法包括:a) 培养宿主细胞;和b) 获得宿主细胞表达的含Fc蛋白。

[0211] 在一些实施方案中,制备含Fc蛋白的方法包括:a) 用包含编码含Fc蛋白的核酸的表达载体转化宿主细胞;b) 培养宿主细胞;和c) 获得宿主细胞表达的含Fc蛋白。

[0212] 使用表达载体转化宿主细胞的方法在本领域中是公知的。参见例如Kim等,Anal Bioanal Chem,397,2010。用于转染宿主细胞的方法,包括但不限于转染,感染,磷酸钙共沉淀,电穿孔,显微注射,脂转染,DEAE-葡聚糖介导的转染或其他已知技术。选择的方法部分取决于要使用的宿主细胞的类型。这些方法和其他合适的方法是本领域技术人员所熟知的。

[0213] 还提供包含至少一个本文所述的含Fc蛋白的、质粒、表达载体、转录或表达盒形式的表达系统和构建体。在某些实施方案中,本文提供的质粒,表达载体,转录或表达盒包含编码至少一个含Fc蛋白的多核苷酸。

[0214] 在一些实施方案中,宿主细胞中使用的表达载体含有用于质粒维持和克隆和表达外源核苷酸序列的序列。在某些实施方案中统称为“侧翼序列”的此类序列通常将包括一个或多个以下核苷酸序列:启动子,一个或多个增强子序列,复制起点,转录终止序列,含有供体和受体剪接位点的完整内含子序列,编码用于多肽分泌的前导序列的序列,核糖体结合位点,聚腺苷酸化序列,用于插入待表达的编码含Fc蛋白的核酸的多接头区,和选择标记元件。下面讨论这些序列中的每一个。

[0215] 侧翼序列可以是同源的(即来自与宿主细胞相同的物种和/或株),异源的(即来自与宿主细胞的物种或株不同的物种),杂交体(即,侧翼序列来自多个来源),合成的或天然的。因此,侧翼序列的来源可以是任何真核生物,任何脊椎动物或无脊椎动物或任何植物,条件是侧翼序列在宿主细胞机制中起作用并且可以被宿主细胞机制激活。

[0216] 可用于本发明载体中的侧翼序列可通过本领域众所周知的几种方法中的任何一种获得。典型地,本文有用的侧翼序列先前已通过作图和/或通过限制性内切酶消化而鉴定,并因此可使用适当的限制性内切核酸酶从适当的组织来源分离。在一些情况下,侧翼序列的全部核苷酸序列可以是已知的。这里,侧翼序列可以使用本文所述的用于核酸合成或克隆的方法来合成。

[0217] 无论侧翼序列的全部或仅一部分是已知的,其均可以使用聚合酶链式反应(PCR)

和/或通过用合适的探针(例如来自相同或另一物种的寡核苷酸和/或侧翼序列片段)筛选基因组文库来获得。在侧翼序列未知的情况下,含有侧翼序列的DNA片段可以从可能含有例如编码序列或甚至另一个或多个基因的较大DNA片段中分离。可以通过限制性核酸内切酶消化产生合适的DNA片段,然后使用琼脂糖凝胶纯化,Qiagen柱层析(Chatsworth,CA),或本领域技术人员已知的其他方法来进行分离。对于本领域的普通技术人员来说,选择合适的酶来实现该目的将是显而易见的。

[0218] 复制起点通常是商业购买的部分表达载体,并且复制起点有助于在宿主细胞中扩增载体。如果选择的载体不含复制起点位点,可以基于已知序列化学合成,并连接到载体中。例如,各种病毒起点(例如SV40,多瘤病毒,腺病毒,水泡性口炎病毒(VSV)或乳头瘤病毒如HPV或BPV)可用于在哺乳动物细胞中克隆载体。通常,哺乳动物表达载体不需要复制起点组分(例如,通常使用SV40起点仅是因为它也含有病毒早期启动子)。

[0219] 转录终止序列通常位于多肽编码区末端3'处并用于终止转录。尽管该序列容易从文库克隆或者甚至作为载体的一部分商业化购买,但其也可以使用用于核酸合成的方法例如本文所述的方法容易地合成。

[0220] 选择性标记基因编码对生长在选择性培养基中的宿主细胞的存活和生长所必需的蛋白质。典型的选择标记基因编码蛋白质,其(a)赋予对抗生素或其他毒素的抗性;(b)补充细胞的营养缺陷;或(c)提供复杂或成分确定培养基中不能提供的关键营养素。具体的选择标记可以是卡那霉素抗性基因,氨苄青霉素抗性基因和四环素抗性基因。有利地,新霉素抗性基因也可用于在真核宿主细胞中选择。

[0221] 其他选择性基因可用于扩增待表达的基因。扩增是在重组细胞的后继世代的染色体中串联重复基因的过程,其中所述基因是产生生长或细胞存活至关重要的蛋白质所需的基因。适用于哺乳动物细胞的选择性标记的实例包括二氢叶酸还原酶(DHFR)和无启动子胸腺嘧啶核苷激酶基因。将哺乳动物细胞转化体置于选择压力下,其中只有转化体由于存在于载体中的选择基因而独特地适应生存。选择压力是通过将转化的细胞培养在以下条件下来施予的,其中在该条件下,培养基中选择剂浓度连续增加,由此导致选择基因和编码另一基因(如抗体轻或重链)的DNA的扩增。结果,从扩增的DNA合成增加量的多肽。

[0222] mRNA的翻译起始通常需要核糖体结合位点,其特征例如Kozak序列。该元件通常位于启动子的3'和待表达的多肽的编码序列的5'位置。在某些实施方案中,一个或多个编码区可以与内部核糖体结合位点(IRES)可操作地连接,从而允许从单个RNA转录物翻译两个开放阅读框。

[0223] 本发明的表达和克隆载体将通常含有被宿主生物体识别并可操作地连接至编码多肽的分子的启动子。启动子是位于结构基因起始密码子上游(即5')(通常在约100至1000bp内)的非转录序列,其控制结构基因的转录。启动子通常分为两类:诱导型启动子和组成型启动子。诱导型启动子响应于培养条件中的一些变化(例如存在或不存在营养物或温度变化),启动其控制下的DNA进行提高水平的转录。另一方面,组成型启动子均匀地转录其可操作连接的基因,即对基因表达很少或没有控制。许多可被各种潜在宿主细胞识别的启动子是众所周知的。通过限制性内切酶消化从源DNA移出启动子并将所需启动子序列插入载体中,可以将合适的启动子可操作地连接至编码例如重链或轻链的DNA。

[0224] 与酵母宿主一起使用的合适启动子在本领域中也是众所周知的。酵母增强子有利

地与酵母启动子一起使用。适用于哺乳动物宿主细胞的启动子是众所周知的,包括但不限于从病毒如多瘤病毒,禽痘病毒,腺病毒(如腺病毒2),牛乳头瘤病毒,禽肉瘤病毒,巨细胞病毒,逆转录病毒,乙型肝炎病毒,最优选猿猴病毒40(SV40)的基因组获得的那些启动子。其他合适的哺乳动物启动子包括异源哺乳动物启动子,例如热休克启动子和肌动蛋白启动子。

