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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 January 2017 (19.01.2017)

WIPO | PCT

(10) International Publication Number
WO 2017/011773 A2

(51) International Patent Classification:
A61K 48/00 (2006.01) *C07H 21/04* (2006.01)

(21) International Application Number:
PCT/US2016/042568

(22) International Filing Date:
15 July 2016 (15.07.2016)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/193,018 15 July 2015 (15.07.2015) US

(71) Applicant: **MODERNATX, INC.** [US/US]; 200 Technology Square, Cambridge, Massachusetts 02139 (US).

(72) Inventors: **BOLEN, Joseph Beene**; 303 3rd Street #724, Cambridge, Massachusetts 02142 (US). **HOGE, Stephen**; 66 Summit Avenue, Brookline, Massachusetts 02446 (US). **MCFADYEN, Iain**; 11 Amherst Street, Arlington, Massachusetts 02474 (US).

(74) Agents: **KIM, Ji-Eun** et al.; Sterne, Kessler, Goldstein & Fox P.L.L.C., 1100 New York Avenue N.W., Washington, District of Columbia 20005 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, 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, 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:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

[Continued on next page]

(54) Title: CODON-OPTIMIZED NUCLEIC ACIDS ENCODING ANTIBODIES

(57) Abstract: The present disclosure discloses compositions comprising codon-optimized nucleotide sequences, in particular mRNAs, that encode antibodies and functional fragments thereof (e.g., antigen-binding fragments or Fc fragments that can be used in fusion proteins). These optimized nucleic acid sequences can be used to expressing therapeutic antibodies in vivo. Also provided are methods of manufacturing the disclosed codon-optimized nucleotide sequences, methods of generating therapeutic antibodies in a subject in need thereof by administering a polynucleotide comprising a codon-optimized nucleotide sequence, and methods for treating and/preventing a disease or condition in a subject using the disclosed compositions and methods.

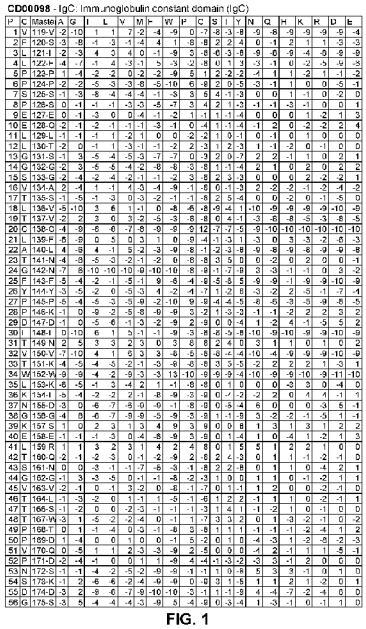


FIG. 1

WO 2017/011773 A2

— *with sequence listing part of description (Rule 5.2(a))*

CODON-OPTIMIZED NUCLEIC ACIDS ENCODING ANTIBODIES

REFERENCE TO SEQUENCE LISTING SUBMITTED ELECTRONICALLY

[0001] The content of the electronically submitted sequence listing in ASCII text file (Name 3529_023PC02_ST25_SQLST.txt; Size: 3,586,874 bytes; and Date of Creation: June 29, 2016) filed with the application is incorporated herein by reference in its entirety.

BACKGROUND

[0002] Therapeutic antibodies represent one of the most important medical therapeutic advances of the last 25 years. Immune based therapies are used routinely against a host of autoimmune diseases, treatment of cancer as well as infectious diseases. However, issues remain that limit the use and dissemination of this therapeutic approach. The high cost of production of these complex biologics can limit their use in the broader population, particularly in the developing world where they could have a great impact. Furthermore, the frequent requirement for repeat administrations of the therapeutic antibodies to attain and maintain efficacy can be an impediment in terms of logistics and patient compliance. Additionally, the long-term stability of these antibody formulations is frequently short and less than optimal. The administration of a synthetic antibody in nucleic acid form that could be delivered to a subject in a safe and cost effective manner could bypass current bottlenecks in antibody therapy, namely the high costs associated with protein production, purification, formulation, and storage/distribution. However, immunoglobulin molecules are very complex genetic systems that are difficult to express and assemble in vivo. Furthermore, there are multiple problems affecting DNA-based in vivo protein expression. For example, introduced DNA can integrate into host cell genomic DNA at some frequency, resulting in alterations and/or damage to the host cell genomic DNA. Alternatively, the heterologous deoxyribonucleic acid (DNA) introduced into a cell can be inherited by daughter cells (whether or not the heterologous DNA has integrated into the chromosome) or by offspring. In addition, assuming proper delivery and no damage or integration into the host genome, there are multiple steps which must occur before the encoded protein is made. Once inside the cell, DNA must be transported into the nucleus where it is transcribed into RNA. The RNA transcribed from DNA must then enter the

cytoplasm where it is translated into protein. Not only do the multiple processing steps from administered DNA to protein create lag times before the generation of the functional protein, each step represents an opportunity for error and damage to the cell. Further, it is known to be difficult to obtain DNA expression in cells as DNA frequently enters a cell but is not expressed or not expressed at reasonable rates or concentrations. This can be a particular problem when DNA is introduced into primary cells or modified cell lines.

[0003] These shortcomings can be avoided by the use of mRNA to deliver a polypeptide of interest, and in particular, by introducing modifications to mRNA sequences that have potential as therapeutics conferring benefits beyond evading, avoiding or diminishing the immune response. See, e.g., U.S. Pat. No. US8999380, U.S. Publ. Nos. US20120065252, US20130102034, US20120237975, which are herein incorporated by reference in their entireties.

[0004] Accordingly, there is a need for optimized nucleic acid molecules, in particular mRNAs or DNAs encoding such mRNAs, that can be administered to subjects and result in effective expression of antibody-based biologics *in vivo* (e.g., full antibodies; constructs comprising one or more antibody components, such as scFv's; or antibody fusion proteins, for example, Fc fusion proteins).

BRIEF SUMMARY

[0005] The present disclosure provides optimized nucleotide sequences (e.g., mRNA sequences) encoding antibodies and functional fragments thereof (e.g., antigen binding fragments or Fc fragments) which can be expressed *in vivo* in a subject in need thereof. Accordingly, in one aspect, the present disclosure provides a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP 9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes TVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQES VTEQDSKDSTYSLSX[1]TLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNR (SEQ ID NO: 2200), wherein X^1 is selected from N and S. In some aspects, the nucleotide sequence encodes SEQ ID NO:2189. Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1,

MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes PKAAPSVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTT PSKQSNNKYAASSYLSLTPEQWKSHX^{2}SYSCQVTHEGSTVEKTVAPX^{3}ECS (SEQ ID NO: 2201),

wherein X^2 is selected from R and K, and X^3 is selected from T and A. In some aspects, the nucleotide sequence encodes SEQ ID NO: 2190. In some aspects, the nucleotide sequence encodes a lambda light chain constant domain of an antibody or a fragment thereof.

[0006] The disclosure also provides a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes SX^{4}GPSVX^{5}PLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGVHTFPAVLQSSG LYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKX^{6}X^7 (SEQ ID NO: 2202) wherein X^4 is an optional ASTK sequence, X^5 is selected from F and L, X^6 is selected from K and R, and X^7 is selected from V and A. In some aspects, the nucleotide sequence encodes SEQ ID NO: 2191. In some aspects, the nucleotide sequence encodes a CH1 domain of an IgG1 antibody or a fragment thereof.

[0007] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes APEX^{8}X^{9}GX^{10}PSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNX^{11}YVDGV EVHNAKTKPREEQYX^{12}STYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK TISKAK (SEQ ID NO: 2203) wherein X^8 is selected from L and A, X^9 is selected from L and A, X^{10} is selected from G and A, and X^{11} is selected from V and W, and X^{12} is selected from N and A. In some aspects, the nucleotide sequence encodes SEQ ID NO: 2192. In some aspects, the nucleotide sequence encodes a CH2 domain of an IgG1 antibody or a fragment thereof.

[0008] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5,

MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes GQPREPQVYTLPPSRX^{13}EX^{14}TKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYK TTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG (SEQ ID NO: 2204) wherein X^{13} is selected from E and D, and X^{14} is selected from M and L. In some aspects, the nucleotide sequence encodes SEQ ID NO: 2193. In some aspects, the nucleotide sequence encodes a CH3 domain of an IgG1 antibody or a fragment thereof.

[0009] The disclosure also provides a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes SASTKGPSVFPLAPCSRSTSESTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPA VLQSSGLYSLSSVVTVX^{15}SSNFGTQTYTCNVDHKPSNTKVDKTV (SEQ ID NO: 2205) wherein X^{15} is selected from P and T. In some aspects, the nucleotide sequence encodes SEQ ID NO: 2194. In some aspects, the nucleotide sequence encodes a CH1 domain of an IgG2 antibody or a fragment thereof.

[0010] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes APPVAGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFNWYVDGX^{16}EV HNAKTKPREEQFNSTFRVVSVLTVVHQDWLNGKEYKCKVSNKGLPX^{17}X^{18}IEKTI SKTK (SEQ ID NO: 2206) wherein X^{16} is selected from V and M, X^{17} is selected from A and S; and X^{18} is selected from P and S. In some aspects, the nucleotide sequence encodes SEQ ID NO: 2195. In some aspects, the nucleotide sequence encodes a CH2 domain of an IgG2 antibody or a fragment thereof.

[0011] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes GQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTP PMLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG

(SEQ ID NO: 2196). In some aspects, the nucleotide sequence encodes a CH3 domain of an IgG2 antibody or a fragment thereof.

[0012] The disclosure also provides a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes SASTKGPSVFPLAPCSRSTSESTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPA VLQSSGLYSLSSVVTVPSSSLGTKTYTCNVDHKPSNTKVDKRV (SEQ ID NO: 2197). In some aspects, the nucleotide sequence encodes a CH1 domain of an IgG4 antibody or a fragment thereof.

[0013] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes APEFLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSQEDPEVQFNWYVDGVEVH NAKTKPREEQFNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKGLPSSIEKTISKA K (SEQ ID NO: 2198). In some aspects, the nucleotide sequence encodes a CH2 domain of an IgG4 antibody or a fragment thereof.

[0014] The disclosure also provides a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes GQPREPQVYTLPPSQEEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTP PVLDSDGSFFLYSRLTVDKSRWQEGNVFSCSVMHEALHNHYTQKSLSLSLG (SEQ ID NO: 2199). In some aspects, the nucleotide sequence encodes a CH3 domain of an IgG4 antibody or a fragment thereof.

[0015] The disclosure also provides a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $X^1X^2X^3LTQX^4X^5X^6VSX^7X^8X^9GX^{10}X^{11}X^{12}X^{13}X^{14}X^{15}C$ (SEQ ID NO: 2235) wherein X^1 is selected from Q, D, E and S; X^2 is selected from S, I, A, and Y; X^3 is selected from V, Q, A, and E; X^4 is selected from P and D; X^5 is selected from P, N, and A; X^6 is selected

from S and A; X^7 is selected from G, T, A, and V; X^8 is selected from A and S; X^9 is selected from P and L; X^{10} is selected from Q, K, and S; X^{11} is selected from R, K, T, and S; X^{12} is selected from V, I, and A; X^{13} is selected from T, K, and R; X^{14} is selected from I and L; and, X^{15} is selected from S at T. In some aspects, the nucleotide sequence encodes a sequence identical to QSVLTQPPSVSGAPGQRVTISC (SEQ ID NO: 2207) except for at least one substitution selected from Q1(DES), S2(IAY), V3(QAE), P7D, P8(NA), S9A, G12(TAV), A13S, P14L, Q16(KS), R17(KTS), V18(IA), T19(KR), I20L, and S21T. In some aspects, the nucleotide sequence encodes QSVLTQPPSVSGAPGQRVTISC (SEQ ID NO: 2207). In some aspects, the nucleotide sequence encodes the first framework region (FW1) of a lambda light chain variable domain.

[0016] Also provided in the disclosure is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $WYQX^1X^2X^3GX^4X^5PX^6X^7X^8I$ (SEQ ID NO: 2236) wherein X^1 is selected from Q and L; X^2 is selected from L,Y, H, and K; X^3 is selected from P and E; X^4 is selected from T, R, K, and Q; X^5 is selected from A and S; X^6 is selected from K, T, V and I; X^7 is selected from L and T; and X^8 is selected from L, M, and V. In some aspects, the nucleotide sequence encodes a sequence identical to WYQQLPGTAPKLLI (SEQ ID NO: 2208) except for at least one substitution selected from Q4L, L5(YHK), P6E, T8(RKQ), A9S, K11(TVI), L12T, and L13(MV). In some aspects, the nucleotide sequence encodes WYQQLPGTAPKLL (SEQ ID NO: 2208). In some aspects, the nucleotide sequence encodes the second framework region (FW2) of a lambda light chain variable domain.

[0017] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $RFSGSX^1SX^2X^3X^4AX^5LX^6IX^7X^8X^9X^{10}X^{11}X^{12}DEAX^{13}YX^{14}C$ (SEQ ID NO: 2237) wherein X^1 is selected from K, N, S, and I; X^2 is selected from G and S; X^3 is selected from T and N; X^4 is selected from S and T; X^5 is selected from S, T, and F; X^6 is selected from A, T, and G; X^7 is selected from T, H, and S; X^8 is selected from G, N, and R; X^9 is selected from L, V, and A; X^{10} is selected from Q, E, and A; X^{11} is selected from A, T,

and I; X^{12} is selected from E and G; X^{13} is selected from D and I; and, X^{14} is selected from Y and F. In some aspects, the nucleotide sequence encodes a sequence identical to RFSGSKSGTSASLAITGLQAEDEADYYC (SEQ ID NO: 2209) except for at least one substitution selected from K6(NSI), G8S, T9N, S10T, S12(TF), A14(TG), T16(HS), G17(NR), L18(VA), Q19(EA), A20(TI), E21G, D25I, and Y27F. In some aspects, the nucleotide sequence encodes RFSGSKSGTSASLAITGLQAEDEADYYC (SEQ ID NO: 2209). In some aspects, the nucleotide sequence encodes the third framework region (FW3) of a lambda light chain variable domain.

[0018] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes FGX^1GTX^2X^3TVL (SEQ ID NO:2238) wherein X^1 is selected from G and T; X^2 is selected from K and Q; and X^3 is selected from L and V. In some aspects, the nucleotide sequence encodes a sequence identical to FGGGTKLTVL (SEQ ID NO: 2210) except for at least one substitution selected from G3T, K6Q, and L7V. In some aspects, the nucleotide sequence encodes FGGGTKLTVL (SEQ ID NO: 2210). In some aspects, the nucleotide sequence encodes the fourth framework region (FW4) of a lambda light chain variable domain.

[0019] The disclosure also provides a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $X^1X^2QX^3TQX^4X^5SX^6X^7SASX^8CDRVTX^9X^{10}C$ (SEQ ID NO: 2239) wherein X^1 is selected from D and A; X^2 is selected from I and V; X^3 is selected from M, L, and V; X^4 is selected from S and F; X^5 is selected from P and T; X^6 is selected from S and T; X^7 is selected from L and V; X^8 is selected from V, I, and A; X^9 is selected from I and M; and, X^{10} is selected from T and S. In some aspects, the nucleotide sequence encodes a sequence identical to DIQMTQSPSSLSASVCDRVTITC (SEQ ID NO: 2211) except for at least one substitution selected from D1A, I2V, M4(LV), S7F, P8T, S10T, L11V, V15(IA), I21M, and T22S. In some aspects, the nucleotide sequence encodes DIQMTQSPSSLSASVCDRVTITC (SEQ ID NO: 2211).

[0020] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $DX^1X^2X^3TQX^4PX^5SX^6X^7X^8X^9X^{10}GX^{11}X^{12}X^{13}X^{14}X^{15}X^{16}C$ (SEQ ID NO: 2243) wherein X^1 is selected from I and V; X^2 is selected from V, L, and Q; X^3 is selected from M and L; X^4 is selected from S and T; X^5 is selected from L and D; X^6 is selected from L and V; X^7 is selected from P, S and A; X^8 is selected from V and M; X^9 is selected from T and S; X^{10} is selected from P and L; X^{11} is selected from E and Q; X^{12} is selected from P and R; X^{13} is selected from A and V; X^{14} is selected from S and T; X^{15} is selected from I, M, and L; and X^{16} is selected from S and N. In some aspects, the nucleotide sequence encodes a sequence identical to DIVMTQSPLSLPVTPGEPASISC (SEQ ID NO: 2215) except for at least one substitution selected from I2V, V3(LQ), M4L, S7T, L9D, L11V, P12(SA), V13M, T14S, P15L, E17Q, P18R, A19V, S20T, I21(ML), and S22N. In some aspects, the nucleotide sequence encodes DIVMTQSPLSLPVTPGEPASISC (SEQ ID NO: 2215).

[0021] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $X^1X^2VX^3TQSPX^4TLSX^5SPGERATLSC$ (SEQ ID NO: 2247) wherein X^1 is selected from E and D; X^2 is selected from I and T; X^3 is selected from L and M; X^4 is selected from G and A; and, X^5 is selected from L and V. In some aspects, the nucleotide sequence encodes a sequence identical to EIVLTQSPGTLSLSPGERATLSC (SEQ ID NO: 2219) except for at least one substitution selected from E1D, I2T, L4M, G9A, and L13V. In some aspects, the nucleotide sequence encodes EIVLTQSPGTLSLSPGERATLSC (SEQ ID NO: 2219). In some aspects, the nucleotide sequence encodes the first framework region (FW1) of a kappa light chain variable domain.

[0022] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $WX^1X^2X^3X^4PX^5KX^6X^7X^8X^9X^{10}IX^{11}$ (SEQ ID NO: 2240) wherein X^1 is selected from Y and F; X^2 is selected from Q and L; X^3 is selected from Q and H; X^4 is selected from K

and I; X^5 is selected from G and E; X^6 is selected from A and V; X^7 is selected from P and V; X^8 is selected from K and Q; X^9 is selected from L, T, S, R, P, and V; X^{10} is selected from L and W; and, X^{11} is selected from Y and S. In some aspects, the nucleotide sequence encodes a sequence identical to WYQQKPGKAPKLLIY (SEQ ID NO: 2212) except for at least one substitution selected from Y2F, Q3L, Q4H, K5I, G7E, A9V, P10V, K11Q, L12(TSRPV), L13W, and Y15S. In some aspects, the nucleotide sequence encodes WYQQKPGKAPKLLIY (SEQ ID NO: 2212).

[0023] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $WX^1X^2QX^3X^4GQX^5PX^6X^7LIX^8$ (SEQ ID NO: 2244) wherein X^1 is selected from Y, F, and W; X^2 is selected from L and Q; X^3 is selected from K and R; X^4 is selected from P and S; X^5 is selected from S and P; X^6 is selected from Q, K, R, and N; X^7 is selected from L and R; and, X^8 is selected from Y and W. In some aspects, the nucleotide sequence encodes a sequence identical to WYLQKPGQSPQLLIY (SEQ ID NO: 2216) except for at least one substitution selected from Y2(FW), L3Q, K5R, P6S, S9P,Q11(KRN), L12R, and Y15W. In some aspects, the nucleotide sequence encodes WYLQKPGQSPQLLIY (SEQ ID NO: 2216).

[0024] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $WX^1X^2QX^3PGQAPRX^4LIX^5$ (SEQ ID NO: 2248) wherein X^1 is selected from Y and F; X^2 is selected from Q and R; X^3 is selected from K and R; X^4 is selected from L and P; and X^5 is selected from Y, R, and K. In some aspects, the nucleotide sequence encodes a sequence identical to WYQQKPGQAPRLLIY (SEQ ID NO: 2220) except for at least one substitution selected from Y2F, Q3R, K5R, L12P, and Y15(RK). In some aspects, the nucleotide sequence encodes WYQQKPGQAPRLLIY (SEQ ID NO: 2220). In some aspects, the nucleotide sequence encodes the second framework region (FW2) of a kappa light chain variable domain.

[0025] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5,

MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes RFSGSX^{1}SGX^{2}X^{3}X^{4}X^{5}X^{6}TISSLX^{7}X^{8}X^{9}DX^{10}AX^{11}YX^{12}C (SEQ ID NO: 2241) wherein X^1 is selected from G and R; X^2 is selected from T and Q; X^3 is selected from D, E, and Y; X^4 is selected from F and Y; X^5 is selected from T and S; X^6 is selected from L and F; X^7 is selected from Q and E; X^8 is selected from P, Q, A, and S; X^9 is selected from E and D; X^{10} is selected from F, I, S, L, V, and T; X^{11} is selected from T, S, and V; and, X^{12} is selected from Y and F. In some aspects, the nucleotide sequence encodes a sequence identical to RFSGSGSGTDFTLTISSLQPEDFATYYC (SEQ ID NO: 2213) except for at least one substitution selected from G6R, T9Q, D10(EY), F11Y, T12S, L13F, Q19E, P20(QAS), E21D, F23(ISLVT), T25(SV), and Y27F. In some aspect, the nucleotide sequence encodes RFSGSGSGTDFTLTISSLQPEDFATYYC (SEQ ID NO: 2213).

[0026] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes RFSGSGSX^{1}TX^{2}FTLX^{3}ISX^{4}X^{5}X^{6}AX^{7}DVX^{8}X^{9}X^{10}X^{11}C (SEQ ID NO: 2245) wherein X^1 is selected from G and A; X^2 is selected from D and A; X^3 is selected from K, R, and T; X^4 is selected from R and S; X^5 is selected from V and L; X^6 is selected from E and Q; X^7 is selected from E and Q; X^8 is selected from G and A; X^9 is selected from V, D, and F; X^{10} is selected from Y and W; and, X^{11} is selected from Y, F, and W. In one aspect, the nucleotide sequence encodes a sequence identical to RFSGSGSGTDFTLKISRVEAEDVGVYYC (SEQ ID NO: 2217) except for at least one substitution selected from G8A, D10A, K14(RT), R17S, V18L, E19Q, E21Q, G24A, V25(DF), Y26W, and Y27(FW). In one aspect, the nucleotide sequence encodes RFSGSGSGTDFTLKISRVEAEDVGVYYC (SEQ ID NO: 2217).

[0027] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes RFSGSGSGTX^{1}X^{2}TLTISX^{3}LX^{4}X^{5}EDFAX^{6}X^{7}YC (SEQ ID NO: 2249) wherein X^1 is selected from D and E; X^2 is selected from F and S; X^3 is selected from R and S; X^4 is selected from E and Q; X^5 is selected from P and S; X^6 is selected from V and T; and, X^7

is selected from Y and F. In one aspect, the nucleotide sequence encodes a sequence identical to RFSGSGSGTDFTLTISRLEPEDFAVYYC (SEQ ID NO: 2221) except for at least one substitution selected from D10E, F11S, R17S, E19Q, P20S, V25T, and Y26F. In one aspect, the nucleotide sequence encodes RFSGSGSGTDFTLTISRLEPEDFAVYYC (SEQ ID NO: 2221). In one aspect, the nucleotide sequence encodes the third framework region (FW3) of a kappa light chain variable domain.

[0028] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $X^1GX^2GTX^3X^4X^5X^6X^7$ (SEQ ID NO: 2242) wherein X^1 is selected from F and L; X^2 is selected from Q, G, and S; X^3 is selected from K and R; X^4 is selected from V and L; X^5 is selected from E, D, and Q; X^6 is selected from I and V; and, X^7 is selected from K and T. In some aspects, the nucleotide sequence encodes a sequence identical to FGQGTKVEIK (SEQ ID NO: 2214) except for at least one substitution selected from F1L, Q3(GS), K6R, V7L, E8(DQ), I9V, and K10T. In some aspects, the nucleotide sequence encodes FGQGTKVEIK (SEQ ID NO: 2214).

[0029] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $FGX^1GTX^2X^3X^4X^5K$ (SEQ ID NO: 2246) wherein X^1 is selected from Q, A, P, and G; X^2 is selected from K and R; X^3 is selected from V and L; X^4 is selected from E and Q; and X^5 is selected from I and L. In some aspects, the nucleotide sequence encodes a sequence identical to FGQGTKVEIK (SEQ ID NO: 2218) except for at least one substitution selected from Q3(APG), K6R, V7L, E8Q, and I9L. In some aspects, the nucleotide sequence encodes FGQGTKVEIK (SEQ ID NO: 2218).

[0030] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1, wherein the nucleotide sequence encodes $FX^1X^2GTX^3X^4X^5IK$ (SEQ ID NO: 2250) wherein X^1 is selected from G and C; X^2 is selected from Q, G, and P; X^3 is selected from K and R; X^4 is selected from V, L, and A; and, X^5 is selected from E and D. In some aspects, the nucleotide sequence encodes a sequence identical to FGQGTKVEIK (SEQ ID NO: 2222) except for at least one

substitution selected from G2C, Q3(GP), K6R, V7(LA), and E8D. In some aspects, the nucleotide sequence encodes FGQGTKVEIK (SEQ ID NO: 2222). In some aspects, the nucleotide sequence encodes the fourth framework region (FW4) of a kappa light chain variable domain.

[0031] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $X^1X^2X^3X^4X^5X^6SGGX^7X^8X^9X^{10}X^{11}GX^{12}SX^{13}X^{14}LX^{15}C$ (SEQ ID NO: 2251) wherein X^1 is selected from E, D, and Q; X^2 is selected from V and A; X^3 is selected from Q, E, and K; X^4 is selected from L and V; X^5 is selected from V and L; X^6 is selected from E and Q; X^7 is selected from G, K, and D; X^8 is selected from L and V; X^9 is selected from V, L, and E; X^{10} is selected from Q, R and K; X^{11} is selected from P, S, and L; X^{12} is selected from G and R; X^{13} is selected from L and R; X^{14} is selected from R and K; and, X^{15} is selected from S and D. In some aspects, the nucleotide sequence encodes a sequence identical to EVQLVESGGGLVQPGGSLRLSC (SEQ ID NO: 2223) except for at least one substitution selected from E1(DQ), V2A, Q3(EK), L4V, V5L, E6Q, G10(KD), L11V, V12(LE), Q13(RK), P14(SL), G16R, L18R, R19K, and S21D. In some aspects, the nucleotide sequence encodes EVQLVESGGGLVQPGGSLRLSC (SEQ ID NO: 2223).

[0032] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $X^1X^2QLX^3QX^4GX^5X^6X^7X^8X^9X^{10}GX^{11}X^{12}X^{13}X^{14}X^{15}SC$ (SEQ ID NO: 2255) wherein X^1 is selected from Q and E; X^2 is selected from V and I; X^3 is selected from V and Q; X^4 is selected from S and P; X^5 is selected from A, S, V, P, T, and G; X^6 is selected from E, G and V; X^7 is selected from V and L; X^8 is selected from K, V, E, and A; X^9 is selected from K, R and Q; X^{10} is selected from P and S; X^{11} is selected from A, E, S, T, and R; X^{12} is selected from S and T; X^{13} is selected from V and L; X^{14} is selected from K and R; and, X^{15} is selected from V, I, L, and M. In some aspects, the nucleotide sequence encodes a sequence identical to QVQLVQSGAEVKKPGASVKVSC (SEQ ID NO: 2227) except for at least one substitution selected from Q1E, V2I, V5Q, S7P, A9(SVPTG), E10(GV), V11L, K12(VEA), K13(RQ), P14S, A16(ESTR), S17T, V18L, K19R, and V20(ILM). In

some aspects, the nucleotide sequence encodes QVQLVQSGAEVKKPGASVKVSC (SEQ ID NO: 2227).

[0033] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $QX^1X^2LX^3X^4X^5GX^6X^7LX^8X^9PX^{10}X^{11}TLX^{12}LTC$ (SEQ ID NO: 2259) wherein X^1 is selected from V and L; X^2 is selected from Q and T; X^3 is selected from Q and R; X^4 is selected from E and Q; X^5 is selected from S and W; X^6 is selected from P and A; X^7 is selected from G and A; X^8 is selected from V and L; X^9 is selected from K and R; X^{10} is selected from S and T; X^{11} is selected from Q and E; and, X^{12} is selected from S and T. In some aspects, the nucleotide sequence encodes a sequence identical to QVQLQESGPGLVKPSQTLSLTC (SEQ ID NO: 2231) except for at least one substitution selected from V2L, Q3T, Q5R, E6Q, S7W, P9A, G10A, V12L, K13R, S15T, Q16E, and S19T. In some aspects, the nucleotide sequence encodes QVQLQESGPGLVKPSQTLSLTC (SEQ ID NO: 2231). In some aspects, the nucleotide sequence encodes the first framework region (FW1) of a heavy chain variable domain.

[0034] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $WX^1RQX^2PX^3KX^4LX^5X^6X^7X^8$ (SEQ ID NO: 2252) wherein X^1 is selected from V, I, and F; X^2 is selected from A, S and T; X^3 is selected from G and E; X^4 is selected from G and R; X^5 is selected from E and D; X^6 is selected from W and L; X^7 is selected from V and I; and, X^8 is selected from A, S, and G. In some aspects, the nucleotide sequence encodes a sequence identical to WVRQAPGKGLEWVA (SEQ ID NO: 2224) except for at least one substitution selected from V2(IF), A5(ST), G7E, G9R, E11D, W12L, V13I, and A14(SG). In some aspects, the nucleotide sequence encodes WVRQAPGKGLEWVA (SEQ ID NO: 2224).