[0225] 可能感兴趣的另外的启动子包括但不限于:SV40早期启动子(Benoist和Chambon, 1981, *Nature* 290:304-310);CMV启动子(Thomsen等,1984, *Proc. Natl. Acad. U.S.A.* 81:659-663);包含在劳斯肉瘤病毒的3'长末端重复中的启动子(Yamamoto等,1980, *Cell* 22:787-797);疱疹胸苷激酶启动子(Wagner等,1981, *Proc. Natl. Acad. Sci. U.S.A.* 78:1444-1445);来自金属硫蛋白基因的启动子和调控序列(Prinster等,1982, *Nature* 296:39-42);或tac启动子(DeBoer等,1983, *Proc. Natl. Acad. Sci. U.S.A.* 80:21-25)。另外感兴趣的是下列动物转录控制区,其显示组织特异性并已用于转基因动物:在胰腺腺泡细胞中有活性的弹性蛋白酶I基因控制区(Swift等,1984, *Cell* 38:639-646;Omitz等,1986, *Cold Spring Harbor Symp. Quant. Biol.* 50:399-409;MacDonald,1987, *Hepatology* 7:425-515);在胰腺 β 细胞中有活性的胰岛素基因控制区(Hanahan,1985, *Nature* 315:115-122);在淋巴样细胞中有活性的免疫球蛋白基因控制区(Grosschedl等,1984, *Cell* 38:647-658;Adames等,1985, *Nature* 318:533-538;Alexander等,1987, *Mol. Cell. Biol.* 7:1436-1444);在睾丸、乳房、淋巴样和肥大细胞中有活性的小鼠乳房肿瘤病毒控制区(Leder等,1986, *Cell* 45:485-495);在肝中有活性的白蛋白基因控制区(Pinkert等,1987, *Genes and Devel.* 1:268-276);在肝中有活性的甲胎蛋白基因控制区(Krumlauf等,1985, *Mol. Cell. Biol.* 5:1639-1648;Hammer等,1987, *Science* 253:53-58);在肝中有活性的 α 1-抗胰蛋白酶基因控制区(Kelsey等,1987, *Genes and Devel.* 1:161-171);在骨髓细胞中有活性的 β -珠蛋白基因控制区(Mogham等,1985, *Nature* 315:338-340;Kollias等,1986, *Cell* 46:89-94);在脑中的少突胶质细胞中有活性的髓磷脂碱性蛋白质基因控制区(Readhead等,1987, *Cell* 48:703-712);在骨骼肌中有活性的肌球蛋白轻链-2基因控制区(Sani,1985, *Nature* 314:283-286);和在下丘脑中有活性的促性腺激素释放激素基因控制区(Mason等,1986, *Science* 234:1372-1378)。

[0226] 可以将增强子序列插入载体中以增加高等真核生物对编码本发明轻链或重链的DNA的转录。增强子是DNA的顺式作用元件,通常长约10-300bp,作用于启动子以增加转录。增强子是相对的定位和位置独立的,已经在转录单位的5'和3'位置发现。可从哺乳动物基因获得的几种增强子序列是已知的(例如,珠蛋白,弹性蛋白酶,白蛋白,甲胎蛋白和胰岛素)。然而,通常使用来自病毒的增强子。本领域已知的SV40增强子,巨细胞病毒早期启动子增强子,多瘤病毒增强子和腺病毒增强子是用于激活真核生物启动子的示例性增强元件。尽管增强子可以在载体中位于编码序列的5'或3',但其通常位于启动子5'的位点。编码适当的天然或异源信号序列(前导序列或信号肽)的序列可以掺入表达载体中以促进抗体的细胞外分泌。信号肽或前导序列的选择取决于待产生抗体的宿主细胞的类型,并且异源信号序列可取代天然信号序列。在哺乳动物宿主细胞中有功能的信号肽的实例包括以下:美国专利号4,965,195中描述的血细胞介素-7(IL-7)的信号序列;Cosman等,1984, *Nature* 312:768中描述的血细胞介素-2受体的信号序列;EP专利号0367566中描述的血细胞介素-4

受体的信号肽;美国专利4,968,607中描述的I型白细胞介素-1受体信号肽;EP专利号0 460 846中描述的II型白细胞介素-1受体信号肽。

[0227] 当载体整合到宿主细胞基因组中时,载体可以含有一种或多种促进表达的元件。实例包括EASE元件(Aldrich等,2003Biotechnol Prog.19:1433-38)和基质附着区(MAR)。MARs介导染色质的结构组织化并且可以使整合的载体隔绝“位置”效应。因此,当载体用于产生稳定的转染体时,MARs特别有用。许多天然和合成的含MAR的核酸在本领域中是已知的,参见美国专利号6,239,328;7326567;6177612;6388066;6245974;7259010;6037525;7422874;7129062。

[0228] 本发明提供的表达载体可以由起始载体如市售载体构建。这种载体可含有或可不含有所有所需的侧翼序列。当载体中不存在一个或多个本文所述的侧翼序列时,它们可以单独获得并连接到载体中。用于获得每个侧翼序列的方法是本领域技术人员熟知的。

[0229] 在构建载体并将编码含Fc蛋白序列的核酸分子插入载体的适当位点后,可将完成的载体插入合适的宿主细胞中用于扩增和/或多肽表达。

[0230] 制备包含编码含Fc蛋白的核酸的载体的方法在本领域是公知的。参见例如美国专利号7,923,221。

[0231] 包含含Fc蛋白和期望的编码和控制序列的合适载体的构建可以采用标准连接技术。将分离的质粒或DNA片段切割,剪裁并以所需的形式重新连接以形成所需的质粒。所采用的方法不依赖于DNA来源或预期的宿主。

[0232] 在一些实施方案中,制备含Fc蛋白的方法还包括确定用于导入宿主细胞的多核苷酸的最佳比例。在一些实施方案中,质谱法用于确定含Fc蛋白产量,并调节该比率以使含Fc蛋白产量最大化。在一些实施方案中,使用双抗原ELISA来确定含Fc蛋白产量,并调节比率以使含Fc蛋白产量最大化。

[0233] 用于培养宿主细胞的方法是本领域技术人员熟知的。参见例如Li等,MAbs,2,2010。用于产生本文实施方案的期望含Fc蛋白的宿主细胞可以在各种培养基中培养。诸如Ham's F10 (Sigma),极限必需培养基(MEM), (Sigma),RPMI-1640 (Sigma)和Dulbecco氏改良的Eagle氏培养基(DMEM), (Sigma)等市售培养基适于培养宿主细胞。此外,Ham等人,Meth Enz,58,1979;Barnes等人,Anal Biochem,102,1980;美国专利号4,767,704;4,657,866;4,927,762;4,560,655;或5,122,469;国际专利申请号W090/03430或W087/00195;或美国专利重发布号30,985中描述的任何培养基均可以使用作为宿主细胞的培养基。这些培养基中的任何一种可以根据需要补充激素和/或其他生长因子(如胰岛素,转铁蛋白或表皮生长因子),盐(如氯化钠,钙,镁和磷酸盐),缓冲液(如HEPES),核苷酸(如腺苷和胸腺嘧啶),抗生素(如GENTAMYCIN™药物),微量元素(定义为通常以微摩尔范围内的终浓度存在的无机化合物),和葡萄糖或等效能源。任何其他必需的补充物也可以以本领域技术人员已知的适当浓度包含。诸如温度、pH等的培养条件是之前与选择用于表达的宿主细胞一起使用的那些,并且对普通技术人员而言将是显而易见的。

[0234] 通常,大规模(例如商业规模)进行含Fc蛋白的生产。为了获得适合商业规模生产的宿主细胞群,本领域普通技术人员将认识到使用逐步方式扩增宿主细胞群的效用。例如,该过程涉及以较小规模生长期望的宿主细胞以允许宿主细胞群体增加,例如种子罐(seed train)。为了进一步增加宿主细胞群,方法一般涉及使用种子罐接种更大的培养罐(例如接

种罐)。通常,使用一系列尺寸递增的接种罐来扩大宿主细胞群,例如接种罐(inoculum train)。该过程将提供合适的宿主细胞群用于生产培养中的培养。在一些实施方案中,生产培养是1000L培养罐。

[0235] 在一些实施方案中,制备含Fc蛋白的方法包括使用分批补料方法培养宿主细胞。在一些实施方案中,制备含Fc蛋白的方法包括使用连续补料方法培养宿主细胞。在一些实施方案中,制备含Fc蛋白的方法包括使用包含分批补料法和连续补料法的补料方法培养宿主细胞。

[0236] 获得含Fc蛋白的方法在本领域中是公知的。参见例如Huse等,J Biochem Biophys Meth,51,2002。含Fc蛋白可以在细胞内产生或直接分泌到培养基中。如果含Fc蛋白在细胞内产生,作为第一步,例如通过离心或超滤,除去颗粒碎片(宿主细胞或裂解的片段)。如果含Fc蛋白分泌到培养基中,则通常首先使用市售蛋白质浓缩过滤器,例如Amicon或Millipore Pellicon超滤单位,浓缩来自这种表达系统的上清液。在一些实施方案中,蛋白酶抑制剂可包括在任何前述步骤中以抑制蛋白水解,并可包含抗生素以防止外来污染物的生长。

[0237] 可以使用例如羟基磷灰石层析,凝胶电泳,透析和亲和层析来纯化由细胞制备的组合物,其中亲和层析是优选的纯化技术。蛋白A作为亲和配体的适合性取决于存在于含Fc蛋白中的任何免疫球蛋白Fc结构域的种类和同种型。蛋白A可以用于纯化基于含有1,2或4个重链的人免疫球蛋白的含Fc蛋白(参见,例如,Lindmark等,J Immunol Meth,62,1983)。蛋白G可以推荐用于所有小鼠同种型和人3(参见例如Guss等,EMBO,5,1986)。亲和配体所连接的基质通常是琼脂糖,但其他基质是可用的。机械稳定的基质,如可控多孔玻璃或聚(苯乙烯-二乙烯基)苯,比琼脂糖可以实现更快的流速和更短的处理时间。取决于待回收的含Fc蛋白,用于蛋白质纯化的其他技术例如在离子交换柱上分级分离,乙醇沉淀,反相HPLC,二氧化硅色谱法,在肝素SEPHAROSE™上的层析,在阴离子或阳离子交换树脂(例如聚天冬氨酸柱)上的层析,色谱聚焦,SDS-PAGE和硫酸铵沉淀,也是可用的。

[0238] 在一些实施方案中,纯化含Fc蛋白的方法包括使用过滤器。在一些实施例中,过滤器是渗滤系统。在一些实施方案中,过滤器是超滤系统。在一些实施方案中,过滤器是病毒过滤系统。在一些实施方案中,纯化含Fc蛋白包括使用一系列过滤步骤。在一些实施方案中,一系列过滤步骤选自以下中的至少一种:渗滤,超滤和病毒过滤。

[0239] 在一些实施方案中,纯化含Fc蛋白的方法包括使用选自过滤,蛋白A纯化,阳离子交换纯化,强阳离子交换纯化,阴离子交换纯化,反相纯化和多模式纯化的一系列蛋白纯化技术。

[0240] 本申请还提供了确定C-末端赖氨酸存在的方法。在一些实施方案中,确定C-末端赖氨酸存在的方法包括使用等电聚焦,例如成像的毛细管等电聚焦。在一些实施方案中,确定C-末端赖氨酸存在的方法包括使用质谱法。

[0241] 治疗方法

[0242] 本申请提供了治疗有需要的个体的疾病的方法,其包括向个体施用本文实施方案中描述的组合物。

[0243] 在一些实施方案中,将有效量的多个含Fc蛋白施用于有需要的个体,其中所述多个含Fc蛋白中基本上所有的含Fc蛋白在每个Fc结构域上具有C-末端赖氨酸。

[0244] 本文描述的药物组合物可以经由各种途径施用,例如胃肠外,包括静脉内,动脉内,腹膜内,肺内,口服,吸入,血管内,肌内,气管内,皮下,眼内,鞘内或透皮。在一些实施方案中,本文所述的药物组合物可以肠胃外施用。在一些实施方案中,本文所述的药物组合物可以静脉内施用。在一些实施方案中,本文所述的药物组合物可以皮下施用。在一些实施方案中,本文所述的药物组合物可以局部 (locally) 施用。在一些实施方案中,本文所述的药物组合物可以外用 (topically) 施用。

[0245] 可以通过本文方法治疗的疾病是可以用含Fc蛋白治疗的任何疾病。在一些实施方案中,该疾病是癌症。在一些实施方案中,所述疾病是自身免疫性疾病。在一些实施方案中,疾病是感染。

[0246] 在一些实施方案中,本文所述的药物组合物与另一种施用形式或治疗组合使用。

[0247] 本领域技术人员将认识到,在本发明的范围和精神内可有若干实施方式。现在将参考以下非限制性实施例,更详细地描述本发明。以下实施例进一步说明本发明,但是,当然不应该以任何方式解释为限制其范围。

实施例

[0248] 实施例1:

[0249] 羧肽酶D负责中国仓鼠卵巢 (CHO) 细胞中的C-末端赖氨酸切割

[0250] 方法:

[0251] 细胞系A是源自基于DUXB-11的DHFR缺陷型DP12宿主的、内部开发的抗体生产 (IgG1) 细胞系。参见Hu等人,Biotechnol Prog,29,2013;Urlaub等人,Proc Natl Acad Sci,77,1980。细胞系B是自CHOK1宿主衍生的内部开发的抗体生产 (IgG1) 细胞系。参见Hu等,Biotechnol Prog,29,2013。CHO细胞在125ml摇瓶容器中在专利DMEM/F12培养基中以150rpm,37°C和5%CO₂培养。每3至4天将细胞以3x10⁵/mL的接种密度传代。