[0035] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes

$WX^1X^2QX^3X^4GX^5X^6LX^7WX^8G$ (SEQ ID NO: 2256) wherein X^1 is selected from V and I; X^2 is selected from R and K; X^3 is selected from A, M, N, R, K, T, and S; X^4 is selected from P, T, and H; X^5 is selected from Q, K, and R; X^6 is selected from G, R and S; X^7 is selected from E, D, K, Q, and A; and, X^8 is selected from M, I, and V. In some aspects, the nucleotide sequence encodes a sequence identical to WVRQAPGQGLEWMG (SEQ ID NO: 2228) except for at least one substitution selected from V2I, R3K, A5(MNRKTS), P6(TH), Q8(KR), G9(RS), E11(DKQA), and M13(IV). In some aspects, the nucleotide sequence encodes WVRQAPGQGLEWMG (SEQ ID NO: 2228).

[0036] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $WX^1RX^2X^3X^4X^5X^6X^7LX^8WX^9X^{10}$ (SEQ ID NO: 2260) wherein X^1 is selected from I and V; X^2 is selected from Q and H; X^3 is selected from L, P, S, and H; X^4 is selected from P and S; X^5 is selected from G and E; X^6 is selected from K and R; X^7 is selected from G and A; X^8 is selected from E and Q; X^9 is selected from I and L; and, X^{10} is selected from G and A. In some aspects, the nucleotide sequence encodes a sequence identical to WIRQLPGKGLEWIG (SEQ ID NO: 2232) except for at least one substitution selected from I2V, Q4H, L5(PSH), P6S, G7E, K8R, G9A, E11Q, I13L, and G14A. In some aspects, the nucleotide sequence encodes WIRQLPGKGLEWIG (SEQ ID NO: 2232). In some aspects, the nucleotide sequence encodes the second framework region (FW2) of a heavy chain variable domain.

[0037] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $X^1X^2X^3X^4SX^5DX^6X^7X^8X^9X^{10}X^{11}X^{12}LX^{13}X^{14}X^{15}X^{16}LX^{17}X^{18}EDTX^{19}X^{20}X^{21}X^{22}C$ (SEQ ID NO: 2253) wherein X^1 is selected from R and K; X^2 is selected from F and V; X^3 is selected from T, I, and A; X^4 is selected from L and I; X^5 is selected from V, R, L, and A; X^6 is selected from R, N, T, D, K, and S; X^7 is selected from S, A and V; X^8 is selected from K, R, and E; X^9 is selected from N, S, R, H, and T; X^{10} is selected from T and S; X^{11} is selected from L, A, and F; X^{12} is selected from Y and F; X^{13} is selected from Q and E; X^{14} is selected from M and V; X^{15} is selected from N, D, and S; X^{16} is selected from S, G,

and I; X^{17} is selected from R and K; X^{18} is selected from A, S, D, V, and P; X^{19} is selected from A and G; X^{20} is selected from V, M, and L; X^{21} is selected from Y and F; and, X^{22} is selected from Y and F. In some aspects, the nucleotide sequence encodes a sequence identical to RFTLSVDRSKNTLYLQMNSLRAEDTAVYYC (SEQ ID NO: 2225) except for at least one substitution selected from R1K, F2V, T3(IA), L4I, V6(RLA), R8(NTDKS), S9(AV), K10(RE), N11(SRHT), T12S, L13(AF), Y14F, Q16E, M17V, N18(DS), S19(GI), R21K, A22(SDVP), A26G, V27(ML), Y28F, and Y29F. In some aspects, the nucleotide sequence encodes RFTLSVDRSKNTLYLQMNSLRAEDTAVYYC (SEQ ID NO: 2225).

[0038] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $X^1X^2X^3X^4X^5X^6X^7X^8SX^9X^{10}TX^{11}X^{12}X^{13}X^{14}X^{15}X^{16}X^{17}LX^{18}X^{19}X^{20}DX^{21}X^{22}X^{23}YX^{24}C$ (SEQ ID NO: 2257) wherein X^1 is selected from R, Q, and K; X^2 is selected from V, I, F, G, and A; X^3 is selected from T, A, and K; X^4 is selected from M, I, L, and F; X^5 is selected from T and S; X^6 is selected from T, A, R, V, S, E, and L; X^7 is selected from D, E, and N; X^8 is selected from T, K, Q, S, P, R, I, N, and E; X^9 is selected from T, K, S, A, I, and V; X^{10} is selected from S, N, D, and T; X^{11} is selected from A, V, and T; X^{12} is selected from Y, and F; X^{13} is selected from M and L; X^{14} is selected from E, Q, and D; X^{15} is selected from L, M, W, and I; X^{16} is selected from R, S, D, L, K, T, and N; X^{17} is selected from S and R; X^{18} is selected from R, K, Q, and T; X^{19} is selected from S, H, F, A, and P; X^{20} is selected from D, E, and S; X^{21} is selected from T and S; X^{22} is selected from A and G; X^{23} is selected from V, F, T, and M; and, X^{24} is selected from Y, F, and L. In some aspects, the nucleotide sequence encodes a sequence identical to RVTMTTDTSTSTAYMELRSLRSDDTAVYYC (SEQ ID NO: 2229) except for at least one substitution selected from R1(QK), V2(IFGA), T3(AK), M4(ILF), T5S, T6(ARVSEL), D7(EN), T8(KQSPRINE), T10(KSAIV), S11(NDT), A13(VT), Y14F, M15L, E16(QD), L17(MWI), R18(SDLKTN), S19R, R21(KQT), S22(HFAP), D23(ES), T25S, A26G, V27(FTM), and Y29(FL). In some aspects, the nucleotide sequence encodes RVTMTTDTSTSTAYMELRSLRSDDTAVYYC (SEQ ID NO: 2229).

[0039] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5,

MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $RX^1X^2X^3X^4X^5DX^6SX^7X^8QX^9X^{10}LX^{11}X^{12}X^{13}X^{14}X^{15}X^{16}X^{17}X^{18}DTAX^{19}X^{20}X^{21}C$ (SEQ ID NO: 2261) wherein X^1 is selected from V and L; X^2 is selected from T and S; X^3 is selected from I and M; X^4 is selected from S and L; X^5 is selected from V, R, and K; X^6 is selected from T and K; X^7 is selected from K and R; X^8 is selected from K and N; X^9 is selected from F and V; X^{10} is selected from S and V; X^{11} is selected from R, T, K, and M; X^{12} is selected from L, I, M, and V; X^{13} is selected from S, T, and N; X^{14} is selected from S and N; X^{15} is selected from V and M; X^{16} is selected from T and D; X^{17} is selected from A and P; X^{18} is selected from A and V; X^{19} is selected from V and T; X^{20} is selected from Y and W; and, X^{21} is selected from Y, F and W. In some aspects, the nucleotide sequence encodes a sequence identical to RVTISVDTSKKQFSLRLSSVTAADTAVYYC (SEQ ID NO: 2233). except for at least one substitution selected from V2L, T3S, I4M, S5L, V6(RK), T8K, K10R, K11N, F13V, S14V, R16(TKM), L17(IMV), S18(TN), S19N, V20M, T21D, A22P, A23V, V27T, Y28W, and Y29(FW). In some aspects, the nucleotide sequence encodes RVTISVDTSKKQFSLRLSSVTAADTAVYYC (SEQ ID NO: 2233). In some aspects, the nucleotide sequence encodes the third framework region (FW3) of a heavy chain variable domain.

[0040] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP 9,MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $WGX^1GX^2X^3VTVS$ (SEQ ID NO: 2254) wherein X^1 is selected from Q, R, and K; X^2 is selected from T, I and A; and, X^3 is selected from L, S, T, M, and P. In some aspects, the nucleotide sequence encodes a sequence identical to WGQGTLVTVS (SEQ ID NO: 2226) except for at least one substitution selected from Q3(RK), T5(IA), and L6(STMP). In some aspects, the nucleotide sequence encodes WGQGTLVTVS (SEQ ID NO: 2226).

[0041] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $WGX^1GTX^2X^3TVS$ (SEQ ID NO: 2258) wherein X^1 is selected from R, Q, K, A and S; X^2 is selected from L, M, T, Q, and P; and, X^3 is selected from V and L. In some aspects,

the nucleotide sequence encodes a sequence identical to WGRGTLVTVS (SEQ ID NO: 2230) except for at least one substitution selected from R3(QKAS), L6(MTQP), and V7L. In some aspects, the nucleotide sequence encodes WGRGTLVTVS (SEQ ID NO: 2230).

[0042] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $WX^1X^2GX^3X^4VTVS$ (SEQ ID NO: 2262) wherein X^1 is selected from G and D; X^2 is selected from Q and R; X^3 is selected from T and S; and, X^4 is selected from T, L, and M. In some aspects, the nucleotide sequence encodes a sequence identical to WGQGTTVTVS (SEQ ID NO: 2234).except for at least one substitution selected from G2D, Q3R, T5S, and T6(LM). In some aspects, the nucleotide sequence encodes WGQGTTVTVS (SEQ ID NO: 2234). In some aspects, the nucleotide sequence encodes the fourth framework region (FW4) of a heavy chain variable domain.

[0043] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes a sequence of formula (GlyxSer)y, wherein x and y are integers between 1 and 100. In some aspects, a polynucleotide disclosed herein further comprises a nucleotide sequence which encodes a sequence of formula $(Gly_xSer)_y$, wherein x and y are integers between 1 and 100. In some aspects, the sequence of formula $(Gly_xSer)_y$ is a linker. In some aspects, the linker comprises the sequence (Gly_4Ser), (Gly_3Ser), (Gly_2Ser), or a combination thereof. In some aspects, the linker comprises the sequence $(Gly_4Ser)_3$. In some aspects, the linker is interposed between a VH domain and a VL domain. In some aspects, the polynucleotide encodes an scFv.

[0044] Also provided is a polynucleotide encoding an antibody or an antigen binding portion thereof comprising (i) a polynucleotide comprising a nucleotide sequence encoding the first framework region (FW1) of a lambda light chain or a kappa light chain variable domain, (iii) a polynucleotide comprising a nucleotide sequence encoding the second framework region (FW2) of a lambda light chain or a kappa light chain variable domain, (iii) a polynucleotide comprising a nucleotide sequence encoding the third

framework region (FW3) of a lambda light chain or a kappa light chain variable domain, (iv) a polynucleotide comprising a nucleotide sequence encoding the fourth framework region (FW4) of a lambda light chain or a kappa light chain variable domain, or (v) any combination thereof.

[0045] Also provided is a polynucleotide encoding an antibody or an antigen binding portion thereof comprising (i) a polynucleotide comprising a nucleotide sequence encoding the first framework region (FW1) of a lambda light chain or a kappa light chain variable domain, (iii) a polynucleotide comprising a nucleotide sequence encoding the second framework region (FW2) of a lambda light chain or a kappa light chain variable domain, (iii) a polynucleotide comprising a nucleotide sequence encoding the third framework region (FW3) of a lambda light chain or a kappa light chain variable domain, and (iv) a polynucleotide comprising a nucleotide sequence encoding the fourth framework region (FW4) of a lambda light chain or a kappa light chain variable domain.

[0046] Also provided is a polynucleotide encoding an antibody or an antigen binding portion thereof comprising (i) a nucleotide sequence encoding the first framework region (FW1) of a heavy chain variable domain, (iii) a nucleotide sequence encoding the second framework region (FW2) of a heavy chain variable domain, (iii) a nucleotide sequence encoding the third framework region (FW3) of a heavy chain variable domain, (iv) a nucleotide sequence encoding the fourth framework region (FW4) of a heavy chain variable domain, or (v) any combination thereof. Also provided is a polynucleotide encoding an antibody or an antigen binding portion thereof comprising (i) a nucleotide sequence encoding the first framework region (FW1) of a heavy chain variable domain, (iii) a nucleotide sequence encoding the second framework region (FW2) of a heavy chain variable domain, (iii) a nucleotide sequence encoding the third framework region (FW3) of a heavy chain variable domain, and (iv) a nucleotide sequence encoding the fourth framework region (FW4) of a heavy chain variable domain. In some aspects, a polynucleotide comprising nucleotides encoding the FW1-FW4 regions of a light chain also comprises nucleotides encoding the FW1-FW4 regions of a light chain. In some aspects, a polypeptide comprising nucleotides encoding the FW1-FW4 regions of a light chain and/or nucleotides encoding the FW1-FW4 regions of a light chain further comprises nucleotides encoding a constant domain (e.g., CL, CH1, CH2, CH3, or a combination thereof).

[0047] The present disclosure also provides a polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to (i) any one of the polynucleotides of SEQ ID NOS:1-88, or (ii) a subsequence of any one of the polynucleotides of SEQ ID NOS: 89-1978 encoding an Ig constant domain, wherein the nucleotide subsequence encodes an Immunoglobulin (Ig) polypeptide that has a significant match to a corresponding sequence of CDD domain CD00098 (FIG. 1). In some aspects, the Ig polypeptide comprises an Ig constant domain of an antibody or a fragment thereof. In some aspects, the Ig constant domain is a CL, CH1, CH2, or CH3 constant domain from an IgG.

[0048] Also provided is a polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to (i) any one of the polynucleotides of SEQ ID NOS:1-8, or 45-52, or (ii) a subsequence of any one of the polynucleotides of SEQ ID NOS:1034-1978 encoding an Ig light chain constant domain (CL), wherein the nucleotide subsequence encodes an Ig polypeptide that has a significant match to a corresponding sequence of CDD domain CD07699 (FIG. 2). In some aspects, the Ig polypeptide comprises a light chain constant region of an antibody or a fragment thereof.

[0049] Also provided is a polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to (i) any one of the polynucleotides of SEQ ID NOS:9-12, 21-24, 33-36, 53-56, 65-68, or 77-80, or (ii) a subsequence of any one of the polynucleotides of SEQ ID NOS: 89-1033 encoding a CH1 constant domain, wherein the nucleotide subsequence encodes an Ig polypeptide that has a significant match to a corresponding sequence of CDD domain CD04985 (FIG. 3). In some aspects, the Ig polypeptide comprises a heavy chain CH1 constant domain of an antibody or a fragment thereof.

[0050] Also provided is a polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to (i) any one of the polynucleotides of SEQ ID NO:13-16, 25-28, 37-40, 57-60, 69-72, or 81-84, or (ii) a subsequence of any one of the polynucleotides of SEQ ID NOS: 89-1033 encoding a CH2 constant domain, wherein the nucleotide subsequence encodes an Ig polypeptide that has a significant match to a

corresponding sequence of CDD domain CD04986 (FIG. 4). In some aspects, the Ig polypeptide comprises a heavy chain CH2 constant domain of an antibody or a fragment thereof.

[0051] Also provided is a polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to (i) any one of the polynucleotides of SEQ ID NO:17-20, 29-32, 41-44, 61-64, 73-76, or 85-88, or (ii) a subsequence of any one of the polynucleotides of SEQ ID NOS: 89-1033 encoding a CH3 constant region, wherein the nucleotide subsequence encodes an Ig polypeptide that has a significant match to a corresponding sequence of CDD domain CD07696 (FIG. 5). In some aspects, the Ig polypeptide comprises a heavy chain CH3 constant domain of an antibody or a fragment thereof.

[0052] Also provided is a polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to a subsequence of any one of the polynucleotides of SEQ ID NOS: 89-1978 encoding a variable domain, wherein the nucleotide subsequence encodes an Ig polypeptide that has a significant match to a corresponding sequence of CDD domain CD00099 (FIG. 6). In some aspects, the Ig polypeptide comprises a variable domain of an antibody or a fragment thereof.

[0053] Also provided is a polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to a subsequence of any one of the polynucleotides of SEQ ID NOS: 89-1033 encoding a VH domain, wherein the nucleotide subsequence encodes an Ig polypeptide that has a significant match to a corresponding sequence of CDD domain CD04981 (FIG. 7). In some aspects, the Ig polypeptide comprises a VH domain of an antibody or a fragment thereof.

[0054] Also provided is a polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical a subsequence of any one of the polynucleotides of SEQ ID NOS: 1034-1978 encoding a VL domain, wherein the nucleotide subsequence encodes an Ig polypeptide that has a significant match to a corresponding sequence of CDD domain CD04980 (FIG. 8) or CD04984 (FIG. 9). In some aspects, the Ig

polypeptide comprises a VL kappa domain or a VL lambda domain of an antibody or a fragment thereof.

[0055] Also provided is a polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to any one of the polynucleotides of SEQ ID NOS:89-1033, wherein the nucleotide sequence encodes an Ig polypeptide that has non-overlapping significant matches to CDD domains CD04981/CD4984, CD04985, and CD04986. In some aspects, the Ig polypeptide comprises the heavy chain of an antibody or a fragment thereof.

[0056] Also provided is a polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to any one of the polynucleotides of SEQ ID NO:1034-1978, wherein the nucleotide sequence encodes an Ig polypeptide that has non-overlapping significant matches to CD04980 and CD07699. In some aspects, the Ig polypeptide comprises the light chain of an antibody or a fragment thereof.

[0057] Also provided is a polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to any one of the polynucleotides of SEQ ID NOS: 1-4, or 45-48, wherein the nucleotide sequence encodes a CL kappa domain or a functional fragment thereof from a therapeutic antibody. In some aspects, the CL kappa domain comprises TVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQ ESVTEQDSKDSTYSLSX^{1}TLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNR (SEQ ID NO: 2200), wherein X^1 is selected from N and S.

[0058] Also provided is a polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to any one of the polynucleotides of SEQ ID NOS: 5-8, or 49-52, wherein the nucleotide sequence encodes a CL lambda domain or a functional fragment thereof from a therapeutic antibody. In some aspects, the CL lambda domain comprises PKAAPSVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVK AGVETTTPSKQSNNKYAASSYLSLTPEQWKSHX^{2}SYSCQVTHEGSTVEKTVAPX^3 ECS (SEQ ID NO: 2201), wherein X^2 is selected from R and K, and X^3 is selected from T and A.

[0059] Also provided is a polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to any one of the polynucleotides of SEQ ID NOS: 9-12, 21-24, 33-36, 53-56, 65-68, or 77-80, wherein the nucleotide sequence encodes a CH1 domain or a functional fragment thereof from a therapeutic antibody. In some aspects, the nucleotide sequence is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO: 9-12, or 53-56 and the CH1 domain is an IgG1 CH1 domain. In some aspects, the IgG1 CH1 domain comprises $SX^4GPSVX^5PLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGVHTFPAVLQSSG$ $LYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKX^6X^7$ (SEQ ID NO: 2202) wherein X^4 is an optional ASTK sequence, X^5 is selected from F and L, X^6 is selected from K and R, and X^7 is selected from V and A. In some aspects, the nucleotide sequence is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO: 21-24, or 65-68 and the CH1 domain is an IgG2 CH1 domain. In some aspects, the IgG2 CH1 domain comprises $SASTKGPSVFPLAPCSRSTSESTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPA$ $VLQSSGLYSLSSVVTVX^{15}SSNFGTQTYTCNVDHKPSNTKVDKTV$ (SEQ ID NO: 2205) wherein X^{15} is selected from P and T. In some aspects, the nucleotide sequence is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO: 33-36, or 77-80 and the CH1 domain is an IgG4 CH1 domain. In some aspects, the IgG4 CH1 domain comprises $SASTKGPSVFPLAPCSRSTSESTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPA$ $VLQSSGLYSLSSVVTVPSSSLGTKTYTCNVDHKPSNTKVDKRV$ (SEQ ID NO: 2197).

[0060] Also provided is a polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to any one of the polynucleotides of SEQ ID NOS: 13-16, 25-28, 37-40, 57-60, 69-72, or 81-84, wherein the nucleotide sequence encodes a CH2 domain or a functional fragment thereof from a therapeutic antibody. In some aspects, the nucleotide sequence is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO: 13-

16, or 57-60 and the CH2 domain is an IgG1 CH2 domain. In some aspects, the IgG1 CH2 domain comprises APEX^{8}X^{9}GX^{10}PSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNX^{11}YVDGV EVHNAKTKPREEQYX^{12}STYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK TISKAK (SEQ ID NO: 2203) wherein X^{8} is selected from L and A, X^{9} is selected from L and A, X^{10} is selected from G and A, and X^{11} is selected from V and W, and X^{12} is selected from N and A. In some aspects, the nucleotide sequence is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO: 25-28, or 69-72 and the CH2 domain is an IgG2 CH2 domain. In some aspects, the IgG2 CH2 domain comprises APPVAGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFNWYVDGX^{16} EVHNAKTKPREEQFNSTFRVVSVLTVVHQDWLNGKEYKCKVSNKGLPX^{17}X^{18}IE KTISKTK (SEQ ID NO: 2206) wherein X^{16} is selected from V and M, X^{17} is selected from A and S; and X^{18} is selected from P and S. In some aspects, the nucleotide sequence is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO: 37-40, or 81-84 and the CH2 domain is an IgG4 CH2 domain In some aspects, the IgG4 CH2 domain comprises APEFLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSQEDPEVQFNWYVDGVEVH NAKTKPREEQFNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKGLPSSIEKTISKA K (SEQ ID NO: 2198).

[0061] Also provided is a polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to any one of the polynucleotides of SEQ ID NOS: 17-20, 29-32, 41-44, 61-64, 73-76, or 85-88, wherein the nucleotide sequence encodes a CH3 domain or a functional fragment thereof from a therapeutic antibody. In some aspects, the nucleotide sequence is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO: 17-20, or 61-64 and the CH3 domain is an IgG1 CH3 domain. In some aspects, the IgG1 CH3 domain comprises GQPREPQVYTLPPSRX^{13}EX^{14}TKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYK TTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG (SEQ ID NO: 2204) wherein X^{13} is selected from E and D, and X^{14} is selected from M

and L. In some aspects, the nucleotide sequence is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO: 29-32, or 73-76 and the CH3 domain is an IgG2 CH3 domain. In some aspects, the IgG2 CH3 domain comprises GQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYK TTPPMLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG (SEQ ID NO: 2196). In some aspects, the nucleotide sequence is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO: 41-44, or 85-88 and the CH3 domain is an IgG4 CH3 domain. In some aspects, the IgG4 CH3 domain comprises GQPREPQVYTLPPSQEEMTKNQVSLTCLVKGF YPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSRLTVDKSRWQEGNVFSCSV MHEALHNHYTQKSLSLSLG (SEQ ID NO: 2199).

[0062] Also provided is a polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to a subsequence from any one of the polynucleotides of SEQ ID NOS: 89-1978, wherein said subsequence encodes (a) one, two, or three VH-CDRs from a therapeutic antibody;(b) one, two, or three VL-CDRs from a therapeutic antibody; (c) one, two, three, or four VH framework (FW) regions from a therapeutic antibody; (d) one, two, three, or four VL framework (FW) regions from a therapeutic antibody; (e) a VH domain from a therapeutic antibody; (f) a VL domain from a therapeutic antibody; (g) a CL domain of a therapeutic antibody; (h) a CH1 domain of a therapeutic antibody; (i) a CH2 domain of a therapeutic antibody; (j) a CH3 domain of a therapeutic antibody; or, (k) a combination thereof. In some aspects, the subsequence encoding one, two, three, or four VH framework (FW) regions from a therapeutic antibody comprises a codon-optimized nucleotide sequence encoding a first framework region (FW1) of a heavy chain variable domain disclosed herein; a codon-optimized nucleotide sequence a second framework region (FW2) of a heavy chain variable domain disclosed herein; a codon-optimized nucleotide sequence encoding a third framework region (FW3) of a heavy chain variable domain disclosed herein; a codon-optimized nucleotide sequence encoding a fourth framework region (FW4) of a heavy chain variable domain disclosed herein; or any combinations thereof. In one aspects, the subsequence encoding one, two, three, or four VL framework (FW) regions from a therapeutic

antibody comprises a codon-optimized nucleotide sequence encoding a first framework region (FW1) of a light chain variable domain disclosed herein; a codon-optimized nucleotide sequence a second framework region (FW2) of a light chain variable domain disclosed herein; a codon-optimized nucleotide sequence encoding a third framework region (FW3) of a light chain variable domain disclosed herein; a codon-optimized nucleotide sequence encoding a fourth framework region (FW4) of a light chain variable domain disclosed herein; or any combinations thereof. In some aspects, the subsequence encoding a CL domain of a therapeutic antibody comprises a polynucleotide comprising a codon-optimized nucleotide sequence encoding a kappa light chain constant domain of an antibody or a fragment thereof or a lambda light chain constant domain of an antibody or a fragment thereof disclosed herein. In some aspects, the subsequence encoding a CH1 domain of a therapeutic antibody comprises a polynucleotide comprising a codon-optimized nucleotide sequence encoding a CH1 domain disclosed herein. In some aspects, the subsequence encoding a CH2 domain of a therapeutic antibody comprises a polynucleotide comprising a codon-optimized nucleotide sequence encoding a CH2 domain disclosed herein. In some aspects, the subsequence encoding a CH3 domain of a therapeutic antibody comprises a polynucleotide comprising a codon-optimized nucleotide sequence encoding CH3 domain disclosed herein. In some aspects, the polynucleotide sequences disclosed above can comprise a nucleotide sequence encoding a linker. In some aspects, the nucleotide sequence encoding a linker is codon-optimized. In some aspects, the polynucleotide comprising a nucleotide sequence encoding a linker encodes an scFv.

[0063] In some aspects, the therapeutic antibody is selected from the group consisting of abagovomab, abciximab, adalimumab, alemtuzumab, alirocumab, amatuximab, anrukinzumab, arcitumomab, basiliximab, bavituximab, benralizumab, bevacizumab, bezlotoxumab, bimagrumab, bococizumab, brentuximab, briakinumab, brodalumab, canakinumab, cantuzumab, carlumab, cetuximab, cixutumumab, clivatuzumab, conatumumab, crenezumab, dacetuzumab, daclizumab, dalotuzumab, denosumab, drozitumab, dupilumab, dusigitumab, eculizumab, elotuzumab, enokizumab, epratuzumab, etaracizumab, evolocumab, farletuzumab, fasinumab, fezakinumab, ficlatuzumab, figitumumab, fresolimumab, fulranumab, ganitumab, gantenerumab, gevokizumab, girentuximab, glembatumumab, ibalizumab, ibritumomab, icrucumab, inotuzumab, intetumumab, itolizumab, ixekizumab, lebrikizumab, lorvotuzumab,

mavrilimumab, mepolizumab, milatuzumab, mogamulizumab, motavizumab, naptumomab, necitumumab, nivolumab, obinutuzumab, ocrelizumab, olaratumab, omalizumab, otelixizumab, oxelumab, pateclizumab, pembrolizumab, pertuzumab, ponezumab, ramucirumab, rilotumumab, rituximab, robatumumab, romosozumab, rontalizumab, samalizumab, sarilumab, secukinumab, sifalimumab, siltuximab, sirukumab, solanezumab, tabalumab, tanezumab, tenatumomab, teplizumab, tigatuzumab, tildrakizumab, tocilizumab, tositumomab, tralokinumab, trastuzumab, urelumab, ustekinumab, vedolizumab, and veltuzumab, and functional fragments thereof. In some aspects, the functional fragment is an antigen-binding fragment. In some aspects, the functional fragment is a non-antigen-binding fragment. In some aspects, the non-antigen-binding fragment is an Fc fragment.

[0064] The present disclosure also provides a polynucleotide comprising a nucleotide sequence encoding an antibody or a fragment thereof, wherein Ala is encoded by GCC, GCG or GCT; Cys is encoded by TGC or TGT; Asp is encoded by GAC; Glu is encoded by GAG or GAA; Phe is encoded by TTC; Gly is encoded by GGC, GGT, or GGG; His is encoded by CAC; Ile is encoded by ATC or ATT; Lys is encoded by AAG; Leu is encoded by CTG, CTC or TTG; Met is encoded by ATG; Asn is encoded by AAC; Pro is encoded by CCC, CCA or CCG; Gln is encoded by CAG or CAA, Arg is encoded by CGG, AGG, CGC, CGT, AGA, CGA, Ser is encoded by AGC, TCC or TCT, Thr is encoded by ACC, ACG or ACT, Val is encoded by GTG, GTC or GTT, Trp is encoded by TGG, and Tyr is encoded by TAC, wherein the nucleotide sequence encodes abagovomab, abciximab, adalimumab, alemtuzumab, alirocumab, amatuximab, anrukinzumab, arcitumomab, basiliximab, bavituximab, benralizumab, bevacizumab, bezlotoxumab, bimagrumab, bococizumab, brentuximab, briakinumab, brodalumab, canakinumab, cantuzumab, carlumab, cetuximab, cixutumumab, clivatuzumab, conatumumab, crenezumab, dacetuzumab, daclizumab, dalotuzumab, denosumab, drozitumab, dupilumab, dusigitumab, eculizumab, elotuzumab, enokizumab, epratuzumab, etaracizumab, evolocumab, farletuzumab, fasinumab, fezakinumab, ficlatuzumab, figitumumab, fresolimumab, fulranumab, ganitumab, gantenerumab, gevokizumab, girentuximab, glembatumumab, ibalizumab, ibritumomab, icrucumab, inotuzumab, intetumumab, itolizumab, ixekizumab, lebrikizumab, lorvotuzumab, mavrilimumab, mepolizumab, milatuzumab, mogamulizumab, motavizumab, naptumomab, necitumumab, nivolumab, obinutuzumab, ocrelizumab, olaratumab,

omalizumab, otelixizumab, oxelumab, pateclizumab, pembrolizumab, pertuzumab, ponezumab, ramucirumab, rilotumumab, rituximab, robatumumab, romosozumab, rontalizumab, samalizumab, sarilumab, secukinumab, sifalimumab, siltuximab, sirukumab, solanezumab, tabalumab, tanezumab, tenatumomab, teplizumab, tigatuzumab, tildrakizumab, tocilizumab, tositumomab, tralokinumab, trastuzumab, urelumab, ustekinumab, vedolizumab, veltuzumab, or an antigen binding fragment thereof. In some aspects, the nucleotide sequence is selected from SEQ ID NOS: 89-1978, and subsequences thereof encoding functional fragments, e.g., antigen binding fragments.