[0252] 使用RNeasy 96试剂盒 (Cat#74181,Qiagen,Valencia,CA) 分离总RNA,并用DNase消化 (Cat#79254,RNase free DNase kit,Qiagen,Valencia,CA) 处理,以去除可能存在于分离的RNA样品中的任何残留DNA。根据制造商的说明书 (Cat#4309169,Applied Biosystems,Foster City,CA) 使用通用qRT-PCR主混合物,进行Taqman。将不同羧肽酶的表达水平相对于管家基因 β -2-微球蛋白 (β 2m) 进行归一化。

[0253] 用于Taqman分析的引物和探针序列如下:

[0254] CpD正向引物:CCC ACA CAT TAC AAA TCT TAC CA (SEQ ID NO.1);

[0255] CpD反向引物:GAG ATT TCG AGG GAC CAA AT (SEQ ID NO.2);

[0256] CpD探针:TTG GGA CAG AGT GCT GAG TAT CGT CA (SEQ ID NO.3);

[0257] CpN正向引物:GTG GGA TCA ATC ACG ATG TC (SEQ ID NO.4);

[0258] CpN反向引物:CCT TGG CAG TGA CAA TGT AAG TA (SEQ ID NO.5);

[0259] CpN探针:ACA TGG GGA TTA CTT CCG TCT GCT G (SEQ ID NO.6);

[0260] CpM正向引物:AAC TTG GAG AGT ACT ACC TGC TTC T (SEQ ID NO.7);

[0261] CpM反向引物:TCA TGC CCA GGG ACT GTA (SEQ ID NO.8);

[0262] CpM探针:ATT GAT CAC GTA GGA CCC TGG CAA A (SEQ ID N6.9);

[0263] CpB正向引物:TGA ATG CGC TGG TGA AAG G (SEQ ID NO.10);

- [0264] CpB反向引物:TCC TGG GCC ATA TGT GTA CTT G (SEQ ID NO.11);
- [0265] CpB探针:CGG TCA AGG AAC TTG CCT CTC TGC A (SEQ ID NO.12);
- [0266] CpE正向引物:TGG CTA CCT GGC AAT AAC AA (SEQ ID NO.13);
- [0267] CpE反向引物:CGA CTC CAG CTC AAA GTC AA (SEQ ID NO.14);
- [0268] CpE探针:AAG TGG CAG TTC CTT TCA AGC CTG C (SEQ ID NO.15);
- [0269] β -微球蛋白正向引物:TCC TCT CAG TGG TCT GCT TGG (SEQ ID NO.16);
- [0270] β -微球蛋白反向引物:TGG CGT GTG TAG ACT TGC ACT T (SEQ ID NO.17);
- [0271] β -微球蛋白探针:TGC CAT CCA GCG TCC CCC A (SEQ ID NO.18).
- [0272] 所有引物和探针在Genentech (Redwood City, CA) 合成和纯化。
- [0273] 进行Western印迹分析以确定CHO细胞中的相对羧肽酶D蛋白表达水平。使用胞质蛋白 β -肌动蛋白作为样品上样对照。收集细胞沉淀并在1X细胞裂解缓冲液 (Cat#9803, Cell Signaling Technology, Danvers, MA) 中裂解。通过Bradford测定 (Cat#1856210, Thermo Scientific, Rockford, IL) 对裂解的上清液的总蛋白进行定量, 并且在加载到还原性SDS-PAGE凝胶上之前在95°C加热3分钟。通过电印迹将分离的蛋白质转移到硝酸纤维素膜上。然后将转移的膜与抗羧肽酶D的兔多克隆抗体 (Cat#SAB2700486, Sigma, Saint Louis, MO) (以检测CpD蛋白表达) 和抗 β -肌动蛋白 (作为上样对照用于相同样品上样) 小鼠抗体 (目录号A2228, Sigma, Saint Louis, MO) 孵育。
- [0274] 小干扰RNA (DSIR) 设计程序 (<http://www.biomedcentral.com>) 用于设计CpD和CpN特异性小干扰RNA (siRNA) 序列:
- [0275] CpD siRNA靶序列:GGAAGAGAACTGCTACTAA (SEQ ID NO.19);
- [0276] CpN siRNA靶序列:GAATGGTGCTTGATGAGAA (SEQ ID NO.20)。
- [0277] 合成上述siRNA序列并分别克隆到pSilencer 3.1-H1载体 (Cat#AM5766, Ambion, Austin, TX) 中以制备siRNA表达构建体。根据制造商的推荐 (Invitrogen, Carlsbad, CA) 在瞬时转染中使用lipofectamine 2000 CD, 将每种构建体导入抗体表达细胞系A和抗体表达细胞系B。转染后24小时, 收集细胞沉淀以通过Taqman评估每种羧肽酶mRNA表达。另外, 收集细胞培养液 (CCF) 并通过Amicon Ultra-15 (30,000 MWCO, Millipore, Bedford, MA) 浓缩, 以用于通过成像毛细管等电聚焦 (IcIEF) 进行电荷变体产物质量分析。
- [0278] 通过使用具有氟碳涂覆的毛细管柱 (100 μ m $\times$ 5 cm) 的iCE280分析仪 (ProteinSimple) 来评估电荷变体分布。两性电解质溶液由0.35%甲基纤维素 (MC), 1.34% 3-10载体两性电解质, 1.34% 6.7-7.7载体两性电解质, pI标记物 (pI 6.61和pI 9.22) 和10mM L-精氨酸游离碱在纯化水中的混合物组成。阳极电解液为80mM磷酸, 阴极电解液为100mM氢氧化钠, 均在0.1%MC中。将样品稀释, 与两性电解质溶液混合, 然后通过引入1500V电势1分钟、接着引入3000V电势8分钟进行聚焦。通过使280nm紫外光通过毛细管并进入电荷耦合器件数码相机的镜头中, 获得聚焦的电荷变体的图像。为了除去重链C末端赖氨酸残基, 在稀释步骤中将羧肽酶B以1:100 (w/w) 的酶与底物比例加入到每个样品中, 然后在37°C温育20分钟。
- [0279] 以下的引导RNA (gRNA) 序列被设计为分别位于CpD外显子2和外显子21上:
- [0280] CpD引导RNA1序列 (gRNA 1) 用于外显子2:GAA GAC GAG ACT TTC AAA GAC GG (SEQ ID NO.21);

[0281] CpD引导RNA2序列(gRNA 2)用于外显子21:TCA GTT AGG TGG CTT AGG TTC GG (SEQ ID NO.22)。

[0282] 当与caspase 9 (Cas9) 一起共转染时,预期这些gRNA序列能够实现包含大部分注释的CpD基因座的**~46kb**缺失。

[0283] 在pLK0.5载体 (Cat#SHC-201, Sigma, St. Louis, MO) 的人U6启动子的控制下,克隆并转录各单个gRNA。两种gRNA和Cas9构建体以等摩尔比共转染。转染后72小时,通过有限稀释将细胞接种到384孔板中。三周后,挑取集落并转移至96孔板,通过聚合酶链式反应(PCR)筛选。用于筛选CpD敲除(**~46kb**缺失)和野生型等位基因的PCR引物如下:

[0284] 用于CpD敲除等位基因的正向引物:AGT TCA TTT ATG AAA GAT CCT GTG G (SEQ ID NO.23);