[0065] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes a fragment of (i) the sequences of SEQ ID NO: 1979-2006; or, (ii) a polypeptide sequence encoded by the nucleotide of any one of claims 1 to 199 and wherein the fragment is about 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 205, 210, 215, 220, 225, 230, 235, 240, 245, 250, 255, 260, 265, 270, 275, 280, 285, 290, 295, 300, 305, 310, 315, 320, 325, 330, 335, 340, 345, 350, 355, 360, 365, 370, 375, 380, 385, 390, 395, 400, 405, 410, 415, 420, 425, 430, 435, 440, 445, 450, 455, 460, 465, 470, 475, 480, 485, 490, 495, 500, 505, 510, 515, 520, 525, 530, 535, 540, 545, or 550 amino acids long.

[0066] Each of the nucleotide sequences disclosed herein is not a wild type nucleotide sequence encoding a therapeutic antibody known in the art. In some aspects, the nucleotide sequence has been optimized according to a method comprising (i) modifying at least one subsequence in a candidate nucleic acid sequence to generate a ramp subsequence; (ii) substituting at least one codon in a candidate nucleic acid sequence with an alternative codon to increase or decrease uridine content to generate a uridine-modified sequence; (iii) substituting at least one codon in a candidate nucleic acid sequence or the uridine-modified sequence with a fast recharging codon; (iv) substituting at least one codon in a candidate nucleic acid sequence with an alternative codon having a higher codon frequency in the synonymous codon set; (v) substituting at least one natural nucleobase in a candidate nucleic acid sequence with an alternative synthetic nucleobase; (vi) substituting at least one internucleoside linkage in a candidate nucleic acid sequence

with a non-natural internucleoside linkage; or, (vii) a combination thereof, wherein the resulting optimized nucleic acid sequence has at least one optimized property with respect to the candidate nucleic acid sequence. In some aspects, the method is multiparametric and comprises one, two, three, four, five or six optimization methods selected from the group consisting of (i) modifying at least one subsequence in a candidate nucleic acid sequence to generate a ramp subsequence; (ii) substituting at least one codon in a candidate nucleic acid sequence with an alternative codon to increase or decrease uridine content to generate a uridine-modified sequence; (iii) substituting at least one codon in a candidate nucleic acid sequence or the uridine-modified sequence with a fast recharging codon; (iv) substituting at least one codon in a candidate nucleic acid sequence with an alternative codon having a higher codon frequency in the synonymous codon set; (v) substituting at least one natural nucleobase in a candidate nucleic acid sequence with an alternative synthetic nucleobase; and (vi) substituting at least one internucleoside linkage in a candidate nucleic acid sequence with a non-natural internucleoside linkage. In preferred aspects the substitutions are to the polynucleotide, as above-described, and the encoded antibody sequence is as described herein, for example (i) the amino acid sequence of any one of SEQ ID NOS:1979-2188 or a functional fragment thereof, (ii) a sequence corresponding to any one of the consensus sequences disclosed herein or a combination thereof.

[0067] In some aspects, the multiparametric method comprises replacing at least 5%, at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% of the codons in the candidate nucleic acid sequence. In some aspects, the candidate nucleic acid sequence is SEQ ID NOS: 1979-2188, or a fragment thereof. In some aspects, the fragment comprises (a) one, two, or three VH-CDRs from SEQ ID NOS: 1979-2083; (b) one, two, or three VL-CDRs from SEQ ID NOS: 2084-2188; (c) one, two, three, or four VH framework (FW) regions from SEQ ID NOS: 1979-2083; (d) one, two, three, or four VL framework (FW) regions from SEQ ID NOS: 2084-2188; (e) a VH domain from SEQ ID NOS: 1979-2083; (f) a VL domain from SEQ ID NOS: 2084-2188; (g) a CL domain from SEQ ID NOS: 2084-2188; (h) a CH1 domain from SEQ ID NOS: 1979-2083; (i) a CH2 domain from SEQ ID NOS:

1979-2083; (j) a CH3 domain from SEQ ID NOS: 1979-2083; or, (k) a combination thereof.

[0068] In some aspects of the polynucleotides disclosed above, the polynucleotide is a DNA. In other aspects, the polynucleotide is an RNA. In some aspects, the RNA is mRNA. In some aspects, the mRNA is synthetic. In some aspects, the polynucleotide comprises at least one nucleotide analogue. In some aspects, the at least one nucleotide analogue is selected from the group consisting of a 5-methoxyuridine, 1-methyl-pseudouridine, 1-ethyl-pseudouridine, 2'-O-methoxyethyl-RNA (2'-MOE-RNA) monomer, a 2'-fluoro-DNA monomer, a 2'-O-alkyl-RNA monomer, a 2'-amino-DNA monomer, a locked nucleic acid (LNA) monomer, a cEt monomer, a cMOE monomer, a 5'-Me-LNA monomer, a 2'-(3-hydroxy)propyl-RNA monomer, an arabino nucleic acid (ANA) monomer, a 2'-fluoro-ANA monomer, an anhydrohexitol nucleic acid (HNA) monomer, an intercalating nucleic acid (INA) monomer, and a combination of two or more of said nucleotide analogues. In some aspects, the polynucleotide comprises at least one backbone modification. In some aspects, the at least one backbone modification is a phosphorothioate internucleotide linkage. In some aspects, all of the internucleotide linkages are phosphorothioate internucleotide linkages. In some aspects (i) at least one uridine has been replaced with 2- pseudouridine, 2-thiouridine, 4-thiouridine, N1-methylpseudouridine, 5-aza-uridine, 2-thio-5-aza-uridine, 4-thio-pseudouridine, 2-thio-pseudouridine, 5-hydroxyuridine, 4-methoxy-pseudouridine, 4-methoxy-2-thio-pseudouridine, 3-methyluridine, 5-carboxymethyl-uridine, 1-carboxymethyl-pseudouridine, 5-propynyl-uridine, 1-propynyl-pseudouridine, 2-methoxy-4-thio-uridine, 5-taurinomethyluridine, 1-taurinomethyl-pseudouridine, 5-taurinomethyl-2-thio-uridine, 1-taurinomethyl-4-thio-uridine, 5-methyl-uridine, 2-methoxyuridine, 4-thio-1-methyl-pseudouridine, 2-thio-1-methyl-pseudouridine, 1-methyl-1-deaza-pseudouridine, 2-thio-1-methyl-1-deaza-pseudouridine, or 2-thio-dihydrouridine; and/or, (ii) at least one adenosine has been replaced with 2-aminopurine, 2,6-diaminopurine, 7-deaza-adenine, 7-deaza-8-aza-adenine, 7-deaza-2-aminopurine, 7-deaza-8-aza-2-aminopurine, 7-deaza-2,6-diaminopurine, 7-deaza-8-aza-2,6-diaminopurine, 1-methyladenosine, N6-methyladenosine, N6-isopentenyladenosine, N6-(cis-hydroxyisopentenyl)adenosine, 2-methylthio-N6-(cis-hydroxyisopentenyl)adenosine, N6-glycinylcarbamoyladenosine, N6-threonylcarbamoyladenosine, 2-methylthio-N6-threonyl carbamoyladenosine, N6,N6-dimethyladenosine, or 7-methyladenine; and/or, (iii) at least one guanosine has been

replaced with inosine, 1-methyl-inosine, wyosine, wybutosine, 7-deaza-guanosine, 7-deaza-8-aza-guanosine, 6-thio-guanosine, 6-thio-7-deaza-guanosine, 6-thio-7-deaza-8-aza-guanosine, 7-methyl-guanosine, 6-thio-7-methyl-guanosine, 7-methylinosine, 6-methoxy-guanosine, 1-methylguanosine, N2-methylguanosine, N2,N2-dimethylguanosine, 8-oxo-guanosine, 7-methyl-8-oxo-guanosine, or 1-methyl-6-thio-guanosine; and/or, (iv) at least one cytidine has been replaced with 5-methylcytidine, 5-aza-cytidine, pseudoisocytidine, 3-methyl-cytidine, N4-acetylcytidine, 5-formylcytidine, N4-methylcytidine, 5-hydroxymethylcytidine, 1-methyl-pseudoisocytidine, pyrrolo-cytidine, pyrrolo-pseudoisocytidine, 2-thio-cytidine, 2-thio-5-methyl-cytidine, 4-thio-pseudoisocytidine, 4-thio-1-methyl-pseudoisocytidine, 4-thio-1-methyl-1-deaza-pseudoisocytidine, 1-methyl-1-deaza-pseudoisocytidine, zebularine, 5-aza-zebularine, 5-methyl-zebularine, 5-aza-2-thio-zebularine, 2-thio-zebularine, 2-methoxy-cytidine, 2-methoxy-5-methyl-cytidine, or 4-methoxy-pseudoisocytidine.

[0069] In some aspects, a polynucleotide disclosed herein has been optimized by replacing 25% of uridines with 4-thiouridine; 50% of uridines with 4-thiouridine; 100% of uridines with 4-thiouridine; 25% of uridines with 2-thiouridine (s2U) and 25% of cytidines with 5-methylcytidine (m5C); 50% of uridines with 2-thiouridine (s2U); 100% of uridines with pseudouridine (Ψ); 100% of uridines with pseudouridine (Ψ) and 100% of cytidines with 5-methylcytidine (5mC); 25% of uridines with 5-methoxyuridine (5moU) and 50% of cytidines with 5-methylcytidine (5mC); 25% of uridines with 5-methoxyuridine (5moU) and 100% of cytidines with 5-methylcytidine (5mC); 100% of uridines with 5-methoxyuridine (5moU); 100% of uridines with 5-methoxyuridine (5moU) and 100% of cytidines with 5-methylcytidine (5mC); 100% of uridines with N1-methylpseudouridine (1mΨ); 100% of uridines with N1-methylpseudouridine (1mΨ) and 100% of cytidines with 5-methylcytidine (5mC); or 100% of uridines replaced with 1-ethyl-pseudouridine.

[0070] The present disclosure also provides a vector or set of vectors comprising a polynucleotide disclosed herein or a complement thereof. Also provided is a method for making a polynucleotide disclosed herein or a complement thereof, comprising chemically synthesizing said polynucleotide. Also provided is a method for producing a protein encoded a polynucleotide disclosed herein, wherein the expression is conducted using an *in vitro* translation system. Also provided is a cell comprising any polynucleotide disclosed or a complement thereof, or the vector or set of vectors

disclosed herein. In some aspects, the cell is an autologous cell or a heterologous cell. Also provided is a pharmaceutical composition comprising (i) a polynucleotide disclosed herein or a complement thereof, (ii) a vector or set of vectors disclosed herein, or (iii) a cell disclosed herein, and a pharmaceutically acceptable vehicle or excipient.

[0071] Also provided is a method of expressing a polypeptide comprising contacting an effective amount of (i) a polynucleotide disclosed herein or a complement thereof or (ii) a vector or set of vectors disclosed herein in a cell, wherein the polypeptide encoded by the polynucleotide is expressed. In some aspects, the polypeptide is expressed in vitro. In some aspects, the polypeptide is expressed in vivo. Also provided is a method to treat a disease or condition in a subject in need thereof comprising administering a therapeutically effective amount of (i) a polynucleotide disclosed herein or a complement thereof, (ii) a vector or set of vectors disclosed herein, (iii) a cell disclosed herein, (iv) a pharmaceutical composition disclosed, or (v) a combination thereof.

[0072] The present disclosure also provides a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein

(a) the nucleotide sequence encodes TVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQES VTEQDSKDSTYSLSX[1]TLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNR (SEQ ID NO: 2200), wherein X^1 is selected from N and S, and wherein the nucleotide sequence encodes a kappa light chain constant domain of an antibody or a fragment thereof; or,

(b) the nucleotide sequence encodes PKAAPSVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTT PSKQSNNKYAASSYLSLTPEQWKSHX[2]SYSCQVTHEGSTVEKTVAPX[3]ECS (SEQ ID NO: 2201), wherein X^2 is selected from R and K, and X^3 is selected from T and A, and wherein the nucleotide sequence encodes a lambda light chain constant domain of an antibody or a fragment thereof.

[0073] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein

(a) the nucleotide sequence encodes SX^{4}GPSVX^{5}PLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGVHTFPAVLQSSG LYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKX^{6}X^7 (SEQ ID NO: 2202) wherein X^4 is an optional ASTK sequence, X^5 is selected from F and L, X^6 is selected from K and R, and X^7 is selected from V and A, and wherein the nucleotide sequence encodes a CH1 domain of an IgG1 antibody or a fragment thereof; and/or,

(b) the nucleotide sequence encodes APEX^{8}X^{9}GX^{10}PSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNX^{11}YVDGV EVHNAKTKPREEQYX^{12}STYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK TISKAK (SEQ ID NO: 2203) wherein X^8 is selected from L and A, X^9 is selected from L and A, X^{10} is selected from G and A, and X^{11} is selected from V and W, and X^{12} is selected from N and A, and wherein the nucleotide sequence encodes a CH2 domain of an IgG1 antibody or a fragment thereof; and/or,

(c) the nucleotide sequence encodes GQPREPQVYTLPPSRX^{13}EX^{14}TKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYK TTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG (SEQ ID NO: 2204) wherein X^{13} is selected from E and D, and X^{14} is selected from M and L, and wherein the nucleotide sequence encodes a CH3 domain of an IgG1 antibody or a fragment thereof; and/or,

(d) the nucleotide sequence encodes SASTKGPSVFPLAPCSRSTSESTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPA VLQSSGLYSLSSVVTVX^{15}SSNFGTQTYTCNVDHKPSNTKVDKTV (SEQ ID NO: 2205) wherein X^{15} is selected from P and T, and wherein the nucleotide sequence encodes a CH1 domain of an IgG2 antibody or a fragment thereof; and/or,

(e) the nucleotide sequence encodes APPVAGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFNWYVDGX^{16}EV HNAKTKPREEQFNSTFRVVSVLTVVHQDWLNGKEYKCKVSNKGLPX^{17}X^{18}IEKTI SKTK (SEQ ID NO: 2206) wherein X^{16} is selected from V and M, X^{17} is selected from A and S; and X^{18} is selected from P and S, and wherein the nucleotide sequence encodes a CH2 domain of an IgG2 antibody or a fragment thereof; and/or,

(f) the nucleotide sequence encodes GQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTP PMLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG

(SEQ ID NO: 2196), wherein the nucleotide sequence encodes a CH3 domain of an IgG2 antibody or a fragment thereof; and/or,

(g) the nucleotide sequence encodes SASTKGPSVFPLAPCSRSTSESTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPA VLQSSGLYSLSSVVTVPSSSLGTKTYTCNVDHKPSNTKVDKRV (SEQ ID NO: 2197), wherein the nucleotide sequence encodes a CH1 domain of an IgG4 antibody or a fragment thereof; and/or,

(h) the nucleotide sequence encodes APEFLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSQEDPEVQFNWYVDGVEVH NAKTKPREEQFNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKGLPSSIEKTISKA K (SEQ ID NO: 2198), wherein the nucleotide sequence encodes a CH2 domain of an IgG4 antibody or a fragment thereof; and/or,

(i) the nucleotide sequence encodes GQPREPQVYTLPPSQEEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTP PVLDSDGSFFLYSRLTVDKSRWQEGNVFSCSVMHEALHNHYTQKSLSLSLG (SEQ ID NO: 2199), wherein the nucleotide sequence encodes a CH3 domain of an IgG4 antibody or a fragment thereof.

[0074] Also provided is polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein

(a) the nucleotide sequence encodes $X^1X^2X^3LTQX^4X^5X^6VSX^7X^8X^9GX^{10}X^{11}X^{12}X^{13}X^{14}X^{15}C$ (SEQ ID NO: 2235) wherein X^1 is selected from Q, D, E and S; X^2 is selected from S, I, A, and Y; X^3 is selected from V, Q, A, and E; X^4 is selected from P and D; X^5 is selected from P, N, and A; X^6 is selected from S and A; X^7 is selected from G, T, A, and V; X^8 is selected from A and S; X^9 is selected from P and L; X^{10} is selected from Q, K, and S; X^{11} is selected from R, K, T, and S; X^{12} is selected from V, I, and A; X^{13} is selected from T, K, and R; X^{14} is selected from I and L; and, X^{15} is selected from S at T, wherein the nucleotide sequence encodes the first framework region (FW1) of a lambda light chain variable domain; and/or,

(b) the nucleotide sequence encodes $X^1X^2QX^3TQX^4X^5SX^6X^7SASX^8CDRVTX^9X^{10}C$ (SEQ ID NO: 2239) wherein X^1 is selected from D and A; X^2 is selected from I and V; X^3 is selected from M, L, and V; X^4

is selected from S and F; X^5 is selected from P and T; X^6 is selected from S and T; X^7 is selected from L and V; X^8 is selected from V, I, and A; X^9 is selected from I and M; and, X^{10} is selected from T and S, wherein the nucleotide sequence encodes the first framework region (FW1) of a kappa light chain variable domain; and/or,

(c) the nucleotide sequence encodes $DX^1X^2X^3TQX^4PX^5SX^6X^7X^8X^9X^{10}GX^{11}X^{12}X^{13}X^{14}X^{15}X^{16}C$ (SEQ ID NO: 2243) wherein X^1 is selected from I and V; X^2 is selected from V, L, and Q; X^3 is selected from M and L; X^4 is selected from S and T; X^5 is selected from L and D; X^6 is selected from L and V; X^7 is selected from P, S and A; X^8 is selected from V and M; X^9 is selected from T and S; X^{10} is selected from P and L; X^{11} is selected from E and Q; X^{12} is selected from P and R; X^{13} is selected from A and V; X^{14} is selected from S and T; X^{15} is selected from I, M, and L; and X^{16} is selected from S and N, wherein the nucleotide sequence encodes the first framework region (FW1) of a kappa light chain variable domain; and/or,

(d) the nucleotide sequence encodes $X^1X^2VX^3TQSPX^4TLSX^5SPGERATLSC$ (SEQ ID NO: 2247) wherein X^1 is selected from E and D; X^2 is selected from I and T; X^3 is selected from L and M; X^4 is selected from G and A; and, X^5 is selected from L and V, wherein the nucleotide sequence encodes the first framework region (FW1) of a kappa light chain variable domain; and/or,

(e) the nucleotide sequence encodes $X^1X^2X^3X^4X^5X^6SGGX^7X^8X^9X^{10}X^{11}GX^{12}SX^{13}X^{14}LX^{15}C$ (SEQ ID NO: 2251) wherein X^1 is selected from E, D, and Q; X^2 is selected from V and A; X^3 is selected from Q, E, and K; X^4 is selected from L and V; X^5 is selected from V and L; X^6 is selected from E and Q; X^7 is selected from G, K, and D; X^8 is selected from L and V; X^9 is selected from V, L, and E; X^{10} is selected from Q, R and K; X^{11} is selected from P, S, and L; X^{12} is selected from G and R; X^{13} is selected from L and R; X^{14} is selected from R and K; and, X^{15} is selected from S and D, wherein the nucleotide sequence encodes the first framework region (FW1) of a heavy chain variable domain; and/or,

(f) the nucleotide sequence encodes $X^1X^2QLX^3QX^4GX^5X^6X^7X^8X^9X^{10}GX^{11}X^{12}X^{13}X^{14}X^{15}SC$ (SEQ ID NO: 2255) wherein X^1 is selected from Q and E; X^2 is selected from V and I; X^3 is selected from V and Q; X^4 is selected from S and P; X^5 is selected from A, S, V, P, T, and G; X^6 is selected from E, G and V; X^7 is selected from V and L; X^8 is selected from K, V, E, and A; X^9 is selected from K, R and Q; X^{10} is selected from P and S; X^{11} is selected from A, E, S, T, and R; X^{12}

is selected from S and T; X^{13} is selected from V and L; X^{14} is selected from K and R; and, X^{15} is selected from V, I, L, and M, wherein the nucleotide sequence encodes the first framework region (FW1) of a heavy chain variable domain; and/or,

(g) the nucleotide sequence encodes $QX^1X^2LX^3X^4X^5GX^6X^7LX^8X^9PX^{10}X^{11}TLX^{12}LTC$ (SEQ ID NO: 2259) wherein X^1 is selected from V and L; X^2 is selected from Q and T; X^3 is selected from Q and R; X^4 is selected from E and Q; X^5 is selected from S and W; X^6 is selected from P and A; X^7 is selected from G and A; X^8 is selected from V and L; X^9 is selected from K and R; X^{10} is selected from S and T; X^{11} is selected from Q and E; and, X^{12} is selected from S and T, wherein the nucleotide sequence encodes the first framework region (FW1) of a heavy chain variable domain; and/or,

(h) the nucleotide sequence encodes $WYQX^1X^2X^3GX^4X^5PX^6X^7X^8I$ (SEQ ID NO: 2236) wherein X^1 is selected from Q and L; X^2 is selected from L,Y, H, and K; X^3 is selected from P and E; X^4 is selected from T, R, K, and Q; X^5 is selected from A and S; X^6 is selected from K, T, V and I; X^7 is selected from L and T; and X^8 is selected from L, M, and V, wherein the nucleotide sequence encodes the second framework region (FW2) of a lambda light chain variable domain; and/or,

(i) the nucleotide sequence encodes $WX^1RQX^2PX^3KX^4LX^5X^6X^7X^8$ (SEQ ID NO: 2252) wherein X^1 is selected from V, I, and F; X^2 is selected from A, S and T; X^3 is selected from G and E; X^4 is selected from G and R; X^5 is selected from E and D; X^6 is selected from W and L; X^7 is selected from V and I; and, X^8 is selected from A, S, and G, wherein the nucleotide sequence encodes the second framework region (FW2) of a heavy chain variable domain; and/or,

(j) the nucleotide sequence encodes $WX^1X^2QX^3X^4GX^5X^6LX^7WX^8G$ (SEQ ID NO: 2256) wherein X^1 is selected from V and I; X^2 is selected from R and K; X^3 is selected from A, M, N, R, K, T, and S; X^4 is selected from P, T, and H; X^5 is selected from Q, K, and R; X^6 is selected from G, R and S; X^7 is selected from E, D, K, Q, and A; and, X^8 is selected from M, I, and V, wherein the nucleotide sequence encodes the second framework region (FW2) of a heavy chain variable domain; and/or,

(k) the nucleotide sequence encodes $WX^1RX^2X^3X^4X^5X^6X^7LX^8WX^9X^{10}$ (SEQ ID NO: 2260) wherein X^1 is selected from I and V; X^2 is selected from Q and H; X^3 is selected from L, P, S, and H; X^4 is selected from P and S; X^5 is selected from G and E; X^6 is selected from K and R; X^7 is selected from G and A; X^8 is selected from E and Q; X^9 is

selected from I and L; and, X^{10} is selected from G and A, wherein the nucleotide sequence encodes the second framework region (FW2) of a heavy chain variable domain; and/or,

(l) the nucleotide sequence encodes $WX^1X^2X^3X^4PX^5KX^6X^7X^8X^9X^{10}IX^{11}$ (SEQ ID NO: 2240) wherein X^1 is selected from Y and F; X^2 is selected from Q and L; X^3 is selected from Q and H; X^4 is selected from K and I; X^5 is selected from G and E; X^6 is selected from A and V; X^7 is selected from P and V; X^8 is selected from K and Q; X^9 is selected from L, T, S, R, P, and V; X^{10} is selected from L and W; and, X^{11} is selected from Y and S, wherein the nucleotide sequence encodes the second framework region (FW2) of a kappa light chain variable domain; and/or,

(m) the nucleotide sequence encodes $WX^1X^2QX^3X^4GQX^5PX^6X^7LIX^8$ (SEQ ID NO: 2244) wherein X^1 is selected from Y, F, and W; X^2 is selected from L and Q; X^3 is selected from K and R; X^4 is selected from P and S; X^5 is selected from S and P; X^6 is selected from Q, K, R, and N; X^7 is selected from L and R; and, X^8 is selected from Y and W, wherein the nucleotide sequence encodes the second framework region (FW2) of a kappa light chain variable domain; and/or,

(n) the nucleotide sequence encodes $WX^1X^2QX^3PGQAPRX^4LIX^5$ (SEQ ID NO: 2248) wherein X^1 is selected from Y and F; X^2 is selected from Q and R; X^3 is selected from K and R; X^4 is selected from L and P; and X^5 is selected from Y, R, and K, wherein the nucleotide sequence encodes the second framework region (FW2) of a kappa light chain variable domain; and/or,

(o) the nucleotide sequence encodes $RFSGSX^1SX^2X^3X^4AX^5LX^6IX^7X^8X^9X^{10}X^{11}X^{12}DEAX^{13}YX^{14}C$ (SEQ ID NO: 2237) wherein X^1 is selected from K, N, S, and I; X^2 is selected from G and S; X^3 is selected from T and N; X^4 is selected from S and T; X^5 is selected from S, T, and F; X^6 is selected from A, T, and G; X^7 is selected from T, H, and S; X^8 is selected from G, N, and R; X^9 is selected from L, V, and A; X^{10} is selected from Q, E, and A; X^{11} is selected from A, T, and I; X^{12} is selected from E and G; X^{13} is selected from D and I; and, X^{14} is selected from Y and F, wherein the nucleotide sequence encodes the third framework region (FW3) of a lambda light chain variable domain; and/or,

(p) the nucleotide sequence encodes $RFSGSX^1SGX^2X^3X^4X^5X^6TISSLX^7X^8X^9DX^{10}AX^{11}YX^{12}C$ (SEQ ID NO: 2241) wherein X^1 is selected from G and R; X^2 is selected from T and Q; X^3 is selected from D, E, and

Y; X^4 is selected from F and Y; X^5 is selected from T and S; X^6 is selected from L and F; X^7 is selected from Q and E; X^8 is selected from P, Q, A, and S; X^9 is selected from E and D; X^{10} is selected from F, I, S, L, V, and T; X^{11} is selected from T, S, and V; and, X^{12} is selected from Y and F, wherein the nucleotide sequence encodes the third framework region (FW3) of a kappa light chain variable domain; and/or,

(q) the nucleotide sequence encodes RFSGSGSX^1TX^2FTLX^3IS$X^4X^5X^6$AX^7DV$X^8X^9X^{10}X^{11}$C (SEQ ID NO: 2245) wherein X^1 is selected from G and A; X^2 is selected from D and A; X^3 is selected from K, R, and T; X^4 is selected from R and S; X^5 is selected from V and L; X^6 is selected from E and Q; X^7 is selected from E and Q; X^8 is selected from G and A; X^9 is selected from V, D, and F; X^{10} is selected from Y and W; and, X^{11} is selected from Y, F, and W, wherein the nucleotide sequence encodes the third framework region (FW3) of a kappa light chain variable domain; and/or,

(r) the nucleotide sequence encodes RFSGSGSGTX^1X^2TLTISX^3LX^4X^5EDFAX^6X^7YC (SEQ ID NO: 2249) wherein X^1 is selected from D and E; X^2 is selected from F and S; X^3 is selected from R and S; X^4 is selected from E and Q; X^5 is selected from P and S; X^6 is selected from V and T; and, X^7 is selected from Y and F, wherein the nucleotide sequence encodes the third framework region (FW3) of a kappa light chain variable domain; and/or,

(s) the nucleotide sequence encodes $X^1X^2X^3X^4$SX^5D$X^6X^7X^8X^9X^{10}X^{11}X^{12}LX^{13}X^{14}X^{15}X^{16}LX^{17}X^{18}EDTX^{19}X^{20}X^{21}X^{22}$C (SEQ ID NO: 2253) wherein X^1 is selected from R and K; X^2 is selected from F and V; X^3 is selected from T, I, and A; X^4 is selected from L and I; X^5 is selected from V, R, L, and A; X^6 is selected from R, N, T, D, K, and S; X^7 is selected from S, A and V; X^8 is selected from K, R, and E; X^9 is selected from N, S, R, H, and T; X^{10} is selected from T and S; X^{11} is selected from L, A, and F; X^{12} is selected from Y and F; X^{13} is selected from Q and E; X^{14} is selected from M and V; X^{15} is selected from N, D, and S; X^{16} is selected from S, G, and I; X^{17} is selected from R and K; X^{18} is selected from A, S, D, V, and P; X^{19} is selected from A and G; X^{20} is selected from V, M, and L; X^{21} is selected from Y and F; and, X^{22} is selected from Y and F, wherein the nucleotide sequence encodes the third framework region (FW3) of a heavy chain variable domain; and/or,

(t) the nucleotide sequence encodes $X^1X^2X^3X^4X^5X^6X^7X^8$SX^9X^{10}T$X^{11}X^{12}X^{13}X^{14}X^{15}X^{16}X^{17}LX^{18}X^{19}X^{20}DX^{21}X^{22}X^{23}YX^{24}$C

(SEQ ID NO: 2257) wherein X^1 is selected from R, Q, and K; X^2 is selected from V, I, F, G, and A; X^3 is selected from T, A, and K; X^4 is selected from M, I, L, and F; X^5 is selected from T and S; X^6 is selected from T, A, R, V, S, E, and L; X^7 is selected from D, E, and N; X^8 is selected from T, K, Q, S, P, R, I, N, and E; X^9 is selected from T, K, S, A, I, and V; X^{10} is selected from S, N, D, and T; X^{11} is selected from A, V, and T; X^{12} is selected from Y, and F; X^{13} is selected from M and L; X^{14} is selected from E, Q, and D; X^{15} is selected from L, M, W, and I; X^{16} is selected from R, S, D, L, K, T, and N; X^{17} is selected from S and R; X^{18} is selected from R, K, Q, and T; X^{19} is selected from S, H, F, A, and P; X^{20} is selected from D, E, and S; X^{21} is selected from T and S; X^{22} is selected from A and G; X^{23} is selected from V, F, T, and M; and, X^{24} is selected from Y, F, and L, wherein the nucleotide sequence encodes the third framework region (FW3) of a heavy chain variable domain; and/or,

(u) the nucleotide sequence encodes $RX^1X^2X^3X^4X^5DX^6SX^7X^8QX^9X^{10}LX^{11}X^{12}X^{13}X^{14}X^{15}X^{16}X^{17}X^{18}DTAX^{19}X^{20}X^{21}C$ (SEQ ID NO: 2261) wherein X^1 is selected from V and L; X^2 is selected from T and S; X^3 is selected from I and M; X^4 is selected from S and L; X^5 is selected from V, R, and K; X^6 is selected from T and K; X^7 is selected from K and R; X^8 is selected from K and N; X^9 is selected from F and V; X^{10} is selected from S and V; X^{11} is selected from R, T, K, and M; X^{12} is selected from L, I, M, and V; X^{13} is selected from S, T, and N; X^{14} is selected from S and N; X^{15} is selected from V and M; X^{16} is selected from T and D; X^{17} is selected from A and P; X^{18} is selected from A and V; X^{19} is selected from V and T; X^{20} is selected from Y and W; and, X^{21} is selected from Y, F and W, wherein the nucleotide sequence encodes the third framework region (FW3) of a heavy chain variable domain; and/or,

(v) the nucleotide sequence encodes $FGX^1GTX^2X^3TVL$ (SEQ ID NO:2238) wherein X^1 is selected from G and T; X^2 is selected from K and Q; and X^3 is selected from L and V, wherein the nucleotide sequence encodes the fourth framework region (FW4) of a lambda light chain variable domain; and/or,

(w) the nucleotide sequence encodes $X^1GX^2GTX^3X^4X^5X^6X^7$ (SEQ ID NO: 2242) wherein X^1 is selected from F and L; X^2 is selected from Q, G, and S; X^3 is selected from K and R; X^4 is selected from V and L; X^5 is selected from E, D, and Q; X^6 is selected from I and V; and, X^7 is selected from K and T, wherein the nucleotide sequence encodes the fourth framework region (FW4) of a kappa light chain variable domain; and/or,

(x) the nucleotide sequence encodes $FGX^1GTX^2X^3X^4X^5K$ (SEQ ID NO: 2246) wherein X^1 is selected from Q, A, P, and G; X^2 is selected from K and R; X^3 is selected from V and L; X^4 is selected from E and Q; and X^5 is selected from I and L, wherein the nucleotide sequence encodes the fourth framework region (FW4) of a kappa light chain variable domain; and/or,

(y) the nucleotide sequence encodes $FX^1X^2GTX^3X^4X^5IK$ (SEQ ID NO: 2250) wherein X^1 is selected from G and C; X^2 is selected from Q, G, and P; X^3 is selected from K and R; X^4 is selected from V, L, and A; and, X^5 is selected from E and D, wherein the nucleotide sequence encodes the fourth framework region (FW4) of a kappa light chain variable domain; and/or,

(z) the nucleotide sequence encodes $WGX^1GX^2X^3VTVS$ (SEQ ID NO: 2254) wherein X^1 is selected from Q, R, and K; X^2 is selected from T, I and A; and, X^3 is selected from L, S, T, M, and P, wherein the nucleotide sequence encodes the fourth framework region (FW4) of a heavy chain variable domain; and/or,

(aa) the nucleotide sequence encodes $WGX^1GTX^2X^3TVS$ (SEQ ID NO: 2258) wherein X^1 is selected from R, Q, K, A and S; X^2 is selected from L, M, T, Q, and P; and, X^3 is selected from V and L, wherein the nucleotide sequence encodes the fourth framework region (FW4) of a heavy chain variable domain; and/or,

(ab) the nucleotide sequence encodes $WX^1X^2GX^3X^4VTVS$ (SEQ ID NO: 2262) wherein X^1 is selected from G and D; X^2 is selected from Q and R; X^3 is selected from T and S; and X^4 is selected from T, L, and M, wherein the nucleotide sequence encodes the fourth framework region (FW4) of a heavy chain variable domain.