[0285] 用于CpD敲除等位基因的反向引物:GGA AAG GAG TCC TTC AGT GAA CAC (SEQ ID NO.24);

[0286] 用于CpD野生型等位基因的正向引物:CCA GTT CTG CTG TTA CAC TTT GAG (SEQ ID NO.25);

[0287] 用于CpD野生型等位基因的反向引物:AAT GTT TCC TCT TTC CTG GAC CTT (SEQ ID NO.26);

[0288] 将稀释的蛋白质样品(1mg/mL)与20mg/mL TCEP溶液以1:1的比例混合,并在60℃温育10分钟,以将抗体还原并解离成轻链和重链。然后使用Agilent 6210飞行时间质谱仪与纳米芯片LC-ESI源联用,通过液相色谱电喷雾电离质谱(LC-ESI-MS)分析每个样品。简言之,将约5ng蛋白质样品注射到定制芯片上,以40nL捕集柱(trap column)用于脱盐并且以等效分析柱尺寸43mm x75μm, Zorbax 300SB-C8 (5μm, Agilent Technologies, Santa Clara, CA) 用于400nL/min分离。参见Lu等人, *Mabs*, 5, 2013。流动相A是在水中的0.1%甲酸,流动相B是在乙腈中的0.1%甲酸。使用Agilent MassHunter定量分析工作站软件B.04.00,对MS数据进行提取并解卷积(deconvolute)。为了估计C-末端赖氨酸的百分比,使用对应于具有或不具有C-末端赖氨酸的抗体重链的解卷积质量的强度。

[0289] 补料分批生产在摇瓶容器中在专利化学成分确定的培养基中进行,第3天、第7天和第10天进行批式补料(bolus feeds)。在第3天进行37℃至35℃的温度转换。第14天使用蛋白A亲和层析和UV检测,测定滴度。

[0290] 结果:

[0291] 进行qPCR分析,以测量在基于DUXB-11的DHFR缺陷型DP12、DHFR阳性CHOK1宿主和来自各宿主的抗体生产细胞系(细胞系A和细胞系B)中可能的内源性羧肽酶的mRNA水平。细胞系A衍生自DHFR缺陷型DP12宿主,细胞系B衍生自DHFR阳性CHOK1宿主。

[0292] 选择5种羧肽酶,CpD,CpN,CpM,CpB和CpE,用于通过qPCR进行mRNA表达分析。CpD在两种CHO宿主和它们各自的抗体表达系中具有最高的mRNA表达水平(图1A和图1B)。在细胞系A和细胞系B中CpD和CpN mRNA水平高于CpM、CpE和CpB的mRNA水平(图1B)。

[0293] 使用具有两种最丰富的羧肽酶(即CpD和CpN)的两种抗体生产细胞系进行siRNA介导的敲低实验(参见例如Rao等人, *Adv Drug Deliv Rev*, 61, 2009)。用分别特异性靶向CpD (CpDi构建体)或CpN (CpNi构建体)的siRNA构建体,瞬时转染每种细胞系。乱码载体对照构建体也被转染作为阴性对照。如图2A所示,在细胞系A中CpD mRNA水平被敲低至39%,在细

胞系B中被敲低至28%。关于CpNi构建体抑制CpN mRNA,观察到了相似结果,其中CpN mRNA水平在细胞系A中降低至43%且在细胞系B中降低至35%(图2B)。

[0294] 计算细胞培养物中存在于抗体上的C-末端赖氨酸的百分比。为了计算细胞培养物中存在于抗体上的C-末端赖氨酸的百分比,将来自CpDi和CpNi转染的细胞系的上清液(用和不用CpB处理)通过蛋白质-A亲和层析纯化,并通过成像毛细管等电聚焦(IcIEF)分析。CpB处理和未处理的样品之间的差异值表示具有完整C末端赖氨酸的抗体的百分比。当CpD mRNA被敲低至约30-40%时,在两种细胞系中均观察到显著增加的C-末端赖氨酸水平(图2A和图3)。在用乱序RNAi构建体(阴性对照)或用将CpN mRNA敲低至约40%的CpNi构建体转染的两种细胞系中,均观察到最小C-末端赖氨酸水平变化(图2B和图3)。这些结果表明,CpD mRNA水平的降低导致了C末端赖氨酸水平增加,而CpN mRNA水平的变化在两种测试的抗体生产系中对C末端赖氨酸没有影响。

[0295] 使用CRISPR技术(参见例如He等人,Biotechnol Bioeng,2014;Jinek等人,Elife,2,2013)进行敲除实验。NCBI数据库中已知的CHOK1 CpD基因组序列约为55kb,包括CH0 CpD编码区序列的大部分。设计两个引导RNA序列以同时靶向中国仓鼠CpD基因的外显子2和21,目的是删除两个外显子之间的46kb序列,如图4A所示。当两种gRNA构建体和Cas9内切核酸酶构建体共转染到抗体表达细胞系B中时,Cas9内切核酸酶切割与gRNA互补的特定DNA序列。这种切割导致双链DNA断裂和随后的非同源末端连接(NHEJ) DNA修复,引起可以消除靶基因表达的缺失或插入。使用在gRNA序列附近和上游的PCR引物,如图4B所示,当两个gRNA之间的间隔序列被删除时,CpD敲除细胞应产生约600个碱基对的PCR产物。筛选总共180个克隆的CpD序列缺失,并基于PCR产物大小和序列分析,获得了CpD等位基因均缺失的两个克隆并用于CpD mRNA和蛋白质表达分析(数据未显示)。如图5A所示,通过qPCR分析,在两个CpD KO克隆中检测不到CpD mRNA。此外,使用抗CpD抗体的蛋白质印迹分析也显示,具有约180kDa分子量的蛋白质在两个CpD KO克隆中不存在,但存在于CpD WT克隆中,如图5B所示。

[0296] 评估CpD KO和WT克隆的生产力和产品质量属性。在两个KO和两个WT克隆(图6)中,通过补料分批摇瓶生产测定评估,在积分活细胞计数细胞(IVCC)、比生产力(Qp)和滴度中没有观察到明显差异,表明CH0细胞中CpD表达的缺乏不影响细胞生产力和生长。重要的是,C-末端赖氨酸水平通过还原抗体的基于质谱的分析来测量,显示在两个WT克隆中观察到了约4-6%的C-末端赖氨酸,而在两个KO克隆中存在100%的C-末端赖氨酸(图7)。因此,证明CpD是唯一负责去除CH0细胞中抗体重链的C-末端赖氨酸的羧肽酶。IcIEF测定也观察到了相同的结果(数据未显示)。CpD KO和WT克隆之间的其他产品质量属性和细胞培养物代谢物谱是相似的(数据未显示)。

序列表

	<110> 豪夫迈·罗氏有限公司 (F. HOFFMANN-LA ROCHE AG)	
	<120> 含 Fc 蛋白的表达	
	<130> 146392032640	
	<140>	
	<141>	
	<150> 62/222, 187	
	<151> 2015-09-22	
	<160> 26	
	<170> FastSEQ for Windows 4.0 版本	
	<210> 1	
	<211> 23	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 1	
	cccacacatt acaaattctta cca	23
	<210> 2	
	<211> 20	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 2	
[0001]	gagatttcga gggaccaa	20
	<210> 3	
	<211> 26	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 3	
	ttgggacaga gtgctgagta tcgtca	26
	<210> 4	
	<211> 20	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 4	
	gtgggatcaa tcacgatgtc	20
	<210> 5	
	<211> 23	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 5	
	ccttggcagt gacaatgtaa gta	23
	<210> 6	

	<211> 25	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 6	
	acatggggat tacttccgtc tgctg	25
	<210> 7	
	<211> 25	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 7	
	aacttggaga gtactacctg ctctt	25
	<210> 8	
	<211> 18	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 8	
	tcatgcccag ggactgta	18
	<210> 9	
	<211> 25	
[0002]	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 9	
	attgatcacg taggaccctg gcaaa	25
	<210> 10	
	<211> 19	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 10	
	tgaatgcgct ggtgaaagg	19
	<210> 11	
	<211> 22	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 11	
	tcctgggcca tatgtgtact tg	22
	<210> 12	
	<211> 25	
	<212> DNA	
	<213> 人工序列	

	<220>	
	<223> 合成构建体	
	<400> 12	
	cgggtcaagga acttgccctct ctgca	25
	<210> 13	
	<211> 20	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 13	
	tggtctacctg gcaataacaa	20
	<210> 14	
	<211> 20	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 14	
	cgactccagc tcaaagtcaa	20
	<210> 15	
	<211> 25	
	<212> DNA	
	<213> 人工序列	
	<220>	
[0003]	<223> 合成构建体	
	<400> 15	
	aagtggcagt tcctttcaag cctgc	25
	<210> 16	
	<211> 21	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 16	
	tcctctcagt ggtctgcttg g	21
	<210> 17	
	<211> 22	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 17	
	tggcgtgtgt agacttgac tt	22
	<210> 18	
	<211> 19	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 18	

	tgccatccag cgtcccca	19
	<210> 19	
	<211> 19	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 19	
	ggaagagAAC tgctactaa	19
	<210> 20	
	<211> 19	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 20	
	gaatggtgct tgatgagaa	19
	<210> 21	
	<211> 23	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 21	
	gaagacgaga cttcaaaaga cgg	23
[0004]	<210> 22	
	<211> 23	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 22	
	tcagttaggt ggcttaggtt cgg	23
	<210> 23	
	<211> 25	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 23	
	agttcattta tgaagatcc tglgg	25
	<210> 24	
	<211> 24	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 24	
	ggaaaggagt ccttcagtga acac	24
	<210> 25	
	<211> 24	

	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 25	
	ccagttctgc tggtacactt tgag	24
[0005]	<210> 26	
	<211> 24	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 合成构建体	
	<400> 26	
	aatgtttcct ctttcctgga cctt	24