[0075] In some aspects, the polynucleotide further comprises a nucleotide sequence which encodes a linker sequence of formula $(Gly_xSer)_y$, wherein x and y are integers between 1 and 100. In some aspects, the linker sequence is interposed between a VH domain and a VL domain. In some aspects, the polynucleotide encodes an scFv. In some aspects, the polynucleotide encodes a therapeutic antibody or an antigen-binding fragment thereof. In some aspects, the therapeutic antibody is selected from the group consisting of abagovomab, abciximab, adalimumab, alemtuzumab, alirocumab, amatuximab, anrukinzumab, arcitumomab, basiliximab, bavituximab, benralizumab, bevacizumab, bezlotoxumab, bimagrumab, bococizumab, brentuximab, briakinumab, brodalumab, canakinumab, cantuzumab, carlumab, cetuximab, cixutumumab, clivatuzumab, conatumumab, crenezumab, dacetuzumab, daclizumab, dalotuzumab,

denosumab, drozitumab, dupilumab, dusigitumab, eculizumab, elotuzumab, enokizumab, epratuzumab, etaracizumab, evolocumab, farletuzumab, fasinumab, fezakinumab, ficlatuzumab, figitumumab, fresolimumab, fulranumab, ganitumab, gantenerumab, gevokizumab, girentuximab, glembatumumab, ibalizumab, ibritumomab, icrucumab, inotuzumab, intetumumab, itolizumab, ixekizumab, lebrikizumab, lorvotuzumab, mavrilimumab, mepolizumab, milatuzumab, mogamulizumab, motavizumab, naptumomab, necitumumab, nivolumab, obinutuzumab, ocrelizumab, olaratumab, omalizumab, otelixizumab, oxelumab, pateclizumab, pembrolizumab, pertuzumab, ponezumab, ramucirumab, rilotumumab, rituximab, robatumumab, romosozumab, rontalizumab, samalizumab, sarilumab, secukinumab, sifalimumab, siltuximab, sirukumab, solanezumab, tabalumab, tanezumab, tenatumomab, teplizumab, tigatuzumab, tildrakizumab, tocilizumab, tositumomab, tralokinumab, trastuzumab, urelumab, ustekinumab, vedolizumab, and veltuzumab, and functional fragments thereof.

[0076] In some aspects, the nucleotide sequence of the polynucleotide is selected from SEQ ID NOS: 1979-2188, and subsequences thereof. In some aspects, the nucleotide sequence is codon-optimized according to any of the methods disclosed in the present application or any other codon optimization methods known in the art. In some aspects, the nucleotide sequence is codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof). In some aspects, the polynucleotide is an mRNA. In some aspects, the mRNA is synthetic.

[0077] In some aspects, (i) at least one uridine of the polynucleotide has been replaced with 2- pseudouridine, 5-methoxyuridine, 1-ethyl-pseudouridine, 2-thiouridine, 4-thiouridine, N1-methylpseudouridine, 5-aza-uridine, 2-thio-5-aza-uridine, 4-thio-pseudouridine, 2-thio-pseudouridine, 5-hydroxyuridine, 4-methoxy-pseudouridine, 4-methoxy-2-thio-pseudouridine, 3-methyluridine, 5-carboxymethyl-uridine, 1-carboxymethyl-pseudouridine, 5-propynyl-uridine, 1-propynyl-pseudouridine, 2-methoxy-4-thio-uridine, 5-taurinomethyluridine, 1-taurinomethyl-pseudouridine, 5-taurinomethyl-2-thio-uridine, 1-taurinomethyl-4-thio-uridine, 5-methyl-uridine, 2-methoxyuridine, 1-methyl-pseudouridine, 4-thio-1-methyl-pseudouridine, 2-thio-1-methyl-pseudouridine, 1-methyl-1-deaza-pseudouridine, 2-thio-1-methyl-1-deaza-pseudouridine, or 2-thio-dihydrouridine; and/or, (ii) at least one adenosine of the polynucleotide has been replaced with 2-aminopurine, 2,6-diaminopurine, 7-deaza-

adenine, 7-deaza-8-aza-adenine, 7-deaza-2-aminopurine, 7-deaza-8-aza-2-aminopurine, 7-deaza-2,6-diaminopurine, 7-deaza-8-aza-2,6-diaminopurine, 1-methyladenosine, N6-methyladenosine, N6-isopentenyladenosine, N6-(cis-hydroxyisopentenyl)adenosine, 2-methylthio-N6-(cis-hydroxyisopentenyl)adenosine, N6-glycinylcarbamoyladenosine, N6-threonylcarbamoyladenosine, 2-methylthio-N6-threonyl carbamoyladenosine, N6,N6-dimethyladenosine, or 7-methyladenine; and/or, (iii) at least one guanosine of the polynucleotide has been replaced with inosine, 1-methyl-inosine, wyosine, wybutosine, 7-deaza-guanosine, 7-deaza-8-aza-guanosine, 6-thio-guanosine, 6-thio-7-deaza-guanosine, 6-thio-7-deaza-8-aza-guanosine, 7-methyl-guanosine, 6-thio-7-methyl-guanosine, 7-methylinosine, 6-methoxy-guanosine, 1-methylguanosine, N2-methylguanosine, N2,N2-dimethylguanosine, 8-oxo-guanosine, 7-methyl-8-oxo-guanosine, or 1-methyl-6-thio-guanosine; and/or, (iv) at least one cytidine of the polynucleotide has been replaced with 5-methylcytidine, 5-aza-cytidine, pseudoisocytidine, 3-methyl-cytidine, N4-acetylcytidine, 5-formylcytidine, N4-methylcytidine, 5-hydroxymethylcytidine, 1-methyl-pseudoisocytidine, pyrrolo-cytidine, pyrrolo-pseudoisocytidine, 2-thio-cytidine, 2-thio-5-methyl-cytidine, 4-thio-pseudoisocytidine, 4-thio-1-methyl-pseudoisocytidine, 4-thio-1-methyl-1-deaza-pseudoisocytidine, 1-methyl-1-deaza-pseudoisocytidine, zebularine, 5-aza-zebularine, 5-methyl-zebularine, 5-aza-2-thio-zebularine, 2-thio-zebularine, 2-methoxy-cytidine, 2-methoxy-5-methyl-cytidine, or 4-methoxy-pseudoisocytidine.

[0078] Also provided is a method to treat a disease or condition in a subject in need thereof comprising administering a therapeutically effective amount of (i) a polynucleotide of the present invention or a complement thereof, (ii) a vector or set of vectors comprising said polynucleotide, (iii) a cell comprising said vector or set of vectors, (iv) a pharmaceutical composition comprising said polynucleotide, vector or set of vectors, or cell, or (v) a combination thereof.

BRIEF DESCRIPTION OF THE DRAWINGS/FIGURES

[0079] FIG. 1 shows a Position Specific Scoring Matrix (PSSM) defining an immunoglobulin constant domain (IgC) in general. The PSSM corresponds to conserved domain CD00098 available at the NCBI CDD database. See Marchler-Bauer et al. (2015), "CDD: NCBI's conserved domain database", Nucleic Acids Res. 43(Database issue):D222-6.

[0080] FIG. 2 shows a PSSM defining an immunoglobulin light chain constant domain (CL), corresponding to conserved domain CD07699 in the CDD database.

[0081] FIG. 3 shows a PSSM defining the first constant domain of the heavy chain of an immunoglobulin (CH1), corresponding to conserved domain CD04985 in the CDD database.

[0082] FIG. 4 shows a PSSM defining the second constant domain of the heavy chain of an immunoglobulin (CH2), corresponding to conserved domain CD04986 in the CDD database.

[0083] FIG. 5 shows a PSSM defining the third constant domain of the heavy chain of an immunoglobulin (CH3), corresponding to conserved domain CD07696 in the CDD database.

[0084] FIG. 6 shows a PSSM defining an immunoglobulin variable domain in general, corresponding to conserved domain CD00099 in the CDD database.

[0085] FIG. 7 shows a PSSM defining an immunoglobulin heavy chain variable domain (VH), corresponding to conserved domain CD04981 in the CDD database.

[0086] FIG. 8 shows a PSSM defining an immunoglobulin light chain variable domain, kappa type (VL kappa), corresponding to conserved domain CD04980 in the CDD database.

[0087] FIG. 9 shows a PSSM defining an immunoglobulin light chain variable domain, lambda type (VL lambda), corresponding to conserved domain CD4984 in the CDD database.

[0088] FIG. 10 shows a multiple sequence alignment of the light chains of 105 therapeutic antibodies conducted using the program ClustalX and default parameters. The location of the boundary of the variable domain (VL) and constant domain (CL), as well as the location of the boundaries of the framework regions (FW1, FW2, FW3, and FW4) and complementarity determining regions (CDR1, CDR and CDR2) are indicated.

[0089] FIG. 11 shows a multiple sequence alignment of the heavy chains of 105 therapeutic antibodies conducted using the program ClustalX and default parameters. The location of the boundaries of the variable domain (VH) and the constant domains (CH1, CH2, CH3), as well as the location of the boundaries of the framework regions (FW1, FW2, FW3, and FW4) and complementarity determining regions (CDR1, CDR and CDR2) are indicated.

[0090] FIG. 12 is a schematic representation the domain organization of an IgG antibody, in particular showing the location of variable regions (VL, VL), constant regions (CL, CH1 CH2, CH3), framework regions (FR), complementarity determining regions (CDR), Hinges, as well as the Fab region and Fc region.

[0091] FIG. 13 is an schematic representation of a typical immunoglobulin fold, showing the location of beta strands (indicated by arrows) and loops connecting the beta strands. The location of the CDRs in loop regions is indicated, as well as the location of the framework regions (FW1 to FW4). Each framework region comprises the labeled beta strands plus their connecting loops.

[0092] FIG. 14 presents a variety of antibody-derived constructs known in the art, wherein each construct comprises one or more domains having an immunoglobulin fold (e.g., VH, VL, CL, CH1, CH2, or CH3 domains).

DETAILED DESCRIPTION

[0093] The present disclosure relates to polynucleotides comprising codon-optimized nucleotide sequences encoding an antibody, a functional fragment thereof (e.g., an antigen-binding fragment thereof or an Fc fragment), a variant thereof, or a combination thereof. These compositions (e.g., mRNAs) can be administered to a subject in need thereof to facilitate *in vivo* expression and assembly of a therapeutic antibody. Each of the nucleotide sequences disclosed herein is not a wild type nucleotide sequence encoding a therapeutic antibody known in the art.

[0094] In order that the present disclosure can be more readily understood, certain terms are first defined. Additional definitions are set forth throughout the detailed description.

I. Definitions

[0095] The headings provided herein are not limitations of the various aspects or aspects of the disclosure, which can be defined by reference to the specification as a whole. Accordingly, the terms defined immediately below are more fully defined by reference to the specification in its entirety. Before describing the present invention in detail, it is to be understood that this invention is not limited to specific compositions or process steps, as such can vary.

[0096] In this specification and the appended claims, the singular forms "a", "an" and "the" include plural referents unless the context clearly dictates otherwise. The terms "a" (or "an"), as well as the terms "one or more," and "at least one" can be used interchangeably herein. In certain aspects, the term "a" or "an" means "single." In other aspects, the term "a" or "an" includes "two or more" or "multiple."

[0097] Furthermore, "and/or" where used herein is to be taken as specific disclosure of each of the two specified features or components with or without the other. Thus, the term "and/or" as used in a phrase such as "A and/or B" herein is intended to include "A and B," "A or B," "A" (alone), and "B" (alone). Likewise, the term "and/or" as used in a phrase such as "A, B, and/or C" is intended to encompass each of the following aspects: A, B, and C; A, B, or C; A or C; A or B; B or C; A and C; A and B; B and C; A (alone); B (alone); and C (alone).

[0098] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure is related. For example, the Concise Dictionary of Biomedicine and Molecular Biology, Juo, Pei-Show, 2nd ed., 2002, CRC Press; The Dictionary of Cell and Molecular Biology, 3rd ed., 1999, Academic Press; and the Oxford Dictionary Of Biochemistry And Molecular Biology, Revised, 2000, Oxford University Press, provide one of skill with a general dictionary of many of the terms used in this disclosure.

[0099] Wherever aspects are described herein with the language "comprising," otherwise analogous aspects described in terms of "consisting of" and/or "consisting essentially of" are also provided.

[0100] The term "about" as used in connection with a numerical value throughout the specification and the claims denotes an interval of accuracy, familiar and acceptable to a person skilled in the art. In general, such interval of accuracy is ± 15 %.

[0101] Units, prefixes, and symbols are denoted in their Système International de Unites (SI) accepted form. Numeric ranges are inclusive of the numbers defining the range. Where a range of values is recited, it is to be understood that each intervening integer value, and each fraction thereof, between the recited upper and lower limits of that range is also specifically disclosed, along with each subrange between such values. The upper and lower limits of any range can independently be included in or excluded from the range, and each range where either, neither or both limits are included is also encompassed within the invention. Where a value is explicitly recited, it is to be

understood that values which are about the same quantity or amount as the recited value are also within the scope of the invention. Where a combination is disclosed, each subcombination of the elements of that combination is also specifically disclosed and is within the scope of the invention. Conversely, where different elements or groups of elements are individually disclosed, combinations thereof are also disclosed. Where any element of an invention is disclosed as having a plurality of alternatives, examples of that invention in which each alternative is excluded singly or in any combination with the other alternatives are also hereby disclosed; more than one element of an invention can have such exclusions, and all combinations of elements having such exclusions are hereby disclosed.

[0102] Nucleotides are referred to by their commonly accepted single-letter codes. Unless otherwise indicated, nucleic acids are written left to right in 5′ to 3′ orientation. Nucleotides are referred to herein by their commonly known one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Accordingly, A represents adenine, C represents cytosine, G represents guanine, T represents thymine, U represents uracil.

[0103] Amino acids are referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Unless otherwise indicated, amino acid sequences are written left to right in amino to carboxy orientation.

[0104] The term "polynucleotide" as used herein refers to polymers of nucleotides of any length, including ribonucleotides, deoxyribonucleotides, analogs thereof, or mixtures thereof. This term refers to the primary structure of the molecule. Thus, the term includes triple-, double- and single-stranded deoxyribonucleic acid ("DNA"), as well as triple-, double- and single-stranded ribonucleic acid ("RNA"). It also includes modified, for example by alkylation, and/or by capping, and unmodified forms of the polynucleotide. More particularly, the term "polynucleotide" includes polydeoxyribonucleotides (containing 2-deoxy-D-ribose), polyribonucleotides (containing D-ribose), including tRNA, rRNA, hRNA, siRNA and mRNA, whether spliced or unspliced, any other type of polynucleotide which is an N- or C-glycoside of a purine or pyrimidine base, and other polymers containing normucleotidic backbones, for example, polyamide (e.g., peptide nucleic acids "PNAs") and polymorpholino polymers, and other synthetic sequence-specific nucleic acid polymers providing that the polymers contain nucleobases in a

configuration which allows for base pairing and base stacking, such as is found in DNA and RNA. In particular aspects, the polynucleotide is an mRNA. In other aspect, the mRNA is a synthetic mRNA. In some aspects, the synthetic mRNA comprises at least one unnatural nucleobase. In some aspects, all nucleobases of a certain class have been replaced with unnatural nucleobases (e.g., all uridines in a polynucleotide disclosed herein can be replaced with a unnatural nucleobase, e.g., 5-methoxyuridine). In some aspects, the polynucleotide (e.g., a synthetic RNA or a synthetic DNA) comprises only natural nucleobases, i.e., A,C, T and U in the case of a synthetic DNA, or A, C, T, and U in the case of a synthetic RNA.

[0105] Standard A-T and G-C base pairs form under conditions which allow the formation of hydrogen bonds between the N3-H and C4-oxy of thymidine and the N1 and C6-NH2, respectively, of adenosine and between the C2-oxy, N3 and C4-NH2, of cytidine and the C2-NH2, N′—H and C6-oxy, respectively, of guanosine. Thus, for example, guanosine (2-amino-6-oxy-9-β-D-ribofuranosyl-purine) can be modified to form isoguanosine (2-oxy-6-amino-9-β-D-ribofuranosyl-purine). Such modification results in a nucleoside base which will no longer effectively form a standard base pair with cytosine. However, modification of cytosine (1-β-D-ribofuranosyl-2-oxy-4-amino-pyrimidine) to form isocytosine (1-β-D-ribofuranosyl-2-amino-4-oxy-pyrimidine-) results in a modified nucleotide which will not effectively base pair with guanosine but will form a base pair with isoguanosine (U.S. Pat. No. 5,681,702 to Collins et al., hereby incorporated by reference in its entirety). Isocytosine is available from Sigma Chemical Co. (St. Louis, Mo.); isocytidine can be prepared by the method described by Switzer et al. (1993) Biochemistry 32:10489-10496 and references cited therein; 2′-deoxy-5-methyl-isocytidine can be prepared by the method of Tor et al., 1993, J. Am. Chem. Soc. 115:4461-4467 and references cited therein; and isoguanine nucleotides can be prepared using the method described by Switzer et al., 1993, supra, and Mantsch et al., 1993, Biochem. 14:5593-5601, or by the method described in U.S. Pat. No. 5,780,610 to Collins et al., each of which is hereby incorporated by reference in its entirety. Other nonnatural base pairs can be synthesized by the method described in Piccirilli et al., 1990, Nature 343:33-37, hereby incorporated by reference in its entirety, for the synthesis of 2,6-diaminopyrimidine and its complement (1-methylpyrazolo-[4,3]pyrimidine-5,7-(4H,6H)-dione. Other such modified nucleotide units which form unique base pairs are known,

such as those described in Leach et al. (1992) J. Am. Chem. Soc. 114:3675-3683 and Switzer et al., supra.

[0106] The terms "nucleic acid sequence" and "nucleotide sequence" are used interchangeably and refer to a contiguous nucleic acid sequence. The sequence can be either single stranded or double stranded DNA or RNA, e.g., an mRNA.

[0107] A polynucleotide, vector, polypeptide, cell, or any composition disclosed herein which is "isolated" is a polynucleotide, vector, polypeptide, cell, or composition which is in a form not found in nature. Isolated polynucleotides, vectors, polypeptides, or compositions include those which have been purified to a degree that they are no longer in a form in which they are found in nature. In some aspects, a polynucleotide, vector, polypeptide, or composition which is isolated is substantially pure.

[0108] The terms "polypeptide," "peptide," and "protein" are used interchangeably herein to refer to polymers of amino acids of any length. The polymer can comprise modified amino acids. The terms also encompass an amino acid polymer that has been modified naturally or by intervention; for example, disulfide bond formation, glycosylation, lipidation, acetylation, phosphorylation, or any other manipulation or modification, such as conjugation with a labeling component. Also included within the definition are, for example, polypeptides containing one or more analogs of an amino acid (including, for example, unnatural amino acids such as homocysteine, ornithine, p-acetylphenylalanine, D-amino acids, and creatine), as well as other modifications known in the art.

[0109] The terms "codon substitution" or "codon replacement" refers to replacing a codon present in candidate nucleotide sequence (e.g., a DNA encoding the heavy chain or light chain of an antibody or a fragment thereof) with another codon. A codon can be substituted in a candidate nucleic acid sequence, for example, via chemical peptide synthesis or through recombinant methods known in the art. Accordingly, references to a "substitution" or "replacement" at a certain location in a nucleic acid sequence (e.g., an mRNA) or within a certain region or subsequence of a nucleic acid sequence (e.g., an mRNA) refer to the substitution of a codon at such location or region with an alternative codon. A candidate nucleic acid sequence can be a wild type nucleic sequence encoding any antibody heavy chain or light chain presented in FIGS. 10 or 11 (SEQ ID NOS: 1979 to 2188) or a functional fragment thereof (e.g., a VH, VL, CL, CH1, CH2, or CH3 domain or a combination thereof), wherein the boundaries of such fragments are provided by FIGS. 10 and 11 and method known in the art as disclosed below. A candidate nucleic

acid sequence can be codon-optimized by replacing all or part of its codons according to a substitution table map (see, .e.g., TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof)).

[0110] The skilled artisan will appreciate that the T bases in the codon maps disclosed below are present in DNA, whereas the T bases would be replaced by U bases in corresponding RNAs. For example, a codon-nucleotide sequence disclosed herein in DNA form, e.g., a vector or an in-vitro translation (IVT) template, would have its T bases transcribed as U based in its corresponding transcribed mRNA. In this respect, both codon-optimized DNA sequences (comprising T) and their corresponding RNA sequences (comprising U) are considered codon-optimized nucleotide sequence of the present invention. A skilled artisan would also understand that equivalent codon-maps can be generated by replaced one or more bases with non-natural bases. Thus, e.g., a TTC codon (DNA map) would correspond to a UUC codon (RNA map), which in turn would correspond to a ΨΨC codon (RNA map in which U has been replaced with pseudouridine).

TABLE 1: DNA codon substitution map. The map indicates possible replacement codons for each one of the 20 natural amino acids. For example, Ala can be encoded by GCC, GCG, or GCT.

Amino acid	Substitution Codon
Ala	(GCC, GCG, GCT)
Cys	(TGC, TGT)
Asp	(GAC)
Glu	(GAG, GAA)
Phe	(TTC)
Gly	(GGC, GGT, GGG)
His	(CAC)
Ile	(ATC, ATT)
Lys	(AAG)
Leu	(CTG, CTC, TTG)
Met	(ATG)
Asn	(AAC)
Pro	(CCC, CCA, CCG)
Gln	(CAG, CAA)
Arg	(CGG, AGG, CGC, CGT, AGA, CGA)

Ser	(AGC, TCC, TCT)
Thr	(ACC, ACG, ACT)
Val	(GTG, GTC, GTT)
Trp	(TGG)
Tyr	(TAC)

[0111] In one aspect, the candidate sequence can be optimized by replacing all the codons encoding a certain amino acid with only one of the alternative codons provided in TABLE 1, i.e., all the valines in the codon-optimized sequence would be encoded by GTG or GTC or GTT.

[0112] In some aspects, codons can be substituted in a candidate sequence according to any of the codon substitution maps disclosed in TABLE 2.

TABLE 2: Codon substitution maps for sequence optimization. Each one of the 16 maps presented indicates possible replacement codons for each one of the 20 natural amino acids.

		MAP1	MAP2	MAP3	MAP4	MAP5	MAP6	MAP7	MAP8
*	End	TGA	TGA	TGA	TGA	TGA	TGA	TGA	TGA
A	Ala	GCC	GCC	GCC	GCC	GCC	GCC	GCT	GCG
C	Cys	TGC	TGT	TGC	TGC	TGC	TGC	TGC	TGT
D	Asp	GAC	GAC	GAC	GAC	GAC	GAC	GAC	GAC
E	Glu	GAG	GAA	GAG	GAG	GAG	GAG	GAG	GAA
F	Phe	TTC	TTC	TTC	TTC	TTC	TTC	TTC	TTC
G	Gly	GGC	GGC	GGC	GGG	GGC	GGG	GGC	GGT
H	His	CAC	CAC	CAC	CAC	CAC	CAC	CAC	CAC
I	Ile	ATC	ATT	ATC	ATT	ATC	ATC	ATT	ATT
K	Lys	AAG	AAG	AAG	AAG	AAG	AAG	AAG	AAG
L	Leu	CTG	CTC	CTC	CTG	CTG	CTG	TTG	TTG
M	Met	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG
N	Asn	AAC	AAC	AAC	AAC	AAC	AAC	AAC	AAC
P	Pro	CCC	CCC	CCC	CCG	CCC	CCG	CCA	CCG
Q	Gln	CAG	CAA	CAG	CAG	CAG	CAG	CAG	CAA
R	Arg	CGG	CGC	CGC	AGA	AGA	AGA	CGT	CGG
S	Ser	AGC	TCC	TCC	AGC	AGC	AGC	TCT	TCT
T	Thr	ACC	ACT	ACC	ACC	ACC	ACC	ACT	ACG
V	Val	GTG	GTC	GTC	GTT	GTG	GTG	GTT	GTC
W	Trp	TGG	TGG	TGG	TGG	TGG	TGG	TGG	TGG
Y	Tyr	TAC	TAC	TAC	TAC	TAC	TAC	TAC	TAC
		MAP9	MAP10	MAP11	MAP12	MAP13	MAP14	MAP15	MAP16
*	End	TGA	TGA	TGA	TGA	TGA	TGA	TGA	TGA
A	Ala	GCG	GCG	GCG	GCT	GCG	GCT	GCC	GCC
C	Cys	TGT	TGC	TGT	TGT	TGT	TGT	TGT	TGC
D	Asp	GAC	GAC	GAC	GAC	GAC	GAC	GAC	GAC
E	Glu	GAA	GAG	GAA	GAA	GAA	GAA	GAA	GAG
F	Phe	TTC	TTC	TTC	TTC	TTC	TTC	TTC	TTC
G	Gly	GGT	GGG	GGT	GGT	GGT	GGT	GGT	GGC
H	His	CAC	CAC	CAC	CAC	CAC	CAC	CAC	CAC
I	Ile	ATT	ATC	ATT	ATC	ATT	ATT	ATC	ATC
K	Lys	AAG	AAG	AAG	AAG	AAG	AAG	AAG	AAG
L	Leu	CTC	CTG	TTG	TTG	CTC	TTG	CTG	CTC

M	**Met**	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG
N	**Asn**	AAC	AAC	AAC	AAC	AAC	AAC	AAC	AAC
P	**Pro**	CCA	CCG	CCA	CCA	CCA	CCA	CCG	CCA
Q	**Gln**	CAA	CAG	CAA	CAA	CAA	CAA	CAA	CAG
R	**Arg**	AGG	AGG	CGT	CGT	CGA	CGT	AGA	CGA
S	**Ser**	TCC	AGC	TCT	TCT	TCC	TCT	AGC	TCC
T	**Thr**	ACG	ACG	ACG	ACT	ACG	ACT	ACC	ACC
V	**Val**	GTC	GTG	GTT	GTC	GTC	GTT	GTC	GTG
W	**Trp**	TGG	TGG	TGG	TGG	TGG	TGG	TGG	TGG
Y	**Tyr**	TAC	TAC	TAC	TAC	TAC	TAC	TAC	TAC

[0113] As used herein, the terms "candidate nucleic acid sequence" and "candidate nucleotide sequence" refer to a nucleotide sequence (e.g., a nucleotide sequence encoding an antibody or a functional fragment thereof) that can be codon-optimized, for example, to improve its translation efficacy. In some aspects, the candidate nucleotide sequence is optimized for improved translation efficacy after *in vivo* administration.