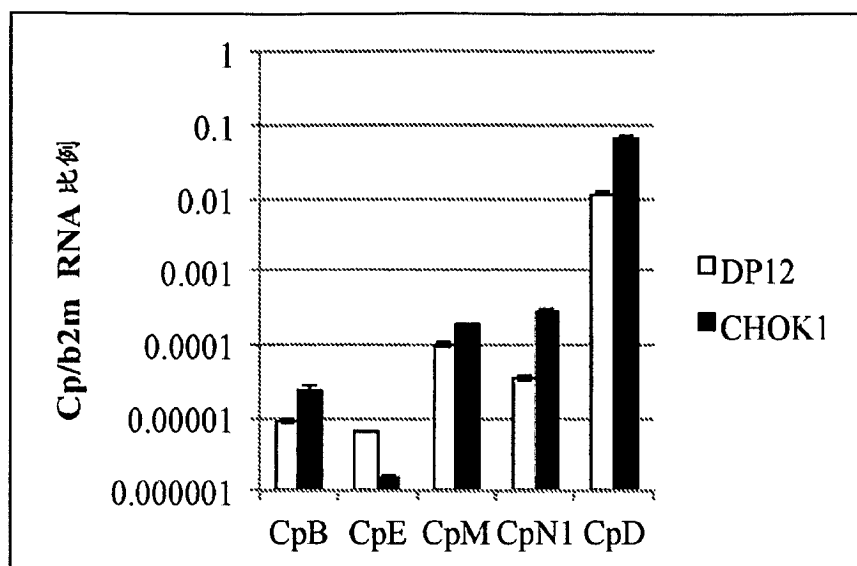


图1A

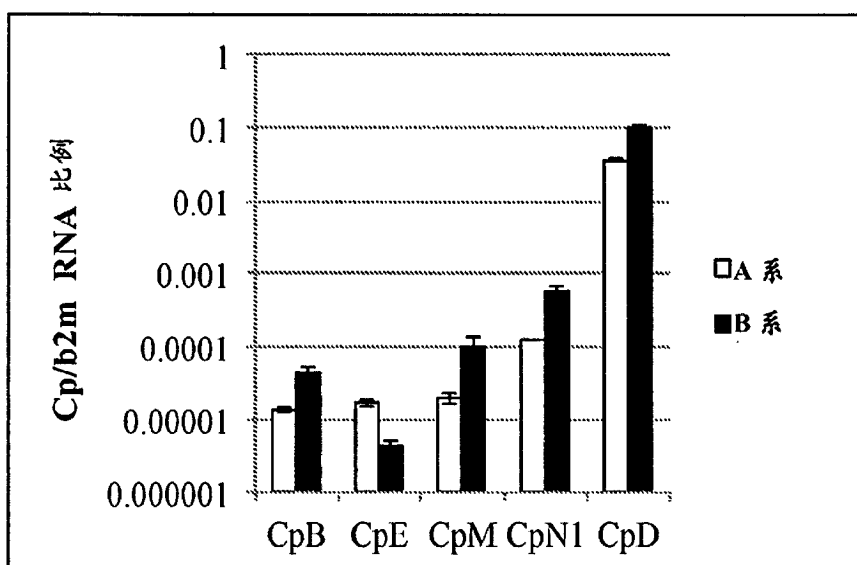


图1B

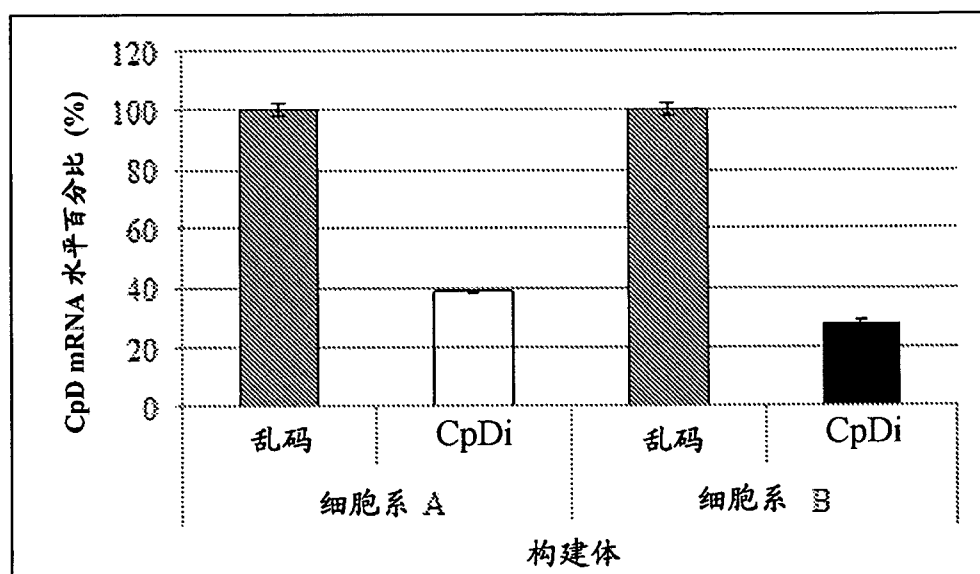


图2A

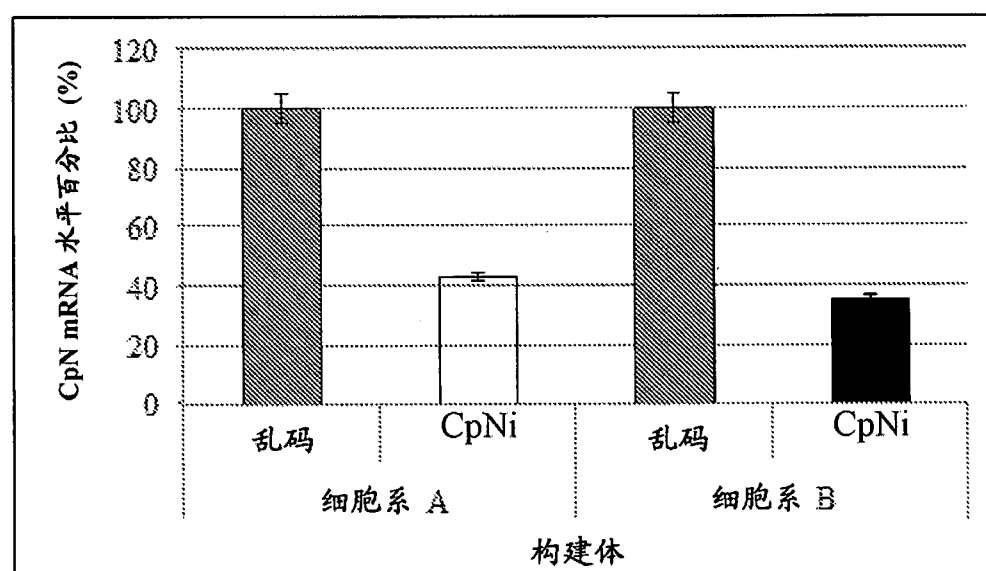


图2B

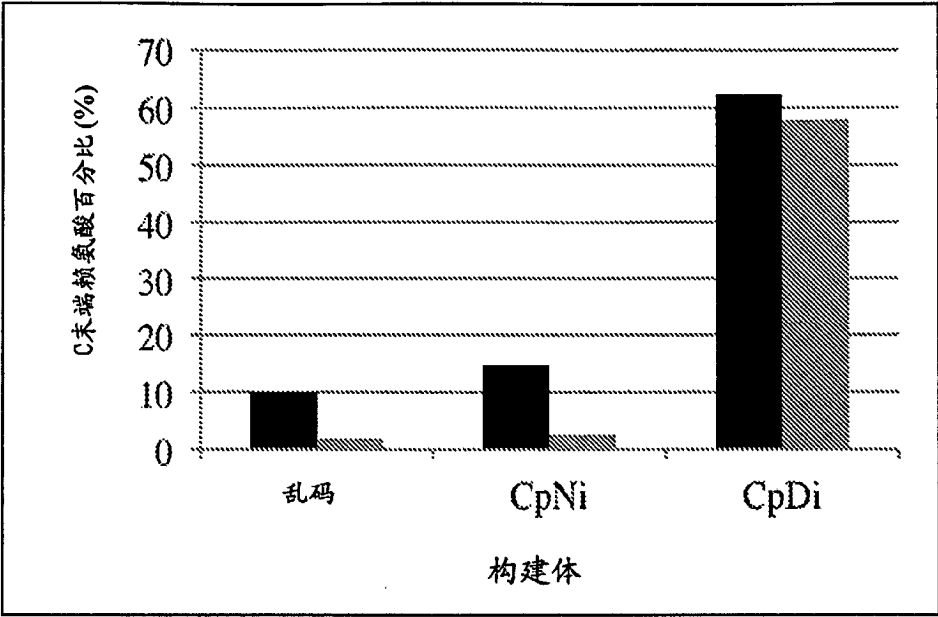


图3

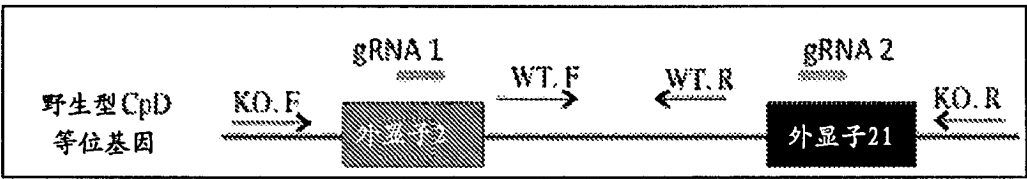


图4A

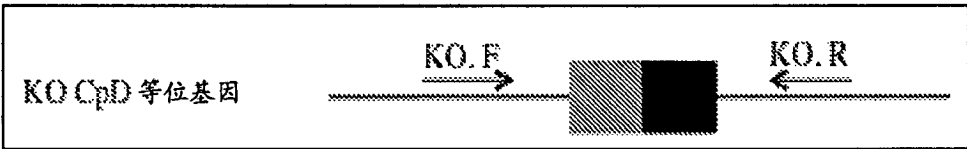


图4B

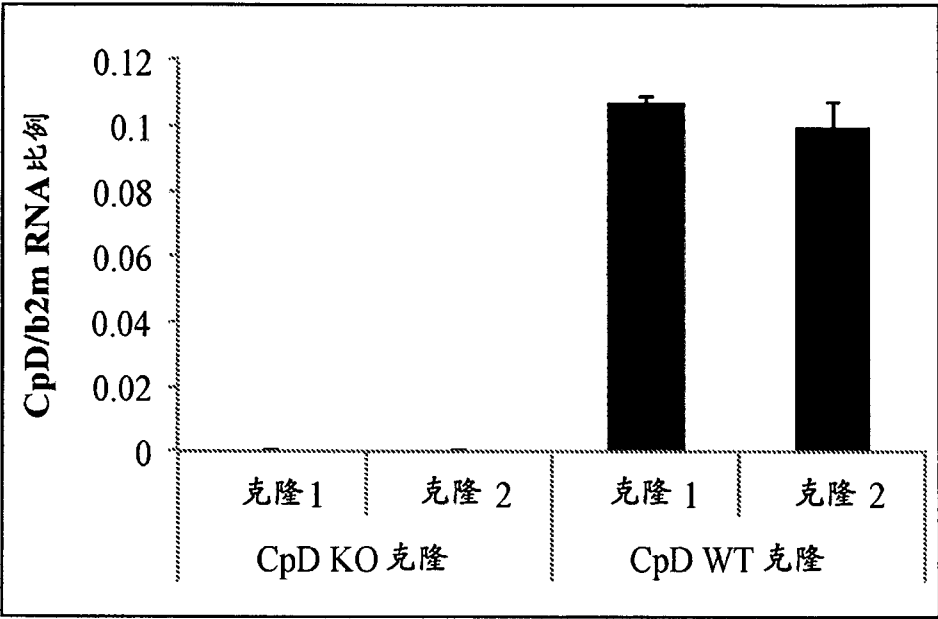


图5A

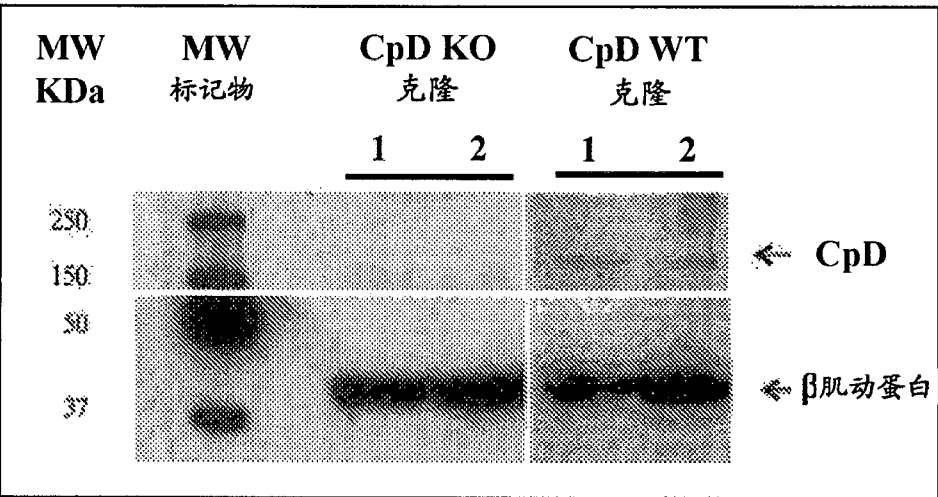


图5B

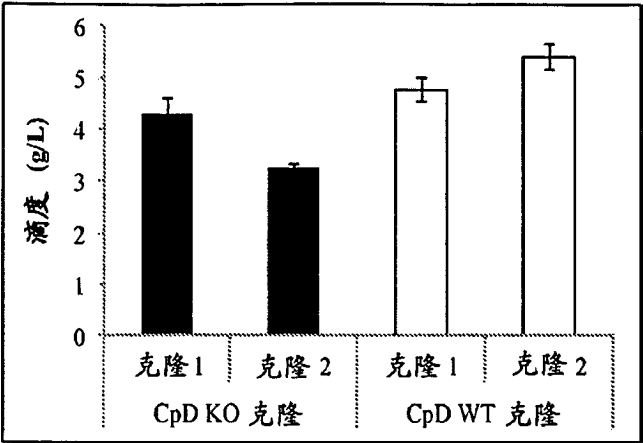