[0114] The term "percent sequence identity" between two polypeptide or polynucleotide sequences refers to the number of identical matched positions shared by the sequences over a comparison window, taking into account additions or deletions (i.e., gaps) that must be introduced for optimal alignment of the two sequences. A matched position is any position where an identical nucleotide or amino acid is presented in both the target and reference sequence. Gaps presented in the target sequence are not counted since gaps are not nucleotides or amino acids. Likewise, gaps presented in the reference sequence are not counted since target sequence nucleotides or amino acids are counted, not nucleotides or amino acids from the reference sequence. When comparing DNA and RNA, thymine (T) and uracil (U) can be considered equivalent.

[0115] The percentage of sequence identity is calculated by determining the number of positions at which the identical amino-acid residue or nucleic acid base occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity. The comparison of sequences and determination of percent sequence identity between two sequences can be accomplished using readily available software both for online use and for download. Suitable software programs are available from various sources, and for alignment of both protein and nucleotide sequences. One suitable program to determine percent sequence identity is bl2seq, part of the BLAST suite of program available from the U.S. government's National Center for Biotechnology Information BLAST web site

(blast.ncbi.nlm.nih.gov). Bl2seq performs a comparison between two sequences using either the BLASTN or BLASTP algorithm. BLASTN is used to compare nucleic acid sequences, while BLASTP is used to compare amino acid sequences. Other suitable programs are, e.g., Needle, Stretcher, Water, or Matcher, part of the EMBOSS suite of bioinformatics programs and also available from the European Bioinformatics Institute (EBI) at www.ebi.ac.uk/Tools/psa.

[0116] Different regions within a single polynucleotide or polypeptide target sequence that aligns with a polynucleotide or polypeptide reference sequence can each have their own percent sequence identity. It is noted that the percent sequence identity value is rounded to the nearest tenth. For example, 80.11, 80.12, 80.13, and 80.14 are rounded down to 80.1, while 80.15, 80.16, 80.17, 80.18, and 80.19 are rounded up to 80.2. It also is noted that the length value will always be an integer.

[0117] In certain aspects, the percentage identity "%ID" of a first amino acid sequence (or nucleic acid sequence) to a second amino acid sequence (or nucleic acid sequence) is calculated as %ID = 100 x (Y/Z), where Y is the number of amino acid residues (or nucleobases) scored as identical matches in the alignment of the first and second sequences (as aligned by visual inspection or a particular sequence alignment program) and Z is the total number of residues in the second sequence. If the length of a first sequence is longer than the second sequence, the percent identity of the first sequence to the second sequence will be higher than the percent identity of the second sequence to the first sequence.

[0118] One skilled in the art will appreciate that the generation of a sequence alignment for the calculation of a percent sequence identity is not limited to binary sequence-sequence comparisons exclusively driven by primary sequence data. It will also be appreciated that sequence alignments can be generated by integrating sequence data with data from heterogeneous sources such as structural data (e.g., crystallographic protein structures), functional data (e.g., location of mutations), or phylogenetic data. A suitable program that integrates heterogeneous data to generate a multiple sequence alignment is T-Coffee, available at www.tcoffee.org, and alternatively available, e.g., from the EBI. It will also be appreciated that the final alignment used to calculate percent sequence identity can be curated either automatically or manually.

[0119] The term "amino acid substitution" refers to replacing an amino acid residue present in a parent sequence (e.g., a candidate sequence or a consensus sequence) with

another amino acid residue. An amino acid can be substituted in a parent sequence, for example, via chemical peptide synthesis or through recombinant methods known in the art. Accordingly, a reference to a "substitution at position X" refers to the substitution of an amino acid present at position X with an alternative amino acid residue. In some aspects, substitution patterns can be described according to the schema AnY, wherein A is the single letter code corresponding to the amino acid naturally present at position n, and Y is the substituting amino acid residue. In other aspects, substitution patterns can be described according to the schema An(YZ), wherein A is the single letter code corresponding to the amino acid residue substituting the amino acid naturally present at position X, and Y and Z are alternative substituting amino acid residue, i.e., A could be substituted by Y or Z. For example, for a sequence such as WYLQKPGQSPQLLIY (SEQ ID NO: 2216), a substitution described as P6S would be a substitution of the proline residue at position 6 of the polypeptide (counting from the amino terminus, i.e., from left to right) with a serine. A substitution described as Q11(KRN) would be a substitution of the glutamine residue at position 11 of the polypeptide with a lysine or an arginine or an asparagine. In the context of the present disclosure, substitutions (even when they referred to as amino acid substitution) are conducted at the nucleic acid level, i.e., substituting an amino acid residue with an alternative amino acid residue is conducted by substituting the codon encoding the first amino acid with a codon encoding the second amino acid.

[0120] A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art, including basic side chains (e.g., lysine, arginine, or histidine), acidic side chains (e.g., aspartic acid or glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, or cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, or tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, or histidine). Thus, if an amino acid in a polypeptide is replaced with another amino acid from the same side chain family, the amino acid substitution is considered to be conservative. In another aspect, a string of amino acids can be conservatively replaced with a structurally similar string that differs in order and/or composition of side chain family members.

[0121] Non-conservative amino acid substitutions include those in which (i) a residue having an electropositive side chain (e.g., Arg, His or Lys) is substituted for, or by, an electronegative residue (e.g., Glu or Asp), (ii) a hydrophilic residue (e.g., Ser or Thr) is substituted for, or by, a hydrophobic residue (e.g., Ala, Leu, Ile, Phe or Val), (iii) a cysteine or proline is substituted for, or by, any other residue, or (iv) a residue having a bulky hydrophobic or aromatic side chain (e.g., Val, His, Ile or Trp) is substituted for, or by, one having a smaller side chain (e.g., Ala or Ser) or no side chain (e.g., Gly).

[0122] Other amino acid substitutions can be readily identified by workers of ordinary skill. For example, for the amino acid alanine, a substitution can be taken from any one of D-alanine, glycine, beta-alanine, L-cysteine and D-cysteine. For lysine, a replacement can be any one of D-lysine, arginine, D-arginine, homo-arginine, methionine, D-methionine, ornithine, or D- ornithine. Generally, substitutions in functionally important regions that can be expected to induce changes in the properties of isolated polypeptides are those in which (i) a polar residue, e.g., serine or threonine, is substituted for (or by) a hydrophobic residue, e.g., leucine, isoleucine, phenylalanine, or alanine; (ii) a cysteine residue is substituted for (or by) any other residue; (iii) a residue having an electropositive side chain, e.g., lysine, arginine or histidine, is substituted for (or by) a residue having an electronegative side chain, e.g., glutamic acid or aspartic acid; or (iv) a residue having a bulky side chain, e.g., phenylalanine, is substituted for (or by) one not having such a side chain, e.g., glycine. The likelihood that one of the foregoing non-conservative substitutions can alter functional properties of the protein is also correlated to the position of the substitution with respect to functionally important regions of the protein: some non-conservative substitutions can accordingly have little or no effect on biological properties.

[0123] The phrase "nucleotide sequence encoding" and variants thereof refers to the nucleic acid (e.g., an mRNA or DNA molecule) coding sequence that comprise a nucleotide sequence which encodes an antibody or functional fragment thereof as set forth herein. The coding sequence can further include initiation and termination signals operably linked to regulatory elements including a promoter and polyadenylation signal capable of directing expression in the cells of an individual or mammal to whom the nucleic acid is administered. The coding sequence can further include sequences that encode signal peptides.

II. Codon-optimized Nucleic Acid Sequences Encoding Antibodies

[0124] The present disclosure is directed to polynucleotides comprising codon-optimized nucleotide sequences (e.g., mRNA sequences) encoding antibodies, antibody functional fragments (e.g., an antigen-binding fragment thereof or an Fc fragment), antibody variants, or combinations thereof. These polypeptides can be used to express the antibodies and functional fragments thereof, for example, *in vivo* in a host organism (e.g., in a particular tissue or cell). The codon-optimized nucleotide sequences presented in the instant disclosure can present improved properties related to expression efficacy, for example, of an mRNA (e.g., a synthetic mRNA) administered *in vivo* to a subject in need thereof. Such properties include, but are not limited to, improving nucleic acid stability (e.g., mRNA stability), increasing translation efficacy in the target tissue, reducing the number of truncated proteins expressed, improving the folding or prevent misfolding of the expressed proteins, reducing toxicity of the expressed products, reducing cell death caused by the expressed products, increasing or decreasing protein aggregation, etc.

[0125] The recombinant expression of large molecules in cell cultures, and in particular molecules comprising several subunits, for example antibodies, can be a challenging task with numerous limitations. These limitations can be avoided by administering the polynucleotides disclosed herein (e.g., mRNAs) which encode a therapeutic agent of interest such as an antibody or an antigen binding fragment thereof to a patient, so the synthesis and delivery of the therapeutic agent takes place endogenously.

[0126] Changing from an *in vitro* expression system (e.g., cell culture) to *in vivo* expression requires the redesign of the nucleic acid encoding the therapeutic agent. Redesigning a naturally occurring gene sequence by choosing different codons without necessarily altering the encoded amino acid sequence can often lead to dramatic increases in protein expression levels (Gustafsson et al., 2004, Journal/Trends Biotechnol 22, 346-53). Variables such as codon adaptation index (CAI), mRNA secondary structures, cis-regulatory sequences, GC content and many other similar variables have been shown to somewhat correlate with protein expression levels (Villalobos et al., 2006, " Journal/BMC Bioinformatics 7, 285). However, due to the degeneracy of the genetic code, there are numerous different nucleotide sequences that can all encode the same therapeutic agent. Each amino acid is encoded by up to six synonymous codons; and the choice between these codons influences gene expression. In addition, codon usage (i.e., the frequency

with which different organisms use codons for expressing a polypeptide sequence) differs among organisms (for example, recombinant production of human or humanized therapeutic antibodies frequently takes place in hamster cell cultures).

[0127] Accordingly, the present disclosure provides nucleotide sequences encoding antibodies and functional fragments thereof that have been optimized for expression in human subjects, and which have structural and/or chemical features that avoid one or more of the problems in the art, for example, features which are useful for optimizing formulation and delivery of nucleic acid-based therapeutics while retaining structural and functional integrity, overcoming the threshold of expression, improving expression rates, half-life and/or protein concentrations, optimizing protein localization, and avoiding deleterious bio-responses such as the immune response and/or degradation pathways.

[0128] The terms "antibody" or "immunoglobulin," are used interchangeably herein, and include whole antibodies and any antigen binding fragment or single chains thereof. A typical antibody comprises at least two heavy (H) chains and two light (L) chains interconnected by disulfide bonds (see FIG. 12). Each heavy chain is comprised of a heavy chain variable region (abbreviated herein as VH or VH) and a heavy chain constant region. The heavy chain constant region is comprised of three domains, CH1, CH2, and CH3. Each light chain is comprised of a light chain variable region (abbreviated herein as VL or VL) and a light chain constant region. The light chain constant region is comprised of one domain, CL. The VH and VL regions can be further subdivided into regions of hypervariability, termed Complementarity Determining Regions (CDR), interspersed with regions that are more conserved, termed framework regions (FW). Each VH and VL is composed of three CDRs and four FWs, arranged from amino-terminus to carboxy-terminus in the following order: FW1, CDR1, FW2, CDR2, FW3, CDR3, and FW4. The variable regions of the heavy and light chains contain a binding domain that interacts with an antigen. The constant regions of the antibodies can mediate the binding of the immunoglobulin to host tissues or factors, including various cells of the immune system (e.g., effector cells) and the first component (C1q) of the classical complement system.

[0129] The term "antibody" encompasses any immunoglobulin molecules that recognize and specifically bind to a target, such as a protein, polypeptide, peptide, carbohydrate, polynucleotide, lipid, or combinations thereof through at least one antigen recognition site within the variable region of the immunoglobulin molecule.

[0130] As used herein, the term "antibody" encompasses intact polyclonal antibodies, intact monoclonal antibodies, antibody fragments (such as Fab, Fab', F(ab')2, and Fv fragments), single chain Fv (scFv) mutants, multispecific antibodies such as bispecific antibodies generated from at least two intact antibodies, chimeric antibodies, humanized antibodies, human antibodies, fusion proteins comprising an antigen determination portion of an antibody, and any other modified immunoglobulin molecule comprising an antigen recognition site so long as the antibodies exhibit the desired biological activity.

[0131] An antibody can be of any the five major classes of immunoglobulins: IgA, IgD, IgE, IgG, and IgM, or subclasses (isotypes) thereof (*e.g.* IgG1, IgG2, IgG3, IgG4, IgA1 and IgA2), based on the identity of their heavy-chain constant domains referred to as alpha, delta, epsilon, gamma, and mu, respectively. The different classes of immunoglobulins have different and well known subunit structures and three-dimensional configurations.

[0132] The term antibody also encompasses molecules comprising an immunoglobulin domain from an antibody (e.g., a VH, CL, CL, CH1, CH2 or CH3 domain) fused to other molecules, i.e., fusion proteins. In some aspects, such fusion protein comprises an antigen-binding moiety (e.g., an scFv). The antibody moiety of a fusion protein comprising g an antigen-binding moiety can be used to direct a therapeutic agent (e.g., a cytotoxin) to a desired cellular or tissue location determined by the specificity of the antigen-binding moiety.

[0133] In other aspects, the fusion protein can comprise a functional fragment of an antibody that is not an antigen-binding fragment, for example, an Fc domain. In this case, the Fc domain can be fused to a therapeutic agent (e.g., a bioactive peptide) and provide a desirable property, for example, increased plasma half-life.

[0134] The term "therapeutic antibody" is used in a broad sense, and encompasses any antibody or a functional fragment thereof that functions to deplete target cells in a patient, as well as molecules that deliver a therapeutic agent to a target cell in a patient (e.g., a cytotoxin or a bioactive peptide). Specific examples of such target cells include tumor cells, virus -infected cells, allogenic cells, pathological immunocompetent cells (e.g., B lymphocytes, T lymphocytes, antigen-presenting cells, etc.) involved in cancers, allergies, autoimmune diseases, allogenic reactions. The therapeutic antibodies can, for instance, mediate a cytotoxic effect or cell lysis, particularly by antibody-dependent cell-mediated cytotoxicity (ADCC). Therapeutic antibodies according to the disclosure can be directed

to epitopes of surface which are overexpressed by cancer cells, or directed to viral epitopes of surface.

[0135] In some aspects, the therapeutic antibody is a blocking antibody. The terms "blocking antibody" or "antagonist antibody" refer to an antibody which inhibits or reduces the biological activity of the antigen it binds. In a certain aspect blocking antibodies or antagonist antibodies substantially or completely inhibit the biological activity of the antigen. In some aspects, the biological activity is reduced by at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%,at least 85%, at least 90%, at least 95%, or even 100%.

[0136] In other aspects, the antibody is a "targeting antibody." The term "targeting antibody" refers to an antibody that delivers an effector molecule or molecules to a target site. In some aspects, the antibody directly delivers the effector molecule (e.g., a cytotoxic agent such as a *Pseudomonas* toxin) to the specific target location. In other aspects, the effector molecule can be released, e.g., after proteolytic cleavage from the targeting antibody, at or near target cells, tissues and organs.

[0137] The term "functional fragment" as used herein, when used in reference to a targeting antibody, is intended to refer to a portion of the targeting antibody which is capable of specifically binding an antigen that is specifically bound by the antibody reference is made to. The term functional fragment also refers to a construct derived from an antibody that functions as a blocking or a targeting antibody, e.g., an scFv. Also included within the definition are non-antigen binding fragments, for example, an Fc fragment. In some aspects, a functional Fc fragment possesses the "effector function" of a native sequence Fc region. Exemplary "effector functions" include Clq binding; complement dependent cytotoxicity; Fc receptor binding; antibody-dependent cell-mediated cytotoxicity (ADCC); phagocytosis; down regulation of cell surface receptors (e.g. B cell receptor; BCR), etc. In some aspects, such effector functions require the Fc region to be combined with a binding domain (e.g. an antibody variable domain) and can be assessed using various assays known in the art. In the context of the present disclosure, an Fc domain or variant thereof fused to a therapeutic agent to provide increased plasma half-life is considered a functional fragment. Whether a fragment is "functional" can be determined using assays known in the art. For example, whether a binding fragment is still capable to specifically binding to its antigen can be determined

using binding assays known in the art (e.g., BIACORE). For example, whether an Fc domain or variant thereof is capable of increasing plasma half-life of a therapeutic agent as part of a fusion protein can be determined using pharmacokinetic methods know in the art.

[0138] The term "antigen binding fragment" refers to a molecule comprising a portion of an intact antibody, and in particular refers to a molecule comprising and least one of the antigenic determining variable regions of an intact antibody. It is known in the art that the antigen binding function of an antibody can be performed by fragments of a full-length antibody. Examples of antibody fragments include, but are not limited to Fab, Fab', F(ab')2, and Fv fragments, linear antibodies, single chain antibodies, and multispecific antibodies formed from antibody fragments.

[0139] The term "non-antigen-binding fragment" refers to a molecule comprising a portion of an intact antibody, and in particular refers to a molecule that does not comprise the antigenic determining variable regions of an intact antibody. Examples of non-antigen binding fragments include Fc, Fc', pFc, pFc' fragments, and variants thereof.

[0140] The term "variant" as used herein with respect to a nucleic acid means (i) a portion or fragment of a referenced nucleotide sequence; (ii) the complement of a referenced nucleotide sequence or portion thereof; (iii) a nucleic acid that is substantially identical to a referenced nucleotide sequence or the complement thereof; (iv) a nucleotide sequence that hybridizes under stringent conditions to the referenced nucleotide sequence, complement thereof, or a sequence substantially identical thereto, or (v) a nucleotide sequence comprising one or more substitutions and encodes a polypeptide retaining at least one biological activity (e.g., antigen binding) of the polypeptide encoded by the referenced nucleotide sequence. "Variant" with respect to a polypeptide refers to a polypeptide that differs in amino acid sequence by the insertion, deletion, or conservative substitution of amino acids, but retains at least one biological activity of a reference polypeptide sequence (e.g., antigen binding).

[0141] A "monoclonal antibody" refers to a homogeneous antibody population involved in the highly specific recognition and binding of a single antigenic determinant, or epitope. This is in contrast to polyclonal antibodies that typically include different antibodies directed against different antigenic determinants. The term "monoclonal antibody" encompasses both intact and full-length monoclonal antibodies as well as antibody fragments (such as Fab, Fab', F(ab')2, Fv), single chain variable fragments

(scFv), fusion proteins comprising an antibody portion, and any other modified immunoglobulin molecule comprising an antigen recognition site.

[0142] The term "human antibody" means an antibody produced by a human or an antibody having an amino acid sequence corresponding to an antibody produced by a human made using any technique known in the art. Thus, the term human antibody also encompasses an antibody expressed *in vivo* in an animal subject, and an antibody having an amino acid sequence corresponding to an antibody originally produced by a human but expressed in a non-human system (e.g., a nucleotide sequence encoding an antibody produced by chemical synthesis and expressed *in vitro* in cultured mammal cells). This definition of a human antibody includes intact or full-length antibodies, fragments thereof, and/or antibodies comprising at least one human heavy and/or light chain polypeptide such as, for example, an antibody comprising murine light chain and human heavy chain polypeptides.

[0143] The term "humanized antibody" refers to an antibody derived from a non-human (*e.g.*, murine) immunoglobulin, which has been engineered to contain minimal non-human (*e.g.*, murine) sequences. Typically, humanized antibodies are human immunoglobulins in which residues from the CDRs are replaced by residues from the CDR of a non-human species (*e.g.*, mouse, rat, rabbit, or hamster) that have the desired specificity, affinity, and capability (Jones *et al.*, 1986, *Nature*, 321:522-525; Riechmann *et al.*, 1988, *Nature*, 332:323-327; Verhoeyen *et al.*, 1988, *Science*, 239:1534-1536). In some instances, the framework (FW) amino acid residues of a human immunoglobulin are replaced with the corresponding residues in an antibody from a non-human species that has the desired specificity, and/or affinity, and/or capability. The humanized antibody can be further modified by the substitution of additional residues either in the Fv framework region and/or within the replaced non-human residues to refine and optimize antibody specificity, affinity, and/or capability. In general, the humanized antibody will comprise substantially all of at least one, and typically two or three, variable domains containing all or substantially all of the CDR regions that correspond to the non-human immunoglobulin, whereas all or substantially all of the FR regions are those of a human immunoglobulin consensus sequence. The humanized antibody can also comprise at least a portion of an immunoglobulin constant region or domain (Fc), typically that of a human immunoglobulin. Examples of methods used to generate humanized antibodies are described in U.S. Pat. Nos. 5,225,539 or 5,639,641.

[0144] The term "chimeric antibodies" refers to antibodies wherein the amino acid sequence of the immunoglobulin molecule is derived from two or more animal species. Typically, the variable region of both light and heavy chains corresponds to the variable region of antibodies derived from one species of mammals (e.g., mouse, rat, rabbit, etc.) with the desired specificity, and/or affinity, and/or capability while the constant regions are homologous to the sequences in antibodies derived from another specie (usually human) to avoid eliciting an immune response in that species.

[0145] In all these types of antibodies described above, i.e., human antibodies, humanized antibodies, chimeric antibodies, etc., the nucleotide sequence encoding the antibody can be a codon-optimized nucleotide sequence.

[0146] A "variable region" of an antibody refers to the variable region of the antibody light chain or the variable region of the antibody heavy chain, either alone or in combination. The variable regions of the heavy and light chain each consist of four FW regions connected by three CDR regions (see FIGS. 12 and 13). The CDRs in each chain are held together in close proximity by the FW regions and, with the CDRs from the other chain, contribute to the formation of the antigen-binding site of antibodies. There are several techniques for determining the location of CDRs. The Kabat numbering system is generally used when referring to a residue in the variable domain (approximately residues 1-107 of the light chain and residues 1-113 of the heavy chain) (*e.g.*, Kabat *et al.*, Sequences of Immunological Interest, 5th Ed. Public Health Service, National Institutes of Health, Bethesda, Md. (1991)). The term "Kabat position" and grammatical variants thereof refer to the numbering system used for heavy chain variable domains or light chain variable domains of the compilation of antibodies in Kabat *et al.*, Sequences of Proteins of Immunological Interest, 5th Ed. Public Health Service, National Institutes of Health, Bethesda, Md. (1991). Using this numbering system, the actual linear amino acid sequence can contain fewer or additional amino acids corresponding to a shortening of, or insertion into, a FW or CDR of the variable domain. For example, a heavy chain variable domain can include a single amino acid insert (residue 52a according to Kabat) after residue 52 of H2 and inserted residues (*e.g.*, residues 82a, 82b, and 82c, *etc.* according to Kabat) after heavy chain FW residue 82.

TABLE 3: Location of loops in variable domains of light (L) and heavy (H) chain of antibodies according to the Kabat, AbM and Chothia numbering systems

Loop	Kabat	AbM	Chothia
L1	L24-L34	L24-L34	L24-L34
L2	L50-L56	L50-L56	L50-L56
L3	L89-L97	L89-L97	L89-L97
H1	H31-H35B	H26-H35B (Kabat Numbering)	H26-H32..34
H1	H31-H35	H26-H35 (Chothia Numbering)	H26-H32
H2	H50-H65	H50-H58	H52-H56
H3	H95-H102	H95-H102	H95-H102

[0147] The Kabat numbering of residues can be determined for a given antibody by alignment at regions of homology of the sequence of the antibody with a "standard" Kabat numbered sequence. Chothia refers instead to the location of the structural loops (Chothia and Lesk, J. Mol. Biol. 196:901-917 (1987)). The end of the Chothia CDR-H1 loop when numbered using the Kabat numbering convention varies between H32 and H34 depending on the length of the loop (this is because the Kabat numbering scheme places the insertions at H35A and H35B; if neither 35A nor 35B is present, the loop ends at 32; if only 35A is present, the loop ends at 33; if both 35A and 35B are present, the loop ends at 34). The AbM hypervariable regions represent a compromise between the Kabat CDRs and Chothia structural loops, and are used by Oxford Molecular's AbM antibody modeling software.

[0148] IMGT (ImMunoGeneTics) also provides a numbering system for the immunoglobulin variable regions, including the CDRs. See *e.g.*, Lefranc, M.P. *et al.*, Dev. Comp. Immunol. 27: 55-77(2003), which is herein incorporated by reference. The IMGT numbering system was based on an alignment of more than 5,000 sequences, structural data, and characterization of hypervariable loops and allows for easy comparison of the variable and CDR regions for all species. According to the IMGT numbering schema VH-CDR1 is at positions 26 to 35, VH-CDR2 is at positions 51 to 57, VH-CDR3 is at positions 93 to 102, VL-CDR1 is at positions 27 to 32, VL-CDR2 is at positions 50 to 52, and VL-CDR3 is at positions 89 to 97.

[0149] The EU index or EU numbering system is based on the sequential numbering of the first human IgG sequenced (the EU antibody). Because the most common reference

for this convention is the Kabat sequence manual (Kabat *et al.*, 1991), the EU index is sometimes erroneously used synonymously with the Kabat index. The EU index does not provide insertions and deletions, and thus in some cases comparisons of IgG positions across IgG subclass and species can be unclear, particularly in the hinge regions.

[0150] The boundaries of the antibody structural elements presented in this disclosure, namely, CDR1, CDR2, and CDR3 and FW1, FW2, FW3 and FW4 of VH or VL domain; VH and VL domain; and constant domain CL, CH1, CH2, and CH3 correspond to the boundaries indicated in the multiple sequence alignments shown in FIGS. 10 and 11. Alternatively, the boundaries can be determined with respect to the domains defined by the Position Specific Scoring Matrices of FIGS. 1 to 9 (first and last amino acid in each PSSM). The boundaries between antibody structural elements can also be defined in accordance with the numbering schemas discussed above. In addition, the boundaries between antibody structural elements can be obtained from the IMGT database, e.g., accessing the database at the URL imgt.org/mAb-DB/query, entering the International Nonproprietary Name (INN) of an antibody, and following the hyperlink to the antibody secondary structure. Alternatively, it is possible to identify the boundaries between the structural elements of an antibody by accessing the Uniform Resource Locator (URL) imgt.org/3Dstructure-DB/cgi/details.cgi?pdbcode=INN, wherein INN is the INN Number corresponding to a certain INN Name. For example, a person skilled in the art would be able to determine that the INN Number corresponding to INN Name adalimumab would be 7860, and therefore the URL providing boundaries between the structural elements in adalimumab would be imgt.org/3Dstructure-DB/cgi/details.cgi?pdbcode=7860. The boundaries between structural elements in an antibody can also be identified from sequence data alone by using the Paratome tool available at URL tools.immuneepitope.org/paratome/. See, Kunik et al. (2012) PLoS Comput. Biol. 8:2; Kunik et al. (2012). Nucleic Acids Res. 40(Web Server issue):W521-4.

[0151] As used herein the term "Fc region" or "Fc domain" includes the polypeptides comprising the constant region of an antibody excluding the first constant region immunoglobulin domain. Thus Fc refers to the last two constant region immunoglobulin domains of an IgG and the flexible hinge N-terminal to these domains. Although the boundaries of the Fc region can vary, the human IgG heavy chain Fc region is usually defined to comprise residues C226 or P230 to its carboxyl-terminus, wherein the numbering is according to the EU index as set forth in Kabat (Kabat *et al.*, Sequences of

Proteins of Immunological Interest, 5th Ed. Public Health Service, National Institutes of Health, Bethesda, Md. (1991)). Fc can refer to this region in isolation, or this region in the context of an antibody, antibody fragment, or Fc fusion protein. Polymorphisms have been observed at a number of different Fc positions, including but not limited to positions 270, 272, 312, 315, 356, and 358 as numbered by the EU index, and thus slight differences between the presented sequence and sequences in the prior art can exist. Numerous amino acid substitutions in the Fc domain are known in the art.

[0152] The term "hinge region" is generally defined as stretching from Glu216 to Pro230 of human IgGl (Burton, Molec. Immunol. (1985) 22:161-206). Hinge regions of other IgG isotypes can be aligned with the IgGl sequence by placing the first and last cysteine residues forming inter-heavy chain S— S bonds in the same positions.

[0153] The term "epitope" as used herein refers to an antigenic protein determinant capable of binding to an antibody or antigen-binding fragment thereof disclosed herein. Epitopes usually consist of chemically active surface groupings of molecules such as amino acids or sugar side chains and usually have specific three dimensional structural characteristics, as well as specific charge characteristics. The part of an antibody or binding molecule that recognizes the epitope is called a paratope. The epitopes of protein antigens are divided into two categories, conformational epitopes and linear epitopes, based on their structure and interaction with the paratope. A conformational epitope is composed of discontinuous sections of the antigen's amino acid sequence. These epitopes interact with the paratope based on the 3-D surface features and shape or tertiary structure of the antigen. By contrast, linear epitopes interact with the paratope based on their primary structure. A linear epitope is formed by a continuous sequence of amino acids from the antigen.

[0154] The term "antibody binding site" refers to a region in the antigen comprising a continuous or discontinuous site (*i.e.*, an epitope) to which a complementary antibody specifically binds. Thus, the antibody binding site can contain additional areas in the antigen which are beyond the epitope and which can determine properties such as binding affinity and/or stability, or affect properties such as antigen enzymatic activity or dimerization. Accordingly, even if two antibodies bind to the same epitope within an antigen, if the antibody molecules establish distinct intermolecular contacts with amino acids outside of the epitope, such antibodies are considered to bind to distinct antibody binding sites.

a. Codon-optimized Nucleotide Sequences Defined by Domain Conservation

[0155] The codon-optimized nucleotide sequences presented in the instant disclosure can be described in terms of identity to conserved domains. Thus, the present disclosure provides polynucleotide sequences comprising codon-optimized nucleotide sequences encoding antibodies or functional fragments thereof, wherein the nucleotide sequences have significant matches to conserved domains defining immunoglobulin structural domains as described in the NCBI Conserved Domain Database (CDD) version 3.13 released January 9, 2015. The conserved domains in the CDD database are described by Position Specific Scoring Matrices (PSSMs).

[0156] For example, in this context, a VH domain in an antibody could be defined as a protein subsequence with a significant match to a Conserved Domain (CD) model with accession code CD04981 as determined by using Reverse Position-Specific BLAST (RPS-BLAST) (NCBI, Bethesda) with default parameters, for example, as implemented in the CD-Search tool available at URL www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi. See Marchler-Bauer & Bryant, Nucleic Acids Res. 32(W): 327-331. CD model CD04981 would be defined according to the Position Specific Scoring Matrix (PSSM) shown in FIG. 7. The same approach would be applied to other structural components of an antibody, namely VL, CL, CH1, CH2, and CH3 domains. The CDD database contains also CD models that generically define an immunoglobulin variable domain (i.e., a CD model that would encompass both VH and VL domains), or an immunoglobulin constant domain (i.e., a CD model that would encompass CL, CH1, CH2, and CH3 domains).

[0157] A PSSM (see FIGS. 1 to 9) is a type of scoring matrix in which amino acid substitution scores are given separately for each position in a protein multiple sequence alignment. PSSM scores are shown as positive or negative integers. Positive scores indicate that the given amino acid substitution occurs more frequently in the alignment than expected by chance, while negative scores indicate that the substitution occurs less frequently than expected. Large positive scores often indicate critical functional residues, which can be active site residues or residues required for other intermolecular interactions.

[0158] In the PSSMs shown if FIGS. 1 to 9, the first column includes the amino acid positions in the domain; the second column is a "PSSM Consensus Sequence"; the third column includes a "PSSM Master Sequence"; and the remaining columns are "PSSM

Scores." In the context of a PSSM, the PSSM consensus sequence for a CD contains, at each position, the most frequently occurring amino acid at that position in the seed alignment of the CD. For a position to be represented in the PSSM consensus sequence, it must contain an aligned residue (as opposed to a gap) in at least 50% of the aligned sequences. Therefore, the PSSM consensus sequence is not a real protein, but rather defines both the most observed residues and the extent of the PSSM; however, the PSSM consensus sequence is not used in calculating frequencies for the PSSM. The master sequence is the top listed sequence in the CD seed alignment. It is a real protein, and is the sequence to which all other sequences in the CD alignment are pairwise aligned. Where possible, the PSSM master sequence is a sequence with a solved 3D structure from the Protein Data Bank (PDB). The PSSM scores are displayed as log-odds scores, basically calculated as the log (base 2) of the observed substitution frequency at a given position divided by the expected substitution frequency at that position. Thus, a positive score (ratio>1) indicates that the observed frequency exceeds the expected frequency, suggesting that this substitution is favored in the CD, whereas a negative score (ratio<1) indicates the opposite, i.e., that the observed substitution frequency is less than the expected frequency, suggesting that the substitution is not favored.

[0159] The term "significant match" refers to a high confidence association between a query protein sequence and a Conserved Domain, resulting in a high confidence level for the inferred function of the query protein sequence. As used herein, a significant match corresponds to an alignment of a Conserved Domain model to a query protein sequence having an expectation value (E-value) equal or lower than a domain–specific threshold E-value, for example, an E-value of at least 10^{-10}, 10^{-20}, 10^{-30}, 10^{-40}, 10^{-50}, or 10^{-60}. Thus, if for example, the query sequence was an antibody sequence encoded by a codon-optimized nucleotide sequence disclosed herein, a significant match to an CD domain defined by a PSSM (e.g., CD04981) would be an RPS-BLAST match with an E-value of at least 10^{-10}, 10^{-20}, 10^{-30}, 10^{-40}, 10^{-50}, or 10^{-60}, and such match would indicate that the matching sequence was a VH domain.

[0160] The term "Ig polypeptide" or "immunoglobulin polypeptide" refers to a polypeptide comprising a immunoglobulin (Ig) fold, i.e., 2-layer sandwich structure of between 7 and 9 antiparallel β-strands arranged in two β-sheets with a Greek key topology (see FIG. 13). The backbone switches repeatedly between the two β-sheets. Typically, the pattern is (N-terminal β-hairpin in sheet 1)-(β-hairpin in sheet 2)-(β-strand

in sheet 1)-(C-terminal β-hairpin in sheet 2). The cross-overs between sheets form an "X", so that the N- and C-terminal hairpins are facing each other.

[0161] In some aspects, the boundaries of a structural domain of an antibody (e.g., a CH1 domain) may not correspond exactly to the boundaries of the domain as defined by the PSSM. Accordingly, in some aspects, a significant match can be established between the amino acid sequence of a structural domain encoded by a codon-optimized nucleotide sequence disclosed herein (e.g., a CH1 domain), which could be the isolated domain or a subsequence of a codon-optimized heavy chain or light chain, and a "corresponding sequence of the CDD domain." For example, a structural domain could have a length of 100 amino acids, and the CDD domain defining such structural domain could encompass the core of the structural domain, e.g., 80 amino acids. In that case, a significant match could be established between the 80 amino acids in the core of the structural domain and the corresponding sequence of the CDD domain, i.e., the 80 positions covered by the PSSM defining the CDD domain.

[0162] In some aspects, a polynucleotide disclosed herein comprises a nucleotide sequence encoding an Ig constant domain of an antibody or a functional fragment thereof (e.g., CL, CH1, CH2, or CH3 constant domain from an IgG) which is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to (i) any one of the codon-optimized nucleotide sequences of SEQ ID NOS:1-88, or (ii) a subsequence of any one of the codon-optimized nucleotide sequences of SEQ ID NOS: 89-1978, wherein the subsequence encodes an Immunoglobulin (Ig) polypeptide that has a significant match to a corresponding sequence of CDD domain CD00098 (FIG. 1).

[0163] In other aspects, a polynucleotide disclosed herein comprises a nucleotide sequence encoding a light chain constant region (CL) of an antibody or a functional fragment thereof which is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to (i) any one of the codon-optimized nucleotide sequences of SEQ ID NOS:1-8, or 45-52, or (ii) a subsequence of any one of codon-optimized nucleotide sequences of SEQ ID NOS:1034-1978, wherein the subsequence encodes an Ig polypeptide that has a significant match to a corresponding sequence of CDD domain CD07699 (FIG. 2).

[0164] In other aspects, a polynucleotide disclosed herein comprises a nucleotide sequence encoding a first heavy chain constant domain (CH1) of an antibody or a

functional fragment thereof which is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to (i) any one of the codon-optimized nucleotide sequences of SEQ ID NOS:9-12, 21-24, 33-36, 53-56, 65-68, or 77-80, or (ii) a subsequence of any one of the codon-optimized nucleotide sequences of SEQ ID NOS: 89-1033, wherein the nucleotide subsequence encodes an Ig polypeptide that has a significant match to a corresponding sequence of CDD domain CD04985 (FIG. 3).

[0165] In other aspects, a polynucleotide disclosed herein comprises a nucleotide sequence encoding a second heavy chain constant domain (CH2) of an antibody or a functional fragment thereof which is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical to (i) any one of the codon-optimized nucleotide sequences of SEQ ID NO:13-16, 25-28, 37-40, 57-60, 69-72, or 81-84, or (ii) a subsequence of any one of the codon-optimized nucleotide sequences of SEQ ID NOS: 89-1033, wherein the nucleotide subsequence encodes an Ig polypeptide that has a significant match to a corresponding sequence of CDD domain CD04986 (FIG. 4).

[0166] In other aspects, a polynucleotide disclosed herein comprises a nucleotide sequence encoding a third heavy chain constant domain (CH3) of an antibody or a functional fragment thereof which is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to (i) any one of the codon-optimized nucleotide sequences of SEQ ID NO:17-20, 29-32, 41-44, 61-64, 73-76, or 85-88, or (ii) a subsequence of any one of the codon-optimized nucleotide sequences of SEQ ID NOS: 89-1033, wherein the nucleotide subsequence encodes an Ig polypeptide that has a significant match to a corresponding sequence of CDD domain CD07696 (FIG. 5).

[0167] In other aspects, a polynucleotide disclosed herein comprises a nucleotide sequence encoding a variable domain of an antibody (VH or VL) or a functional fragment thereof which is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical to a subsequence of any one of the codon-optimized nucleotide sequences of SEQ ID NOS: 89-1978, wherein the nucleotide subsequence encodes an Ig polypeptide that has a significant match to a corresponding sequence of CDD domain CD00099 (FIG. 6).

[0168] Also provided is a polynucleotide comprising a nucleotide sequence encoding a heavy chain variable domain (VH) of an antibody or a functional fragment thereof which is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical to a subsequence of any one of the codon-optimized nucleotide sequences of SEQ ID NOS: 89-1033, wherein the nucleotide subsequence encodes an Ig polypeptide that has a significant match to a corresponding sequence of CDD domain CD04981 (FIG. 7).

[0169] Also provided is a polynucleotide comprising a nucleotide sequence encoding a light chain variable domain (either a VL kappa domain or a VL lambda domain) of an antibody or a functional fragment thereof which is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical a subsequence of any one of the codon-optimized nucleotide sequences of SEQ ID NOS: 1034-1978, wherein the nucleotide subsequence encodes an Ig polypeptide that has a significant match to a corresponding sequence of CDD domain CD04980 (FIG. 8) or CD04984 (FIG. 9).

[0170] Also provided is a polynucleotide comprising a nucleotide sequence encoding a heavy chain of an antibody or a functional fragment thereof which is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to any one of the codon-optimized nucleotide sequences of SEQ ID NOS:89-1033, wherein the nucleotide sequence encodes an Ig polypeptide that has non-overlapping significant matches to CDD domains CD04981/CD4984, CD04985, and CD04986.

[0171] Also provided is a polynucleotide comprising a nucleotide sequence encoding light chain of an antibody or a functional fragment thereof which is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to any one of the codon-optimized nucleotide sequences of SEQ ID NO:1034-1978, wherein the nucleotide sequence encodes an Ig polypeptide that has non-overlapping significant matches to CD04980 and CD07699.

[0172] A person of skill in the art would appreciate that matching a query sequence to a CD domain is not limited to using RPS-BLAST as alignment tool, and that calculating the significance of the match is not limited to using the expectation value of the alignment.

b. Codon-optimized Nucleotide Sequences Defined by Sequence Conservation

[0173] The polynucleotide sequences disclosed herein can comprise codon-optimized nucleotide sequences which are defined in terms of sequence identity between the antibodies and fragment thereof encoded by such codon-optimized nucleotide sequences and the sequences or subsequences of therapeutic antibodies known in the art. These therapeutic antibodies known in the art can be defined according to their INN Names, or according to their publicly available protein sequences. Accordingly, the present invention provides codon-optimized nucleotide sequences encoding VH, VL, CL (kappa and lambda), CH1, CH2, or CH3 domain, or combinations thereof defined according to their similarity (level of sequence identity) to therapeutic antibodies known in the art (see TABLE 4).

[0174] In some aspects, the therapeutic antibody known in the art is abagovomab, abciximab, adalimumab, alemtuzumab, alirocumab, amatuximab, anrukinzumab, arcitumomab, basiliximab, bavituximab, benralizumab, bevacizumab, bezlotoxumab, bimagrumab, bococizumab, brentuximab, briakinumab, brodalumab, canakinumab, cantuzumab, carlumab, cetuximab, cixutumumab, clivatuzumab, conatumumab, crenezumab, dacetuzumab, daclizumab, dalotuzumab, denosumab, drozitumab, dupilumab, dusigitumab, eculizumab, elotuzumab, enokizumab, epratuzumab, etaracizumab, evolocumab, farletuzumab, fasinumab, fezakinumab, ficlatuzumab, figitumumab, fresolimumab, fulranumab, ganitumab, gantenerumab, gevokizumab, girentuximab, glembatumumab, ibalizumab, ibritumomab, icrucumab, inotuzumab, intetumumab, itolizumab, ixekizumab, lebrikizumab, lorvotuzumab, mavrilimumab, mepolizumab, milatuzumab, mogamulizumab, motavizumab, naptumomab, necitumumab, nivolumab, obinutuzumab, ocrelizumab, olaratumab, omalizumab, otelixizumab, oxelumab, pateclizumab, pembrolizumab, pertuzumab, ponezumab, ramucirumab, rilotumumab, rituximab, robatumumab, romosozumab, rontalizumab, samalizumab, sarilumab, secukinumab, sifalimumab, siltuximab, sirukumab, solanezumab, tabalumab, tanezumab, tenatumomab, teplizumab, tigatuzumab, tildrakizumab, tocilizumab, tositumomab, tralokinumab, trastuzumab, urelumab, ustekinumab, vedolizumab, or veltuzumab, or a functional fragment thereof. In other aspects, the therapeutic antibody is one of the therapeutic antibodies disclosed in TABLE 4.

TABLE 4: List of Therapeutic antibodies, including their target antigens and indication for treatment.

Antibody Name	Type	Source	Target antigen	Indication
3F8	Mab	mouse	GD2	Neuroblastoma
8H9	Mab	mouse	B7-H3	neuroblastoma, sarcoma, metastatic brain cancers
Abagovomab	Mab	mouse	CA-125 (imitation)	ovarian cancer
Abciximab	Fab	chimeric	CD41 (integrin alpha-IIb)	platelet aggregation inhibitor
Abrilumab	Mab	human	integrin α4β7	inflammatory bowel disease, ulcerative colitis, Crohn's disease
Actoxumab	Mab	human	Clostridium difficile	Clostridium difficile infection
Adalimumab	Mab	human	TNF-α	Rheumatoid arthritis, Crohn's Disease, Psoriasis, Spondylitis
Adecatumumab	Mab	human	EpCAM	prostate and breast cancer
Aducanumab	Mab	human	beta-amyloid	Alzheimer's disease
Afelimomab	F(ab')2	mouse	TNF-α	Sepsis
Afutuzumab	Mab	humanized	CD20	Lymphoma
Alacizumab	F(ab')2	humanized	VEGFR2	Cancer
ALD518		humanized	IL-6	rheumatoid arthritis
Alemtuzumab	Mab	humanized	CD52	Multiple sclerosis
Alirocumab	Mab	human	NARP-1	Hypercholesterolemia
Altumomab	Mab	mouse	CEA	colorectal cancer (diagnosis)
Amatuximab	Mab	chimeric	Mesothelin	Cancer
Anatumomab	Fab	mouse	TAG-72	non-small cell lung carcinoma
Anifrolumab	Mab	human	interferon α/β receptor	systemic lupus erythematosus
Anrukinzumab	Mab	humanized	IL-13	
Apolizumab	Mab	humanized	HLA-DR ?	hematological cancers
Arcitumomab	Fab'	mouse	CEA	gastrointestinal cancers (diagnosis)
Aselizumab	Mab	humanized	L-selectin (CD62L)	severely injured patients
Atinumab	Mab	human	RTN4	
Atlizumab	Mab	humanized	IL-6 receptor	rheumatoid arthritis
Atorolimumab	Mab	human	Rhesus factor	hemolytic disease of the newborn[citation needed]
Bapineuzumab	Mab	humanized	beta amyloid	Alzheimer's disease
Basiliximab	Mab	chimeric	CD25	prevention of organ transplant rejections
Bavituximab	Mab	chimeric	Phosphatidylserine	cancer, viral infections
Bectumomab	Fab'	mouse	CD22	non-Hodgkin's

Antibody Name	Type	Source	Target antigen	Indication
				lymphoma (detection)
Belimumab	Mab	human	BAFF	non-Hodgkin lymphoma etc.
Benralizumab	Mab	humanized	CD125	Asthma
Bertilimumab	Mab	human	CCL11 (eotaxin-1)	severe allergic disorders
Besilesomab	Mab	mouse	CEA-related antigen	inflammatory lesions and metastases (detection)
Bevacizumab	Mab	humanized	VEGF-A	metastatic cancer, retinopathy of prematurity
Bezlotoxumab	Mab	human	Clostridium difficile	Clostridium difficile infection
Biciromab	Fab'	mouse	fibrin II, beta chain	thromboembolism (diagnosis)
Bimagrumab	Mab	human	ACVR2B	myostatin inhibitor
Bivatuzumab	Mab	humanized	CD44 v6	squamous cell carcinoma
Blinatumomab	BiTE	mouse	CD19	Cancer
Blosozumab	Mab	humanized	SOST	Osteoporosis
Brentuximab	Mab	chimeric	CD30 (TNFRSF8)	hematologic cancers
Briakinumab	Mab	human	IL-12, IL-23	psoriasis, rheumatoid arthritis, IBD, multiple sclerosis
Brodalumab	Mab	human	IL-17	inflammatory diseases
Canakinumab	Mab	human	IL-1?	rheumatoid arthritis
Cantuzumab	Mab	humanized	MUC1	Cancers
Caplacizumab	Mab	humanized	VWF	
Capromab	Mab	mouse	prostatic carcinoma cells	prostate cancer (detection)
Carlumab	Mab	human	MCP-1	oncology/immune indications
Catumaxomab	3functional	rat/mouse hybrid	EpCAM, CD3	ovarian cancer, malignant ascites, gastric cancer
CC49	Mab	mouse	TAG-72	tumor detection
cBR96	Mab	humanized	Lewis-Y antigen	Cancer
Cedelizumab	Mab	humanized	CD4	prevention of organ transplant rejections, autoimmune diseases
Certolizumab	Fab'	humanized	TNF-α	Crohn's disease
Cetuximab	Mab	chimeric	EGFR	metastatic colorectal cancer and head and neck cancer
Ch.14.18	Mab	chimeric	???	Neuroblastoma
Citatuzumab	Fab	humanized	EpCAM	ovarian cancer and other solid tumors
Cixutumumab	Mab	human	IGF-1 receptor	solid tumors
Clazakizumab	Mab	humanized	Oryctolagus cuniculus	rheumatoid arthritis
Clenoliximab	Mab	chimeric	CD4	rheumatoid arthritis

Antibody Name	Type	Source	Target antigen	Indication
Clivatuzumab	Mab	humanized	MUC1	pancreatic cancer
Conatumumab	Mab	Human	TRAIL-R2	Cancer
Concizumab	Mab	humanized	TFPI	Bleeding
Crenezumab	Mab	humanized	1-40-β-amyloid	Alzheimer's disease
CR6261	Mab	Human	Influenza A hemagglutinin	infectious disease/influenza A
Dacetuzumab	Mab	humanized	CD40	hematologic cancers
Daclizumab	Mab	humanized	CD25	prevention of organ transplant rejections
Dalotuzumab	Mab	humanized	ILGF I receptor	cancer etc.
Daratumumab	Mab	Human	CD38	
Demcizumab	Mab	humanized	DLL4	Cancer
Denosumab	Mab	Human	RANKL	osteoporosis, bone metastases etc.
Detumomab	Mab	Mouse	B-lymphoma cell	Lymphoma
Dinutuximab	Mab	chimeric	ganglioside GD2	Neuroblastoma
Diridavumab	Mab	Human	Hemagglutinin	influenza A
Dorlimomab	F(ab')2	Mouse		
Drozitumab	Mab	Human	DR5	cancer etc.
Duligotumab	Mab	Human	HER3	
Dupilumab	Mab	Human	IL4	atopic diseases
Durvalumab	Mab	Human	CD274	cancer
Dusigitumab	Mab	Human	ILGF2	cancer
Ecromeximab	Mab	chimeric	GD3 ganglioside	malignant melanoma
Eculizumab	Mab	humanized	C5	paroxysmal nocturnal hemoglobinuria
Edobacomab	Mab	Mouse	Endotoxin	sepsis caused by Gram-negative bacteria
Edrecolomab	Mab	Mouse	EpCAM	colorectal carcinoma
Efalizumab	Mab	humanized	LFA-1 (CD11a)	psoriasis (blocks T-cell migration)
Efungumab	scFv	human	Hsp90	invasive Candida infection
Eldelumab	Mab	human	IF gamma-induced protein	Crohn's disease, ulcerative colitis
Elotuzumab	Mab	humanized	SLAMF7	multiple myeloma
Elsilimomab	Mab	mouse	IL-6	
Emibetuzumab	Mab	humanized	HHGFR	cancer
Enavatuzumab	Mab	humanized	TWEAK receptor	cancer etc.
Enfortumab	Mab	human	AGS-22M6	cancer expressing Nectin-4
Enlimomab	Mab	mouse	ICAM-1 (CD54)	
Enokizumab	Mab	humanized	IL9	asthma
Enoticumab	Mab	human	DLL4	
Ensituximab	Mab	chimeric	5AC	cancer
Epitumomab	Mab	mouse	Episialin	

Antibody Name	Type	Source	Target antigen	Indication
Epratuzumab	Mab	humanized	CD22	cancer, SLE
Erlizumab	F(ab')2	humanized	ITGB2 (CD18)	heart attack, stroke, traumatic shock
Ertumaxomab	3functional	rat/mouse hybrid	HER2/neu, CD3	breast cancer etc.
Etaracizumab	Mab	humanized	integrin αvβ3	melanoma, prostate cancer, ovarian cancer etc.
Etrolizumab	Mab	humanized	integrin α7 β7	inflammatory bowel disease
Evinacumab	Mab	human	angiopoietin 3	dyslipidemia
Evolocumab	Mab	human	PCSK9	hypercholesterolemia
Exbivirumab	Mab	human	hepatitis B surface antigen	hepatitis B
Fanolesomab	Mab	mouse	CD15	appendicitis (diagnosis)
Faralimomab	Mab	mouse	interferon receptor	
Farletuzumab	Mab	humanized	folate receptor 1	ovarian cancer
Fasinumab	Mab	human	HNGF	
FBTA05	3functional	rat/mouse hybrid	CD20	chronic lymphocytic leukaemia
Felvizumab	Mab	humanized	respiratory syncytial virus	respiratory syncytial virus infection
Fezakinumab	Mab	human	IL-22	rheumatoid arthritis, psoriasis
Ficlatuzumab	Mab	humanized	HGF	cancer etc.
Figitumumab	Mab	human	IGF-1 receptor	adrenocortical carcinoma, non-small cell lung carcinoma etc.
Flanvotumab	Mab	human	TYRP1(glycoprotein 75)	melanoma
Fletikumab	Mab	human	IL 20	rheumatoid arthritis
Fontolizumab	Mab	humanized	IFN-γ	Crohn's disease etc.
Foralumab	Mab	human	CD3 epsilon	
Foravirumab	Mab	human	rabies virus glycoprotein	rabies (prophylaxis)
Fresolimumab	Mab	human	TGF-β	idiopathic pulmonary fibrosis, glomerulosclerosis, cancer
Fulranumab	Mab	human	NGF	pain
Futuximab	Mab	chimeric	EGFR	
Galiximab	Mab	chimeric	CD80	B-cell lymphoma
Ganitumab	Mab	human	IGF-I	Cancer
Gantenerumab	Mab	human	beta amyloid	Alzheimer's disease
Gavilimomab	Mab	mouse	CD147 (basigin)	graft versus host disease
Gemtuzumab	Mab	humanized	CD33	acute myelogenous leukemia
Gevokizumab	Mab	humanized	IL-1β	diabetes etc.
Girentuximab	Mab	chimeric	CA-IX	clear cell renal cell carcinoma[73]
Glembatumumab	Mab	human	GPNMB	melanoma, breast cancer
Golimumab	Mab	human	TNF-α	rheumatoid arthritis,

Antibody Name	Type	Source	Target antigen	Indication
				psoriatic arthritis, ankylosing spondylitis
Gomiliximab	Mab	Chimeric	CD23 (IgE receptor)	allergic asthma
Guselkumab	Mab	Human	IL23	Psoriasis
Ibalizumab	Mab	humanized	CD4	HIV infection
Ibritumomab	Mab	Mouse	CD20	non-Hodgkin's lymphoma
Icrucumab	Mab	Human	VEGFR-1	cancer etc.
Igovomab	F(ab')2	Mouse	CA-125	ovarian cancer (diagnosis)
IMAB362	Mab	Human	CLDN18.2	gastrointestinal adenocarcinomas and pancreatic tumor
Imciromab	Mab	Mouse	cardiac myosin	cardiac imaging
Imgatuzumab	Mab	humanized	EGFR	Cancer
Inclacumab	Mab	Human	selectin P	
Indatuximab	Mab	Chimeric	SDC1	Cancer
Infliximab	Mab	Chimeric	TNF-α	rheumatoid arthritis, psoriatic arthritis, Crohn's disease, UC
Intetumumab	mab	Human	CD51	solid tumors (prostate cancer, melanoma)
Inolimomab	Mab	Mouse	CD25 (	graft versus host disease
Inotuzumab	Mab	humanized	CD22	Cancer
Ipilimumab	Mab	Human	CD152	Melanoma
Iratumumab	Mab	Human	CD30 (TNFRSF8)	Hodgkin's lymphoma
Itolizumab	Mab	humanized	CD6	
Ixekizumab	Mab	humanized	IL-17A	autoimmune diseases
Keliximab	Mab	Chimeric	CD4	chronic asthma
Labetuzumab	Mab	humanized	CEA	colorectal cancer
Lambrolizumab	Mab	humanized	PDCD1	antineoplastic agent
Lampalizumab	Mab	humanized	CFD	
Lebrikizumab	Mab	humanized	IL-13	Asthma
Lemalesomab	Mab	Mouse	NCA-90	diagnostic agent
Lerdelimumab	Mab	Human	TGF beta 2	reduction of scarring after glaucoma surgery
Lexatumumab	Mab	Human	TRAIL-R2	Cancer
Libivirumab	Mab	Human	HBV surface antigen	hepatitis B
Lifastuzumab	Mab	humanized	PO4-Na co-transporter	Cancer
Ligelizumab	Mab	humanized	IGHE	severe asthma and chronic spontaneous urticaria
Lintuzumab	Mab	humanized	CD33	Cancer
Lirilumab	Mab	Human	KIR2D	
Lodelcizumab	Mab	humanized	PCSK9	Hypercholesterolemia
Lorvotuzumab	Mab	humanized	CD56	Cancer
Lucatumumab	Mab	Human	CD40	multiple myeloma, non-

Antibody Name	Type	Source	Target antigen	Indication
				Hodgkin's lymphoma, Hodgkin's lymphoma
Lulizumab pegol	Mab	humanized	CD28	autoimmune diseases
Lumiliximab	Mab	Chimeric	CD23 (IgE receptor)	chronic lymphocytic leukemia
Mapatumumab	Mab	Human	TRAIL-R1	Cancer
Margetuximab	Mab	humanized	ch4D5	Cancer
Maslimomab		mouse	T-cell receptor	
Mavrilimumab	Mab	human	GMCSF receptor α-chain	rheumatoid arthritis
Matuzumab	Mab	humanized	EGFR	colorectal, lung and stomach cancer
Mepolizumab	Mab	humanized	IL-5	asthma and white blood cell diseases
Metelimumab	Mab	human	TGF beta 1	systemic scleroderma
Milatuzumab	Mab	humanized	CD74	multiple myeloma and other hematological malignancies
Minretumomab	Mab	mouse	TAG-72	
Mitumomab	Mab	mouse	GD3 ganglioside	small cell lung carcinoma
Mogamulizumab	Mab	humanized	CCR4	Cancer
Morolimumab	Mab	human	Rhesus factor	
Motavizumab	Mab	humanized	respiratory syncytial virus	respiratory syncytial virus (prevention)
Moxetumomab	Mab	mouse	CD22	Cancer
Muromonab-CD3	Mab	mouse	CD3	prevention of organ transplant rejections
Nacolomab	Fab	mouse	C242 antigen	colorectal cancer
Namilumab	Mab	human	CSF2	
Naptumomab	Fab	mouse	5T4	non-small cell lung carcinoma, renal cell carcinoma
Narnatumab	Mab	human	RON	Cancer
Natalizumab	Mab	humanized	integrin α4	multiple sclerosis, Crohn's disease
Nebacumab	Mab	human	Endotoxin	Sepsis
Necitumumab	Mab	human	EGFR	non-small cell lung carcinoma
Nerelimomab	Mab	mouse	TNF-α	
Nesvacumab	Mab	human	angiopoietin 2	Cancer
Nimotuzumab	Mab	humanized	EGFR	squamous cell carcinoma, head and neck cancer, glioma
Nivolumab	Mab	human	IgG4	Cancer
Nofetumomab	Fab	mouse		cancer (diagnosis)
Obiltoxaximab	Mab	chimeric	Bacillus anthracis anthrax	Bacillus anthracis spores
Ocaratuzumab	Mab	humanized	CD20	Cancer
Ocrelizumab	Mab	humanized	CD20	rheumatoid arthritis, lupus erythematosus etc.