图6A

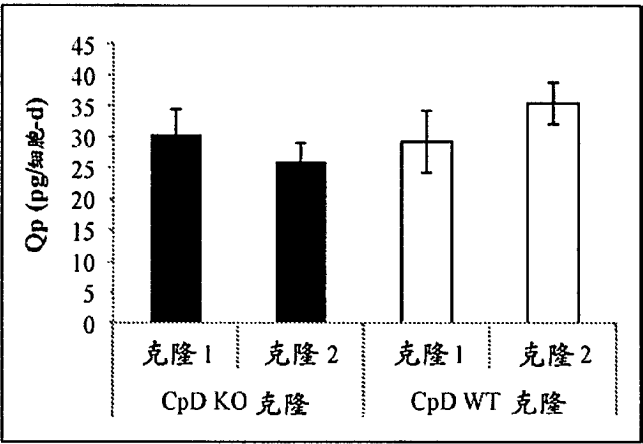


图6B

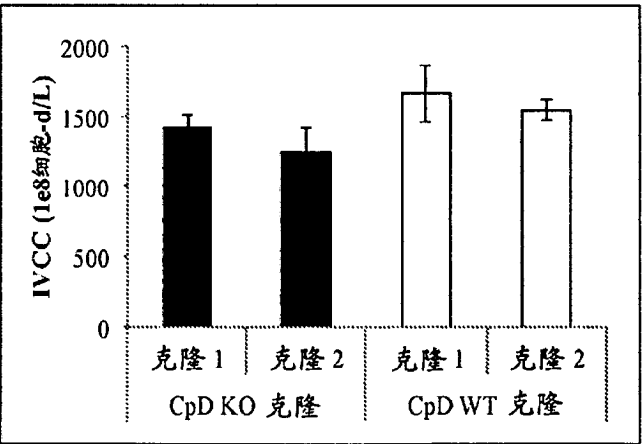


图6C

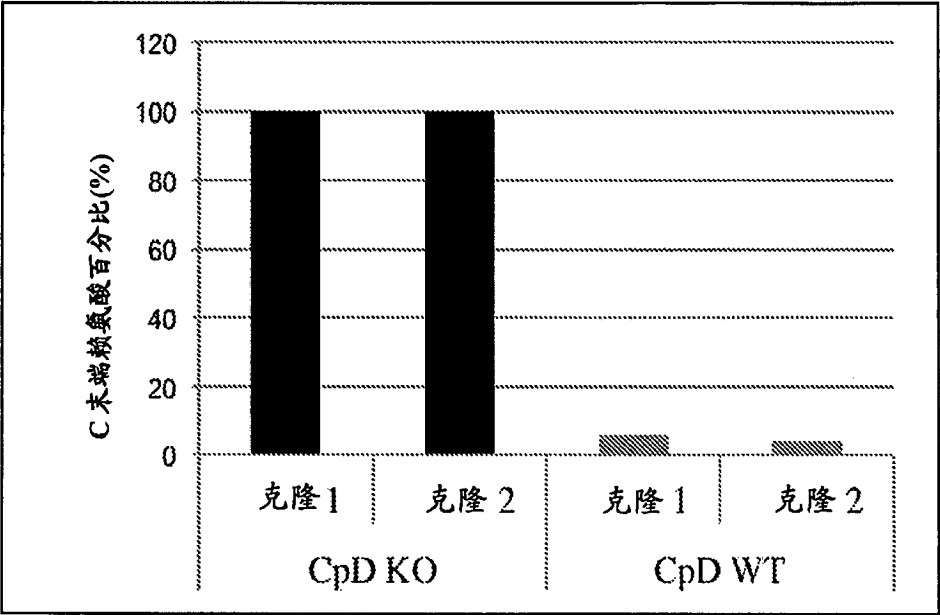


图7

Abstract

The present invention provides compositions comprising a Fc-containing protein wherein substantially all the Fc domains have a C-terminal lysine. Further provided are host cell for producing said compositions, methods of making said host cells and compositions, and method of use thereof.