Antibody Name	Type	Source	Target antigen	Indication
Odulimomab	Mab	mouse	LFA-1 (CD11a)	prevention of organ transplant rejections, immunological diseases
Ofatumumab	Mab	human	CD20	chronic lymphocytic leukemia etc.
Olaratumab	Mab	human	PDGF-R α	Cancer
Olokizumab	Mab	humanized	IL6	
Omalizumab	Mab	humanized	IgE Fc region	allergic asthma
Onartuzumab	Mab	humanized	HSFR kinase	Cancer
Ontuxizumab	Mab	chimeric/humanized	TEM1	Cancer
Oportuzumab	scFv	humanized	EpCAM	Cancer
Oregovomab	Mab	mouse	CA-125	ovarian cancer
Orticumab	Mab	human	oxLDL	
Otelixizumab	Mab	chimeric/humanized	CD3	diabetes mellitus type 1
Otlertuzumab	Mab	humanized	CD37	Cancer
Oxelumab	Mab	human	OX-40	Asthma
Ozanezumab	Mab	humanized	NOGO-A	ALS and multiple sclerosis
Ozoralizumab	Mab	humanized	TNF-α	inflammation
Pagibaximab	Mab	chimeric	lipoteichoic acid	sepsis (Staphylococcus)
Palivizumab	Mab	humanized	F protein of RSV	respiratory syncytial virus (prevention)
Panitumumab	Mab	human	EGFR	colorectal cancer
Pankomab	Mab	humanized	glycosylation of MUC1	ovarian cancer
Panobacumab	Mab	human	Pseudomonas aeruginosa	Pseudomonas aeruginosa infection
Parsatuzumab	Mab	human	EGFL7	Cancer
Pascolizumab	Mab	humanized	IL-4	Asthma
Pateclizumab	Mab	humanized	LTA	TNF
Patritumab	Mab	human	HER3	Cancer
Pembrolizumab	Mab	humanized	PDCD1	cancer etc.
Pemtumomab		mouse	MUC1	Cancer
Perakizumab	Mab	humanized	IL17A	arthritis
Pertuzumab	Mab	humanized	HER2/neu	Cancer
Pexelizumab	scFv	humanized	C5	reduction of side effects of cardiac surgery
Pidilizumab	Mab	humanized	PD-1	cancer and infectious diseases
Pinatuzumab	Mab	humanized	CD22	cancer
Pintumomab	Mab	mouse	adenocarcinoma antigen	adenocarcinoma (imaging)
Placulumab	Mab	human	human TNF	
Polatuzumab	Mab	humanized	CD79B	cancer
Ponezumab	Mab	humanized	human beta-amyloid	Alzheimer's disease
Priliximab	Mab	chimeric	CD4	Crohn's disease, multiple sclerosis
Pritoxaximab	Mab	chimeric	E. coli shiga toxin type-1	

Antibody Name	Type	Source	Target antigen	Indication
Pritumumab	Mab	human	Vimentin	brain cancer
PRO 140		humanized	CCR5	HIV infection
Quilizumab	Mab	humanized	IGHE	
Racotumomab	Mab	mouse	N-glycolylneuraminic acid	cancer
Radretumab	Mab	human	Fn extra domain-B	cancer
Rafivirumab	Mab	human	rabies virus glycoprotein	rabies (prophylaxis)
Ramucirumab	Mab	human	VEGFR2	solid tumors
Ranibizumab	Fab	humanized	VEGF-A	macular degeneration (wet form)
Raxibacumab	Mab	human	anthrax toxin PA	anthrax (prophylaxis and treatment)
Regavirumab	Mab	human	CMV glycoprotein B	cytomegalovirus infection
Reslizumab	Mab	humanized	IL-5	inflammations of the airways, skin and gastrointestinal tract
Rilotumumab	Mab	human	HGF	solid tumors
Rituximab	Mab	chimeric	CD20	lymphomas, leukemias, some autoimmune disorders
Robatumumab	Mab	human	IGF-1 receptor	cancer
Roledumab	Mab	human	RHD	
Romosozumab	Mab	humanized	Sclerostin	osteoporosis
Rontalizumab	Mab	humanized	IFN-α	systemic lupus erythematosus
Rovelizumab	Mab	humanized	CD11, CD18	haemorrhagic shock etc.
Ruplizumab	Mab	humanized	CD154 (CD40L)	rheumatic diseases
Samalizumab	Mab	humanized	CD200	cancer
Sarilumab	Mab	human	IL6	rheumatoid arthritis, ankylosing spondylitis
Satumomab	Mab	mouse	TAG-72	cancer (diagnosis)
Secukinumab	Mab	human	IL-17A	uveitis, rheumatoid arthritis psoriasis
Seribantumab	Mab	human	ERBB3	cancer
Setoxaximab	Mab	chimeric	E. coli shiga toxin type-2	
Sevirumab		human	Cytomegalovirus	cytomegalovirus infection
Sibrotuzumab	Mab	humanized	FAP	cancer
SGN-CD19A	Mab	humanized	CD19	acute lymphoblastic leukemia and B-cell non-Hodgkin lymphoma
SGN-CD33A	Mab	humanized	CD33	Acute myeloid leukemia
Sifalimumab	mab	humanized	IFN-α	SLE, dermatomyositis, polymyositis
Siltuximab	Mab	chimeric	IL-6	cancer
Simtuzumab	Mab	humanized	LOXL2	fibrosis
Siplizumab	Mab	humanized	CD2	psoriasis, graft-versus-host disease (prevention)
Sirukumab	Mab	human	IL-6	rheumatoid arthritis

Antibody Name	Type	Source	Target antigen	Indication
Sofituzumab	CA 125			ovarian cancer
Solanezumab	Mab	humanized	beta amyloid	Alzheimer's disease
Solitomab	Mab	mouse	EpCAM	
Sonepcizumab		humanized	sphingosine-1-phosphate	choroidal and retinal neovascularization
Sontuzumab	Mab	humanized	Episialin	
Stamulumab	Mab	human	Myostatin	muscular dystrophy
Sulesomab	Fab'	mouse	NCA-90	osteomyelitis (imaging)
Suvizumab	Mab	humanized	HIV-1	viral infections
Tabalumab	Mab	human	BAFF	B-cell cancers
Tacatuzumab	Mab	humanized	alpha-fetoprotein	cancer
Tadocizumab	Fab	humanized	integrin αIIbβ3	percutaneous coronary intervention
Talizumab	Mab	humanized	IgE	allergic reaction
Tanezumab	Mab	humanized	NGF	pain
Taplitumomab	Mab	mouse	CD19	cancer[citation needed]
Tarextumab	Mab	human	Notch receptor	cancer
Tefibazumab	Mab	humanized	clumping factor A	Staphylococcus aureus infection
Telimomab	Fab	mouse		
Tenatumomab	Mab	mouse	tenascin C	cancer
Teneliximab	Mab	chimeric	CD40	
Teplizumab	Mab	humanized	CD3	diabetes mellitus type 1
Teprotumumab	Mab	human	CD221	hematologic tumors
TGN1412		humanized	CD28	chronic lymphocytic leukemia, rheumatoid arthritis
Ticilimumab	Mab	human	CTLA-4	cancer
Tildrakizumab	Mab	humanized	IL23	immunologically mediated inflammatory disorders
Tigatuzumab	Mab	humanized	TRAIL-R2	cancer
TNX-650		humanized	IL-13	Hodgkin's lymphoma
Tocilizumab	Mab	humanized	IL-6 receptor	rheumatoid arthritis
Toralizumab	Mab	humanized	CD154 (CD40L)	rheumatoid arthritis, lupus nephritis etc.
Tositumomab		mouse	CD20	follicular lymphoma
Tovetumab	Mab	human	CD140a	cancer
Tralokinumab	Mab	human	IL-13	asthma etc.
Trastuzumab	Mab	humanized	HER2/neu	breast cancer
TRBS07	3funct		GD2	melanoma
Tregalizumab	Mab	Humanized	CD4	
Tremelimumab	Mab	Human	CTLA-4	cancer
Tucotuzumab	Mab	Humanized	EpCAM	cancer
Tuvirumab		Human	hepatitis B virus	chronic hepatitis B
Ublituximab	Mab	Chimeric	MS4A1	cancer

Antibody Name	Type	Source	Target antigen	Indication
Urelumab	Mab	Human	4-1BB	cancer etc.
Urtoxazumab	Mab	Humanized	Escherichia coli	diarrhoea caused by E. coli
Ustekinumab	Mab	Human	IL-12, IL-23	multiple sclerosis, psoriasis, psoriatic arthritis
Vantictumab	Mab	Human	Frizzled receptor	cancer
Vapaliximab	Mab	Chimeric	AOC3 (VAP-1)	
Varlilumab	Mab	Human	CD27	
Vatelizumab	Mab	Humanized	ITGA2	
Vedolizumab	Mab	Humanized	integrin α4β7	Crohn's disease, ulcerative colitis
Veltuzumab	Mab	Humanized	CD20	non-Hodgkin's lymphoma
Vepalimomab	Mab	Mouse	AOC3 (VAP-1)	inflammation
Vesencumab	Mab	Human	NRP1	
Visilizumab	Mab	Humanized	CD3	Crohn's disease, ulcerative colitis
Volociximab	Mab	Chimeric	integrin α5β1	solid tumors
Vorsetuzumab	Mab	Humanized	CD70	cancer
Votumumab	Mab	Human	tumor antigen CTAA16.88	colorectal tumors
Zalutumumab	Mab	Human	EGFR	squamous cell carcinoma of the head and neck
Zanolimumab	Mab	Human	CD4	rheumatoid arthritis, psoriasis, T-cell lymphoma
Zatuximab	Mab	Chimeric	HER1	cancer
Ziralimumab	Mab	Human	CD147 (basigin)	
Zolimomab	mab	Mouse	CD5	systemic lupus erythematosus, graft-versus-host disease

[0175] The sequences of the therapeutic antibodies with their names in bold face in TABLE 4 are included in the multiple sequence alignments of FIGS. 10 and 11.

[0176] Accordingly, the present disclosure also provides a polynucleotide comprising a nucleotide sequence encoding a CL kappa domain from a therapeutic antibody presented in TABLE 4 or a functional fragment thereof or which is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical to any one of the codon-optimized nucleotide sequences of SEQ ID NOS: 1-4, or 45-48. In some aspects, the CL kappa domain comprises the amino acid sequence TVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNAL

QSGNSQESVTEQDSKDSTYSLSX[1]TLTLSKADYEKHKVYACEVTHQGLSSPVTKS FNR (SEQ ID NO: 2200), wherein X^1 is selected from Asparagine (N) and Serine (S).

[0177] Also provided is a polynucleotide comprising a nucleotide sequence encoding a CL lambda domain from a therapeutic antibody presented in TABLE 4 or a functional fragment thereof which is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to any one of the polynucleotides of SEQ ID NOS: 5-8, or 49-52. In some aspects, the CL lambda domain comprises the amino acid sequence PKAAPSVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTT PSKQ SNNKYAASSYLSLTPEQWKSHX[2]SYSCQVTHEGSTVEKTVAPX[3]ECS (SEQ ID NO: 2201), wherein X^2 is selected from Arginine (R) and Lysine (K), and X^3 is selected from Threonine (T) and Alanine (A).

[0178] The present disclosure also provides a polynucleotide comprising a nucleotide sequence encoding a heavy chain first constant domain (CH1) from a therapeutic antibody presented in TABLE 4 or a functional fragment thereof which is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to any one of the codon-optimized nucleotide sequences of SEQ ID NOS: 9-12, 21-24, 33-36, 53-56, 65-68, or 77-80. In some aspects, the nucleotide sequence is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to any one of the codon-optimized nucleotide sequences of SEQ ID NOs: 9-12, or 53-56, wherein the CH1 domain is an IgG1 CH1 domain. In some aspects, the IgG1 CH1 domain comprises the amino acid sequence SX[4]GPSVX[5]PLAPSSKSTSGGTAAL GCLVKDYFPEPVTVSWNSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNV NHKPSNTKVDKX[6]X[7] (SEQ ID NO: 2202) wherein X^4 is an optional ASTK sequence, X^5 is selected from Phenylalanine (F) and Leucine (L), X^6 is selected from Lysine (K) and Arginine (R), and X^7 is selected from Valine (V) and Alanine (A). In some aspects, the nucleotide sequence is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to any one of the codon-optimized nucleotide sequences of SEQ ID NO: 21-24, or 65-68, wherein the CH1 domain is an IgG2 CH1 domain. In some aspects, the IgG2 CH1 domain comprises the amino acid sequence SASTKGPSVF PLAPCSRSTSESTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYS

LSSVVTVX[15]SSNFGTQTYTCNVDHKPSNTKVDKTV (SEQ ID NO: 2205) wherein X^{15} is selected from Proline (P) and Threonine (T). In some aspects, the nucleotide sequence is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to any one of the codon-optimized nucleotide sequences of SEQ ID NO: 33-36, or 77-80, wherein the CH1 domain is an IgG4 CH1 domain. In some aspects, the IgG4 CH1 domain comprises the amino acid sequence SASTKGPSVFPLAPCSRSTSESTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPA VLQSSGLYSLSSVVTVPSSSLGTKTYTCNVDHKPSNTKVDKRV (SEQ ID NO: 2197).

[0179] The present disclosure also provides a polynucleotide comprising a nucleotide sequence encoding a CH2 domain from a therapeutic antibody presented in TABLE 4 or a functional fragment thereof which is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to any one of the codon-optimized nucleotide sequences of SEQ ID NOS: 13-16, 25-28, 37-40, 57-60, 69-72, or 81-84. In some aspects, the nucleotide sequence is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to any one of the codon-optimized nucleotide sequences of SEQ ID NO: 13-16, or 57-60, wherein the CH2 domain is an IgG1 CH2 domain. In some aspects, the IgG1 CH2 domain comprises the amino acid sequence APEX[8]X[9]GX[10]PSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFN X[11]YVDGVEVHNAKTKPREEQYX[12]STYRVVSVLTVLHQDWLNGKEYKCKVSNK ALPAPIEKTISKAK (SEQ ID NO: 2203) wherein X^8 and X^9 are selected from Leucine (L) and Alanine (A), X^{10} is selected from Glycine (G) and Alanine (A), and X^{11} is selected from Valine (V) and Tryptophan (W), and X^{12} is selected from Asparagine (N) and Alanine (A). In some aspects, the nucleotide sequence is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identical to any one of the codon-optimized nucleotide sequences of SEQ ID NO: 25-28, or 69-72, wherein the CH2 domain is an IgG2 CH2 domain. In some aspects, the IgG2 CH2 domain comprises the amino acid sequence APPVAGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFNWYVDGX[16]EV HNAKTKPREEQFNSTFRVVSVLTVVHQDWLNGKEYKCKVSNKGLPX[17]X[18]IEKTI SKTK (SEQ ID NO: 2206) wherein X^{16} is selected from Valine (V) and Methionine (M),

X^{17} is selected from Alanine (A) and Serine (S); and X^{18} is selected from Proline (P) and Serine (S). In some aspects, the nucleotide sequence is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to any one of the codon-optimized nucleotide sequences of SEQ ID NO: 37-40, or 81-84, wherein the CH2 domain is an IgG4 CH2 domain In some aspects, the IgG4 CH2 domain comprises the amino acid sequence APEFLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSQEDPEVQFNWYVD GVEVHNAKTKPREEQFNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKGLPSSIEK TISKAK (SEQ ID NO: 2198).

[0180] The present disclosure also provides a polynucleotide comprising a nucleotide sequence encoding a CH3 domain from a therapeutic antibody presented in TABLE 4 or a functional fragment thereof which is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to any one of the codon-optimized nucleotide sequences of SEQ ID NOS: 17-20, 29-32, 41-44, 61-64, 73-76, or 85-88. In some aspects, the nucleotide sequence is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to any one of the codon-optimized nucleotide sequences of SEQ ID NO: 17-20, or 61-64, wherein the CH3 domain is an IgG1 CH3 domain. In some aspects, the IgG1 CH3 domain comprises the amino acid sequence GQPREPQVYTLPPSRX^{13}EX^{14}TKNQVSLTCLVKGFYPSDIAVEWESNGQPE NNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLS LSPG (SEQ ID NO: 2204) wherein X^{13} is selected from Glutamic acid (E) and Aspartic acid (D), and X^{14} is selected from Methionine (M) and Leucine (L). In some aspects, the nucleotide sequence is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to any one of the codon-optimized nucleotide sequences of SEQ ID NO: 29-32, or 73-76, wherein the CH3 domain is an IgG2 CH3 domain. In some aspects, the IgG2 CH3 domain comprises the amino acid sequence GQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTP PMLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG (SEQ ID NO: 2196). In some aspects, the nucleotide sequence is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%,

98%, 99% or 100% identical to any one of the codon-optimized nucleotide sequences of SEQ ID NO: 41-44, or 85-88, wherein the CH3 domain is an IgG4 CH3 domain. In some aspects, the IgG4 CH3 domain comprises the amino acid sequence GQPREPQVYTLPPSQEEMTKNQVSLTCLVKGF YPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSRLTVDKSRWQEGNVFSCSV MHEALHNHYTQKSLSLSLG (SEQ ID NO: 2199).

[0181] The present disclosure also provides a polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to a subsequence from any one of the codon-optimized nucleotide sequences of SEQ ID NOS: 89-1978, which correspond to codon-optimized heavy chains and light chain of therapeutic antibodies known in the art.

[0182] In some aspects, the nucleotide sequences encoding the therapeutic antibodies disclosed herein (e.g., any of the nucleotide sequences encoding the antibodies disclosed in TABLE 4), or functional fragments thereof can be codon-optimized by applying a codon substitution map to the wild type amino acid sequences of the therapeutic antibodies, wherein Ala is encoded by GCC, GCG or GCT; Cys is encoded by TGC or TGT; Asp is encoded by GAC; Glu is encoded by GAG or GAA; Phe is encoded by TTC; Gly is encoded by GGC, GGT, or GGG; His is encoded by CAC; Ile is encoded by ATC or ATT; Lys is encoded by AAG; Leu is encoded by CTG, CTC or TTG; Met is encoded by ATG; Asn is encoded by AAC; Pro is encoded by CCC, CCA or CCG; Gln is encoded by CAG or CAA, Arg is encoded by CGG, AGG, CGC, CGT, AGA, CGA, Ser is encoded by AGC, TCC or TCT, Thr is encoded by ACC, ACG or ACT, Val is encoded by GTG, GTC or GTT, Trp is encoded by TGG, and Tyr is encoded by TAC.

[0183] In other aspects, the nucleotide sequences encoding the therapeutic antibodies disclosed herein (e.g., any of the nucleotide sequences encoding the antibodies disclosed in TABLE 4), or functional fragments thereof is codon-optimized by applying a codon substitution map of TABLE 2, e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof.

[0184] In some aspects, a codon-optimized nucleotide sequence disclosed herein encodes:

(a) one, two, or three VH-CDRs from a therapeutic antibody (e.g., a therapeutic antibody disclosed in TABLE 4);

(b) one, two, or three VL-CDRs from a therapeutic antibody (e.g., a therapeutic antibody disclosed in TABLE 4);

(c) one, two, three, or four VH framework (FW) regions from a therapeutic antibody (e.g., a therapeutic antibody disclosed in TABLE 4);

(d) one, two, three, or four VL framework (FW) regions from a therapeutic antibody (e.g., a therapeutic antibody disclosed in TABLE 4);

(e) a VH domain from a therapeutic antibody (e.g., a therapeutic antibody disclosed in TABLE 4);

(f) a VL domain from a therapeutic antibody (e.g., a therapeutic antibody disclosed in TABLE 4);

(g) a CL domain of a therapeutic antibody (e.g., a therapeutic antibody disclosed in TABLE 4);

(h) a CH1 domain of a therapeutic antibody (e.g., a therapeutic antibody disclosed in TABLE 4);

(i) a CH2 domain of a therapeutic antibody (e.g., a therapeutic antibody disclosed in TABLE 4);

(j) a CH3 domain of a therapeutic antibody (e.g., a therapeutic antibody disclosed in TABLE 4); or,

(k) a combination thereof.

[0185] In some aspects, a codon-optimized nucleotide sequence disclosed herein encoding one, two, three, or four VH framework (FW) regions from a therapeutic antibody (e.g., a therapeutic antibody disclosed in TABLE 4) comprises a codon-optimized nucleotide sequence encoding a first framework region (FW1) of a heavy chain variable domain disclosed herein; and/or a codon-optimized nucleotide sequence a second framework region (FW2) of a heavy chain variable domain disclosed herein; and/or a codon-optimized nucleotide sequence encoding a third framework region (FW3) of a heavy chain variable domain disclosed herein; and/or a codon-optimized nucleotide sequence encoding a fourth framework region (FW4) of a heavy chain variable domain disclosed herein; or any combinations thereof.

[0186] In some aspects, a codon-optimized nucleotide sequence disclosed herein encoding one, two, three, or four VL framework (FW) regions from a therapeutic antibody (e.g., a therapeutic antibody disclosed in TABLE 4) comprises a codon-optimized nucleotide sequence encoding a first framework region (FW1) of a light chain

variable domain disclosed herein; and/or a codon-optimized nucleotide sequence a second framework region (FW2) of a light chain variable domain disclosed herein; and/or a codon-optimized nucleotide sequence encoding a third framework region (FW3) of a light chain variable domain disclosed herein; and/or a codon-optimized nucleotide sequence encoding a fourth framework region (FW4) of a light chain variable domain disclosed herein; or any combinations thereof.

[0187] In some aspects, a codon-optimized nucleotide sequence disclosed herein encoding a CL domain of a therapeutic antibody (e.g., a therapeutic antibody disclosed in TABLE 4) comprises a codon-optimized nucleotide sequence encoding a kappa light chain constant domain of an antibody or a fragment thereof and/or a lambda light chain constant domain of an antibody or a fragment thereof disclosed herein.

[0188] In some aspects, a codon-optimized nucleotide sequence disclosed herein encoding a CH domain of a therapeutic antibody (e.g., a therapeutic antibody disclosed in TABLE 4) comprises a codon-optimized nucleotide sequence encoding a CH1 domain disclosed herein, and/or a codon-optimized nucleotide sequence encoding a CH2 domain disclosed herein; and/or a codon-optimized nucleotide sequence encoding CH3 domain disclosed herein.

[0189] The polynucleotide sequences disclosed herein also comprise nucleotide sequences which are defined in terms of sequence identity to codon-optimized nucleotide sequences of therapeutic antibodies known in the art. TABLE 5 provides a list of heavy chains and lights of therapeutic antibodies, and their respective codon-optimized counterparts. For example, SEQ ID NO:1979 would be the amino acid sequence of the heavy chain of a therapeutic antibody (candidate antibody), and SEQ ID NOS:89-97 would be 9 codon-optimized nucleotide sequences encoding SEQ ID NO:1979. In some aspects, a codon-optimized nucleotide sequence disclosed herein comprises a full sequence from SEQ ID NOs: 2084 to 2188. In other aspects, a codon-optimized nucleotide sequence disclosed herein comprises a subsequence of a sequence from SEQ ID NOs: 2084 to 2188, wherein the subsequence encodes an immunoglobulin domain (e.g., a VH, VL, CL, CH1, CH2, CH3 or a combination thereof).

TABLE 5: Therapeutic antibody heavy chains and lights and their respective codon-optimized nucleotide sequences.

Heavy Chain SEQ ID NO	Optimized SEQ ID NO	Heavy Chain SEQ ID NO	Optimized SEQ ID NO	Light Chain SEQ ID NO	Optimized SEQ ID NO	Light Chain SEQ ID NO	Optimized SEQ ID NO
1979	89-97	2032	566-574	2084	1034-1042	2137	1511-1519
1980	98-106	2033	575-583	2085	1043-1051	2138	1520-1528
1981	107-115	2034	584-592	2086	1052-1060	2139	1529-1537
1982	116-124	2035	593-601	2087	1061-1069	2140	1538-1546
1983	125-133	2036	602-610	2088	1070-1078	2141	1547-1555
1984	134-142	2037	611-619	2089	1079-1087	2142	1556-1564
1985	143-151	2038	620-628	2090	1088-1096	2143	1565-1573
1986	152-160	2039	629-637	2091	1097-1105	2144	1574-1582
1987	161-169	2040	638-646	2092	1106-1114	2145	1583-1591
1988	170-178	2041	647-655	2093	1115-1123	2146	1592-1600
1989	179-187	2042	656-664	2094	1124-1132	2147	1601-1609
1990	188-196	2043	665-673	2095	1133-1141	2148	1610-1618
1991	197-205	2044	674-682	2096	1142-1150	2149	1619-1627
1992	206-214	2045	683-691	2097	1151-1159	2150	1628-1636
1993	215-223	2046	692-700	2098	1160-1168	2151	1637-1645
1994	224-232	2047	701-709	2099	1169-1177	2152	1646-1654
1995	233-241	2048	710-718	2100	1178-1186	2153	1655-1663
1996	242-250	2049	719-727	2101	1187-1195	2154	1664-1672
1997	251-259	2050	728-736	2102	1196-1204	2155	1673-1681
1998	260-268	2051	737-745	2103	1205-1213	2156	1682-1690
1999	269-277	2052	746-754	2104	1214-1222	2157	1691-1699
2000	278-286	2053	755-763	2105	1223-1231	2158	1700-1708
2001	287-295	2054	764-772	2106	1232-1240	2159	1709-1717
2002	296-304	2055	773-781	2107	1241-1249	2160	1718-1726
2003	305-313	2056	782-790	2108	1250-1258	2161	1727-1735
2004	314-322	2057	791-799	2109	1259-1267	2162	1736-1744
2005	323-331	2058	800-808	2110	1268-1276	2163	1745-1753
2006	332-340	2059	809-817	2111	1277-1285	2164	1754-1762
2007	341-349	2060	818-826	2112	1286-1294	2165	1763-1771
2008	350-358	2061	827-835	2113	1295-1303	2166	1772-1780
2009	359-367	2062	836-844	2114	1304-1312	2167	1781-1789
2010	368-376	2063	845-853	2115	1313-1321	2168	1790-1798
2011	377-385	2064	854-862	2116	1322-1330	2169	1799-1807
2012	386-394	2065	863-871	2117	1331-1339	2170	1808-1816
2013	395-403	2066	872-880	2118	1340-1348	2171	1817-1825
2014	404-412	2067	881-889	2119	1349-1357	2172	1826-1834
2015	413-421	2068	890-898	2120	1358-1366	2173	1835-1843
2016	422-430	2069	899-907	2121	1367-1375	2174	1844-1852
2017	431-439	2070	908-916	2122	1376-1384	2175	1853-1861
2018	440-448	2071	917-925	2123	1385-1393	2176	1862-1870
2019	449-457	2072	926-934	2124	1394-1402	2177	1871-1879
2020	458-466	2073	935-943	2125	1403-1411	2178	1880-1888
2021	467-475	2074	944-952	2126	1412-1420	2179	1889-1897
2022	476-484	2075	953-961	2127	1421-1429	2180	1898-1906
2023	485-493	2076	962-970	2128	1430-1438	2181	1907-1915
2024	494-502	2077	971-979	2129	1439-1447	2182	1916-1924
2025	503-511	2078	980-988	2130	1448-1456	2183	1925-1933
2026	512-520	2079	989-997	2131	1457-1465	2184	1934-1942
2027	521-529	2080	998-1006	2132	1466-1474	2185	1943-1951
2028	530-538	2081	1007-1015	2133	1475-1483	2186	1952-1960
2029	539-547	2082	1016-1024	2134	1484-1492	2187	1961-1969
2030	548-556	2083	1025-1033	2135	1493-1501	2188	1970-1978
2031	557-565			2136	1502-1510		

[0190] Thus, the present disclosure also provides a polynucleotide comprising a nucleotide sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to any one of the codon-optimized nucleotide sequences of SEQ ID NOs: 1-4 or 45-48, wherein the nucleotide sequence encodes a CL kappa domain having an amino acid sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to a representative CL kappa domain (SEQ ID NO: 2189) of a therapeutic antibody disclosed herein.

[0191] Also provided is a polynucleotide comprising a nucleotide sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to any one of the codon-optimized nucleotide sequences of SEQ ID NOs: 5-8 or 49-52, wherein the nucleotide sequence encodes a CL lambda domain having an amino acid sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to a representative CL lambda domain (SEQ ID NO: 2190) of a therapeutic antibody disclosed herein.

[0192] In some aspects, the representative CL lambda or CL kappa domain comprises the CL domain of a therapeutic antibody light chain selected from SEQ ID NOs: 2084 to 2188.

[0193] Also provided is a polynucleotide comprising a nucleotide sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to

(i) SEQ ID NOs: 9-12 or 53-56, wherein the nucleotide sequence encodes an amino acid sequence about 80%, 81%, 82%, 83%, 94%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO: 2191, wherein SEQ ID NO: 2191 is the amino acid sequence of a representative CH1 domain from an IgG1 therapeutic antibody disclosed herein;

(ii) SEQ ID NOs: 13-16 or 57-60, wherein the nucleotide sequence encodes an amino acid sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO: 2192, wherein SEQ ID NO: 2192 is the amino acid sequence of a representative CH2 domain from an IgG1 therapeutic antibody disclosed herein;

(iii) SEQ ID NOs: 17-20 or 61-64, wherein the nucleotide sequence encodes an amino acid sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO: 2193, wherein SEQ ID NO: 2193 is the amino acid sequence of a representative CH3 domain from an IgG1 therapeutic antibody disclosed herein;

(iv) SEQ ID NOs: 21-24 or 65-68, wherein the nucleotide sequence encodes an amino acid sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO: 2194, wherein SEQ ID NO: 2194 is the amino acid sequence of a representative CH1 domain from an IgG2 therapeutic antibody disclosed herein;

(v) SEQ ID NOs: 25-28 or 69-72, wherein the nucleotide sequence encodes an amino acid sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO: 2195, wherein SEQ ID NO: 2195 is the amino acid sequence of a representative CH2 domain from an IgG2 therapeutic antibody disclosed herein;

(vi) SEQ ID NOs: 29-32 or 73-76, wherein the nucleotide sequence encodes an amino acid sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO: 2196, wherein SEQ ID NO: 2196 is the amino acid sequence of a representative CH3 domain from an IgG2 therapeutic antibody disclosed herein;

(vii) SEQ ID NOs: 33-36 or 77-80, wherein the nucleotide sequence encodes an amino acid sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO: 2197, wherein SEQ ID NO: 2197 is the amino acid sequence of a representative CH1 domain from an IgG4 therapeutic antibody disclosed herein;

(viii) SEQ ID NOs: 37-40 or 81-84, wherein the nucleotide sequence encodes an amino acid sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO: 2198, wherein SEQ ID NO: 2197 is the amino acid sequence of a representative CH2 domain from an IgG4 therapeutic antibody disclosed herein;

(ix) SEQ ID NOs: 41-44 or 85-88, wherein the nucleotide sequence encodes an amino acid sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID

NO: 2199, wherein SEQ ID NO: 2199 is the amino acid sequence of a representative CH3 domain from an IgG4 therapeutic antibody disclosed herein; or

(x) a combination thereof.

[0194] The present disclosure also provides a polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to a nucleotide subsequence of a codon-optimized nucleotide sequence encoding a therapeutic antibody disclosed herein (e.g. a codon-optimized nucleotide sequence encoding a therapeutic antibody heavy chain of SEQ ID Nos:89-1033, or a codon-optimized therapeutic nucleotide sequence encoding a therapeutic antibody light chain of SEQ ID NO:1034-1978) wherein the nucleotide subsequence encodes a variable region (VH or VL) protein sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to the corresponding VH or VL region of the candidate antibody sequence (i.e., amino acid SEQ ID NOs:2084-2188, wherein SEQ ID NOs:1979-2083 correspond to heavy chains, and SEQ ID NOs:2084-2188 correspond to light chains).

[0195] The present disclosure also provides nucleotide sequences that are about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to a corresponding codon-optimized nucleotide sequence disclosed in TABLE 5. In some aspects, such nucleotide sequence is a subsequence or a concatenated set of subsequences of one or more codon-optimized sequences disclosed in TABLE 5, e.g., a nucleotide sequence encoding a VH domain, a VL domain, a CL domain, a CH1 domain, a CH2 domain, a CH3 domain, or a combination thereof (e.g., an scFv). The boundaries between these different structural elements can be determined according to FIG. 11 and FIG. 12. FIG. 11 presents a multiple sequence alignment of all the light chain amino acid sequences presented in TABLE 5 (SEQ ID NOs: 2084-2188), whereas FIG. 12 present a multiple sequence alignment of all the heavy chain amino acid sequences presented in TABLE 5 (SEQ ID NOs: 1979-2083). As discussed above, the boundaries between structural elements in an antibody sequence can also be determined according to alternative methods known in the art.

[0196] In some aspects, VH and/or VL domains from the sequences disclosed in TABLE 5 can be combined to yield bispecific, trispecific, tetraspecific, o multispecific antibody

constructs. In some aspects, VH and/or VL domains from the sequences disclosed in TABLE 5 can be combined to yield bifunctional, trifunctional, tetrafunctional, or multifunctional antibody constructs.

[0197] In some aspects, the codon-optimized nucleotide subsequences of the heavy chains and light chains disclosed in TABLE 5 encoding a VH domain, a VL domain, a CL domain, a CH1 domain, a CH2 domain, a CH3 domain, or a combination thereof (as well as variants comprising 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 codon substitutions), can be assembled to generate a polynucleotide sequence or set of polynucleotide sequences encoding an antibody construct known in the art, e.g., the antibody constructs presented in FIG. 14, e.g., an scFv, an scFav, a minibody, an scDv-Fc, a diabody, an sc-diabody, a ZIP miniantibody, an $(scFv)_2$/BITE, a $(Fab)_2$/$sc(Fab)_2$, a V_{HH}, a triabody. a tribody, a tribi-minibody, a collabody, a $(Fab)_3$/DNL, a tetrabody, a tandem diabody (tandab), an $[sc(Fv)_2]_2$, a di-diabody, etc.

[0198] In some aspects, the polynucleotide sequences disclosed above can comprise a nucleotide sequence encoding a linker. In some aspects, the nucleotide sequence encoding a linker is codon-optimized. In some aspects, the polynucleotide comprising a nucleotide sequence encoding a linker encodes an scFv.

c. Codon-Optimized Nucleotide Sequences Defined by Consensus Sequences

[0199] The codon-optimized nucleotide sequences presented in the instant disclosure can also be described with respect to consensus sequences identified in therapeutic antibodies known in the art. The term "consensus sequence," as used herein refers to a composite or genericized sequence defined based on information as to which amino acid residues within the sequence are amenable to modification without detriment to antigen binding. This information can be obtained from multiple sequence alignments according to methods known in the art. Thus, in a "consensus sequence" for a VL or VH chain, certain amino acid positions are occupied by one of multiple possible amino acid residues at that position. Amino acid residues that can be occupied by various amino acid residues are represented as X_n in the consensus sequences presented below. For example, if an Arginine (R) or a Serine (S) is present at a particular position in the multiple sequence alignment, then that particular position within the consensus sequence can be either Arginine or Serine (R or S).

[0200] The phrase "a polynucleotide comprising a consensus nucleotide sequence" means that the polynucleotide can comprise any of the nucleotide sequences described by the consensus nucleotide sequence.

[0201] The present disclosure provides a polynucleotide comprising a consensus nucleotide sequence corresponding to a lambda light chain constant domain of an antibody or a fragment thereof. Accordingly, the disclosure provides a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes TVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQES VTEQDSKDSTYSLSX[1]TLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNR (SEQ ID NO: 2200), wherein X^1 is selected from Asparagine (N) and Serine (S). In some aspects, the nucleotide sequence encodes SEQ ID NO:2189, or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO:2189. In some aspects, the nucleotide sequence encodes a variant identical to SEQ ID NO:2189 except for 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 mutations.

[0202] The disclosure also provides a polynucleotide comprising a consensus nucleotide sequence corresponding to a kappa light chain constant domain of an antibody or a fragment thereof. Thus, the disclosure provides a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes PKAAPSVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTT PSKQSNNKYAASSYLSLTPEQWKSHX[2]SYSCQVTHEGSTVEKTVAPX[3]ECS (SEQ ID NO: 2201),

wherein X^2 is selected from Arginine (R) and Lysine (K), and X^3 is selected from Threonine (T) and Alanine (A). In some aspects, the nucleotide sequence encodes SEQ ID NO: 2190, or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO:2190. In some aspects, the nucleotide sequence encodes a variant identical to SEQ

ID NO:2190 except for 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 mutations.

[0203] The disclosure also provides a polynucleotide comprising a consensus nucleotide sequence corresponding to a CH1 domain of an IgG1 antibody or a fragment thereof. Accordingly, the disclosure provides a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $SX^4GPSVX^5PLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGVHTFPAVLQSSG$ $LYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKX^6X^7$ (SEQ ID NO: 2202) wherein X^4 is an optional ASTK sequence, X^5 is selected from Phenylalanine (F) and Leucine (L), X^6 is selected from Lysine (K) and Arginine (R), and X^7 is selected from Valine (V) and Alanine (A). In some aspects, the nucleotide sequence encodes SEQ ID NO: 2191, or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO: 2191. In some aspects, the nucleotide sequence encodes a variant identical to SEQ ID NO:2191 except for 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 mutations.

[0204] Also provided is a polynucleotide comprising a consensus nucleotide sequence corresponding to a CH2 domain of an IgG1 antibody or a fragment thereof. In this respect, the disclosure provides a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $APEX^8X^9GX^{10}PSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNX^{11}YVDGV$ $EVHNAKTKPREEQYX^{12}STYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK$ TISKAK (SEQ ID NO: 2203) wherein X^8 and X^9 are selected from Leucine (L) and Alanine (A), X^{10} is selected from Glycine (G) and Alanine (A), and X^{11} is selected from Valine (V) and Tryptophan (W), and X^{12} is selected from Asparagine (N) and Alanine (A). In some aspects, the nucleotide sequence encodes SEQ ID NO: 2192, or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO: 2192. In some

aspects, the nucleotide sequence encodes a variant identical to SEQ ID NO:2192 except for 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 mutations.

[0205] Also provided is a polynucleotide comprising a consensus nucleotide sequence corresponding to a CH3 domain of an IgG1 antibody or a fragment thereof. Accordingly, the disclosure provides a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes GQPREPQVYTLPPSRX^{13}EX^{14}TKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYK TTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG (SEQ ID NO: 2204) wherein X^{13} is selected from Glutamic acid (E) and Aspartic acid (D), and X^{14} is selected from Methionine (M) and Leucine (L). In some aspects, the nucleotide sequence encodes SEQ ID NO: 2193, or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO: 2193. In some aspects, the nucleotide sequence encodes a variant identical to SEQ ID NO:2193 except for 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 mutations.

[0206] The disclosure also provides a polynucleotide comprising a consensus nucleotide sequence corresponding to a CH1 domain of an IgG2 antibody or a fragment thereof. Accordingly, the disclosure provides a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes SASTKGPSVFPLAPCSRSTSESTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPA VLQSSGLYSLSSVVTVX^{15}SSNFGTQTYTCNVDHKPSNTKVDKTV (SEQ ID NO: 2205) wherein X^{15} is selected from Proline (P) and Threonine (T). In some aspects, the nucleotide sequence encodes SEQ ID NO: 2194, or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO: 2194. In some aspects, the nucleotide sequence encodes a variant identical to SEQ ID NO: 2194 except for 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 mutations.

[0207] The disclosure also provides a polynucleotide comprising a consensus nucleotide sequence corresponding to a CH2 domain of an IgG2 antibody or a fragment thereof.

Accordingly, the disclosure provides polynucleotides comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16or any combination thereof), wherein the nucleotide sequence encodes APPVAGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFNWYVDGX^{16}EV HNAKTKPREEQFNSTFRVVSVLTVVHQDWLNGKEYKCKVSNKGLPX^{17}X^{18}IEKTI SKTK (SEQ ID NO: 2206) wherein X^{16} is selected from Valine (V) and Methionine (M), X^{17} is selected from Alanine (A) and Serine (S); and X^{18} is selected from Proline (P) and Serine (S). In some aspects, the nucleotide sequence encodes SEQ ID NO: 2195, or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO: 2195. In some aspects, the nucleotide sequence encodes a variant identical to SEQ ID NO: 2195 except for 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 mutations.

[0208] Also provided is a polynucleotide comprising a consensus nucleotide sequence corresponding to a CH3 domain of an IgG2 antibody or a fragment thereof. Accordingly, the disclosure provides a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes GQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTP PMLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG (SEQ ID NO: 2196). In some aspects, the nucleotide sequence encodes a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO: 2196. In some aspects, the nucleotide sequence encodes a variant identical to SEQ ID NO: 2196 except for 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 mutations.

[0209] Also provided is a polynucleotide comprising a consensus nucleotide sequence corresponding to a CH1 domain of an IgG4 antibody or a fragment thereof. Accordingly, the disclosure provides a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes SASTKGPSVFPLAPCSRSTSESTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPA

VLQSSGLYSLSSVVTVPSSSLGTKTYTCNVDHKPSNTKVDKRV (SEQ ID NO: 2197). In some aspects, the nucleotide sequence encodes a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO: 2197. In some aspects, the nucleotide sequence encodes a variant identical to SEQ ID NO: 2197 except for 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 mutations.

[0210] Also provided is a polynucleotide comprising a consensus nucleotide sequence corresponding to a CH2 domain of an IgG4 antibody or a fragment thereof. Thus, the description provides a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes APEFLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSQEDPEVQFNWYVDGVEVH NAKTKPREEQFNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKGLPSSIEKTISKA K (SEQ ID NO: 2198). In some aspects, the nucleotide sequence encodes a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO: 2198. In some aspects, the nucleotide sequence encodes a variant identical to SEQ ID NO: 2198 except for 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 mutations.

[0211] Also provided is a polynucleotide comprising a consensus nucleotide sequence corresponding to a CH3 domain of an IgG4 antibody or a fragment thereof. Accordingly, the disclosure also provides a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes GQPREPQVYTLPPSQEEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTP PVLDSDGSFFLYSRLTVDKSRWQEGNVFSCSVMHEALHNHYTQKSLSLSLG (SEQ ID NO: 2199). In some aspects, the nucleotide sequence encodes a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO: 2199. In some aspects, the nucleotide sequence encodes a variant identical to SEQ ID NO: 2199 except for 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 mutations.

[0212] In addition to the consensus nucleotide sequence defining constant regions of therapeutic antibodies and fragment thereof, the present disclosure also provides consensus sequences defining the variable regions of therapeutic antibodies, in particular, consensus sequences defining their framework regions. In one aspect, the present disclosure provides consensus sequences defining the framework regions of lambda light chains, as shown below.

[0213] The disclosure provides a polynucleotide comprising a consensus nucleotide sequence corresponding to the first framework region (FW1) of a lambda light chain variable domain. Accordingly, the description provides a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $X^1X^2X^3LTQX^4X^5X^6VSX^7X^8X^9GX^{10}X^{11}X^{12}X^{13}X^{14}X^{15}C$ (SEQ ID NO: 2235) wherein X^1 is selected from Q, D, E and S; X^2 is selected from S, I, A, and Y; X^3 is selected from V, Q, A, and E; X^4 is selected from P and D; X^5 is selected from P, N, and A; X^6 is selected from S and A; X^7 is selected from G, T, A, and V; X^8 is selected from A and S; X^9 is selected from P and L; X^{10} is selected from Q, K, and S; X^{11} is selected from R, K, T, and S; X^{12} is selected from V, I, and A; X^{13} is selected from T, K, and R; X^{14} is selected from I and L; and, X^{15} is selected from S at T. In some aspects, the nucleotide sequence encodes a sequence identical to QSVLTQPPSVSGAPGQRVTISC (SEQ ID NO: 2207) except for at least one, two, three, four or five substitutions selected from Q1(DES), S2(IAY), V3(QAE), P7D, P8(NA), S9A, G12(TAV), A13S, P14L, Q16(KS), R17(KTS), V18(IA), T19(KR), I20L, and S21T. In some aspects, the nucleotide sequence encodes QSVLTQPPSVSGAPGQRVTISC (SEQ ID NO: 2207), or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO: 2207.

[0214] Also provided is a polynucleotide comprising a consensus nucleotide sequence corresponding to the second framework region (FW2) of a lambda light chain variable domain. Accordingly, the disclosure provides a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $WYQX^1X^2X^3GX^4X^5PX^6X^7X^8I$ (SEQ ID NO: 2236) wherein X^1 is selected from

Q and L; X^2 is selected from L,Y, H, and K; X^3 is selected from P and E; X^4 is selected from T, R, K, and Q; X^5 is selected from A and S; X^6 is selected from K, T, V and I; X^7 is selected from L and T; and X^8 is selected from L, M, and V. In some aspects, the nucleotide sequence encodes a sequence identical to WYQQLPGTAPKLLI (SEQ ID NO: 2208) except for at least one, two, three, four or five substitutions selected from Q4L, L5(YHK), P6E, T8(RKQ), A9S, K11(TVI), L12T, and L13(MV). In some aspects, the nucleotide sequence encodes WYQQLPGTAPKLL (SEQ ID NO: 2208), or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO: 2208.

[0215] Also provided is a polynucleotide comprising a consensus nucleotide sequence corresponding to the third framework region (FW3) of a lambda light chain variable domain. Accordingly, the disclosure provides a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $RFSGSX^1SX^2X^3X^4AX^5LX^6IX^7X^8X^9X^{10}X^{11}X^{12}DEAX^{13}YX^{14}C$ (SEQ ID NO: 2237) wherein X^1 is selected from K, N, S, and I; X^2 is selected from G and S; X^3 is selected from T and N; X^4 is selected from S and T; X^5 is selected from S, T, and F; X^6 is selected from A, T, and G; X^7 is selected from T, H, and S; X^8 is selected from G, N, and R; X^9 is selected from L, V, and A; X^{10} is selected from Q, E, and A; X^{11} is selected from A, T, and I; X^{12} is selected from E and G; X^{13} is selected from D and I; and, X^{14} is selected from Y and F. In some aspects, the nucleotide sequence encodes a sequence identical to RFSGSKSGTSASLAITGLQAEDEADYYC (SEQ ID NO: 2209) except for at least one, two, three, four or five substitutions selected from K6(NSI), G8S, T9N, S10T, S12(TF), A14(TG), T16(HS), G17(NR), L18(VA), Q19(EA), A20(TI), E21G, D25I, and Y27F. In some aspects, the nucleotide sequence encodes RFSGSKSGTSASLAITGLQAEDEADYYC (SEQ ID NO: 2209), or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO:2209.

[0216] Also provided is a polynucleotide comprising a consensus nucleotide sequence corresponding to the fourth framework region (FW4) of a lambda light chain variable domain. Accordingly, the disclosure provides a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6,

MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $FGX^1GTX^2X^3TVL$ (SEQ ID NO:2238) wherein X^1 is selected from G and T; X^2 is selected from K and Q; and X^3 is selected from L and V. In some aspects, the nucleotide sequence encodes a sequence identical to FGGGTKLTVL (SEQ ID NO: 2210) except for at least one, two, or three substitutions selected from G3T, K6Q, and L7V. In some aspects, the nucleotide sequence encodes FGGGTKLTVL (SEQ ID NO: 2210), or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO:2210.

[0217] In another aspect, the present disclosure provides consensus sequences defining the framework regions of kappa light chains. Clustering analysis indicates that framework regions of kappa light chains can be defined according to three different consensus sequences (analysis not shown). Thus, the disclosure provides polynucleotides comprising at least one of three consensus nucleotide sequences defining the first framework region (FW1) of a kappa light chain variable domain; at least one of three consensus nucleotide sequences defining the second framework region (FW2) of a kappa light chain variable domain; at least one of three consensus nucleotide sequences defining the third framework region (FW3) of a kappa light chain variable domain; and at least one of three consensus nucleotide sequences defining the fourth framework region (FW4) of a kappa light chain variable domain

[0218] Accordingly, in one aspect, the disclosure provides a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $X^1X^2QX^3TQX^4X^5SX^6X^7SASX^8CDRVTX^9X^{10}C$ (SEQ ID NO: 2239) (LC kappa, FW1, consensus sequence 1), wherein X^1 is selected from D and A; X^2 is selected from I and V; X^3 is selected from M, L, and V; X^4 is selected from S and F; X^5 is selected from P and T; X^6 is selected from S and T; X^7 is selected from L and V; X^8 is selected from V, I, and A; X^9 is selected from I and M; and, X^{10} is selected from T and S. In some aspects, the nucleotide sequence encodes a sequence identical to DIQMTQSPSSLSASVCDRVTITC (SEQ ID NO: 2211) except for at least one, two, three, four or five substitutions selected from D1A, I2V, M4(LV), S7F, P8T, S10T, L11V, V15(IA), I21M, and T22S. In some aspects, the nucleotide sequence encodes

DIQMTQSPSSLSASVCDRVTITC (SEQ ID NO: 2211), or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO:2211.

[0219] In another aspect, the disclosure provides a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $DX^1X^2X^3TQX^4PX^5SX^6X^7X^8X^9X^{10}GX^{11}X^{12}X^{13}X^{14}X^{15}X^{16}C$ (SEQ ID NO: 2243) (LC kappa, FW1, consensus sequence 2) wherein X^1 is selected from I and V; X^2 is selected from V, L, and Q; X^3 is selected from M and L; X^4 is selected from S and T; X^5 is selected from L and D; X^6 is selected from L and V; X^7 is selected from P, S and A; X^8 is selected from V and M; X^9 is selected from T and S; X^{10} is selected from P and L; X^{11} is selected from E and Q; X^{12} is selected from P and R; X^{13} is selected from A and V; X^{14} is selected from S and T; X^{15} is selected from I, M, and L; and X^{16} is selected from S and N. In some aspects, the nucleotide sequence encodes a sequence identical to DIVMTQSPLSLPVTPGEPASISC (SEQ ID NO: 2215) except for at least one, two, three, four, or five substitutions selected from I2V, V3(LQ), M4L, S7T, L9D, L11V, P12(SA), V13M, T14S, P15L, E17Q, P18R, A19V, S20T, I21(ML), and S22N. In some aspects, the nucleotide sequence encodes DIVMTQSPLSLPVTPGEPASISC (SEQ ID NO: 2215), or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO:2215.

[0220] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $X^1X^2VX^3TQSPX^4TLSX^5SPGERATLSC$ (SEQ ID NO: 2247) (LC kappa, FW1, consensus sequence 3) wherein X^1 is selected from E and D; X^2 is selected from I and T; X^3 is selected from L and M; X^4 is selected from G and A; and, X^5 is selected from L and V. In some aspects, the nucleotide sequence encodes a sequence identical to EIVLTQSPGTLSLSPGERATLSC (SEQ ID NO: 2219) except for at least one, two, three, four, or five substitutions selected from E1D, I2T, L4M, G9A, and L13V. In some aspects, the nucleotide sequence encodes EIVLTQSPGTLSLSPGERATLSC (SEQ ID

NO: 2219), or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO:2215.

[0221] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $WX^1X^2X^3X^4PX^5KX^6X^7X^8X^9X^{10}IX^{11}$ (SEQ ID NO: 2240) (LC kappa, FW2, consensus sequence 1) wherein X^1 is selected from Y and F; X^2 is selected from Q and L; X^3 is selected from Q and H; X^4 is selected from K and I; X^5 is selected from G and E; X^6 is selected from A and V; X^7 is selected from P and V; X^8 is selected from K and Q; X^9 is selected from L, T, S, R, P, and V; X^{10} is selected from L and W; and, X^{11} is selected from Y and S. In some aspects, the nucleotide sequence encodes a sequence identical to WYQQKPGKAPKLLIY (SEQ ID NO: 2212) except for at least one, two, three, four, or five substitutions selected from Y2F, Q3L, Q4H, K5I, G7E, A9V, P10V, K11Q, L12(TSRPV), L13W, and Y15S. In some aspects, the nucleotide sequence encodes WYQQKPGKAPKLLIY (SEQ ID NO: 2212), or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO:2212.

[0222] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $WX^1X^2QX^3X^4GQX^5PX^6X^7LIX^8$ (SEQ ID NO: 2244) (LC kappa, FW2, consensus sequence 2) wherein X^1 is selected from Y, F, and W; X^2 is selected from L and Q; X^3 is selected from K and R; X^4 is selected from P and S; X^5 is selected from S and P; X^6 is selected from Q, K, R, and N; X^7 is selected from L and R; and, X^8 is selected from Y and W. In some aspects, the nucleotide sequence encodes a sequence identical to WYLQKPGQSPQLLIY (SEQ ID NO: 2216) except for at least one, two, three, four or five substitutions selected from Y2(FW), L3Q, K5R, P6S, S9P,Q11(KRN), L12R, and Y15W. In some aspects, the nucleotide sequence encodes WYLQKPGQSPQLLIY (SEQ ID NO: 2216), or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%,

89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO:2216..

[0223] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $WX^1X^2QX^3PGQAPRX^4LIX^5$ (SEQ ID NO: 2248) (LC kappa, FW2, consensus sequence 3) wherein X^1 is selected from Y and F; X^2 is selected from Q and R; X^3 is selected from K and R; X^4 is selected from L and P; and X^5 is selected from Y, R, and K. In some aspects, the nucleotide sequence encodes a sequence identical to WYQQKPGQAPRLLIY (SEQ ID NO: 2220) except for at least one, two, three, four or five substitutions selected from Y2F, Q3R, K5R, L12P, and Y15(RK). In some aspects, the nucleotide sequence encodes WYQQKPGQAPRLLIY (SEQ ID NO: 2220), or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO:2215.

[0224] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $RFSGSX^1SGX^2X^3X^4X^5X^6TISSLX^7X^8X^9DX^{10}AX^{11}YX^{12}C$ (SEQ ID NO: 2241) (LC kappa, FW3, consensus sequence 1) wherein X^1 is selected from G and R; X^2 is selected from T and Q; X^3 is selected from D, E, and Y; X^4 is selected from F and Y; X^5 is selected from T and S; X^6 is selected from L and F; X^7 is selected from Q and E; X^8 is selected from P, Q, A, and S; X^9 is selected from E and D; X^{10} is selected from F, I, S, L, V, and T; X^{11} is selected from T, S, and V; and, X^{12} is selected from Y and F. In some aspects, the nucleotide sequence encodes a sequence identical to RFSGSGSGTDFTLTISSLQPEDFATYYC (SEQ ID NO: 2213) except for at least one, two, three, four, or five substitutions selected from G6R, T9Q, D10(EY), F11Y, T12S, L13F, Q19E, P20(QAS), E21D, F23(ISLVT), T25(SV), and Y27F. In some aspect, the nucleotide sequence encodes RFSGSGSGTDFTLTISSLQPEDFATYYC (SEQ ID NO: 2213), or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO:2213.

[0225] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes RFSGSGSX^{1}TX^{2}FTLX^{3}ISX^{4}X^{5}X^{6}AX^{7}DVX^{8}X^{9}X^{10}X^{11}C (SEQ ID NO: 2245) (LC kappa, FW3, consensus sequence 2)wherein X^1 is selected from G and A; X^2 is selected from D and A; X^3 is selected from K, R, and T; X^4 is selected from R and S; X^5 is selected from V and L; X^6 is selected from E and Q; X^7 is selected from E and Q; X^8 is selected from G and A; X^9 is selected from V, D, and F; X^{10} is selected from Y and W; and, X^{11} is selected from Y, F, and W. In one aspect, the nucleotide sequence encodes a sequence identical to RFSGSGSGTDFTLKISRVEAEDVGVYYC (SEQ ID NO: 2217) except for at least one, two, three, four or five substitutions selected from G8A, D10A, K14(RT), R17S, V18L, E19Q, E21Q, G24A, V25(DF), Y26W, and Y27(FW). In one aspect, the nucleotide sequence encodes RFSGSGSGTDFTLKISRVEAEDVGVYYC (SEQ ID NO: 2217),or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO:2217.

[0226] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes RFSGSGSGTX^{1}X^{2}TLTISX^{3}LX^{4}X^{5}EDFAX^{6}X^{7}YC (SEQ ID NO: 2249) (LC kappa, FW3, consensus sequence 3) wherein X^1 is selected from D and E; X^2 is selected from F and S; X^3 is selected from R and S; X^4 is selected from E and Q; X^5 is selected from P and S; X^6 is selected from V and T; and, X^7 is selected from Y and F. In one aspect, the nucleotide sequence encodes a sequence identical to RFSGSGSGTDFTLTISRLEPEDFAVYYC (SEQ ID NO: 2221) except for at least one, two, three, four, or five substitutions selected from D10E, F11S, R17S, E19Q, P20S, V25T, and Y26F. In one aspect, the nucleotide sequence encodes RFSGSGSGTDFTLTISRLEPEDFAVYYC (SEQ ID NO: 2221), or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO:2221.

[0227] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5,

MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $X^1GX^2GTX^3X^4X^5X^6X^7$ (SEQ ID NO: 2242) (LC kappa, FW4, consensus sequence 1) wherein X^1 is selected from F and L; X^2 is selected from Q, G, and S; X^3 is selected from K and R; X^4 is selected from V and L; X^5 is selected from E, D, and Q; X^6 is selected from I and V; and, X^7 is selected from K and T. In some aspects, the nucleotide sequence encodes a sequence identical to FGQGTKVEIK (SEQ ID NO: 2214) except for at least one, two, three, four or five substitutions selected from F1L, Q3(GS), K6R, V7L, E8(DQ), I9V, and K10T. In some aspects, the nucleotide sequence encodes FGQGTKVEIK (SEQ ID NO: 2214), or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO:2214.

[0228] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $FGX^1GTX^2X^3X^4X^5K$ (SEQ ID NO: 2246) (LC kappa, FW4, consensus sequence 2) wherein X^1 is selected from Q, A, P, and G; X^2 is selected from K and R; X^3 is selected from V and L; X^4 is selected from E and Q; and X^5 is selected from I and L. In some aspects, the nucleotide sequence encodes a sequence identical to FGQGTKVEIK (SEQ ID NO: 2218) except for at least one, two, three, four, or five substitutions selected from Q3(APG), K6R, V7L, E8Q, and I9L. In some aspects, the nucleotide sequence encodes FGQGTKVEIK (SEQ ID NO: 2218), or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO:2218.

[0229] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $FX^1X^2GTX^3X^4X^5IK$ (SEQ ID NO: 2250) (LC kappa, FW4, consensus sequence 3) wherein X^1 is selected from G and C; X^2 is selected from Q, G, and P; X^3 is selected from K and R; X^4 is selected from V, L, and A; and, X^5 is selected from E and D. In some aspects, the nucleotide sequence encodes a sequence identical to FGQGTKVEIK (SEQ

ID NO: 2222) except for at least one, two, three, four or five substitutions selected from G2C, Q3(GP), K6R, V7(LA), and E8D. In some aspects, the nucleotide sequence encodes FGQGTKVEIK (SEQ ID NO: 2222), or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO:2222.

[0230] In another aspect, the present disclosure provides consensus sequences defining the framework regions of heavy chains. As in the case of the kappa light chains, clustering analysis indicates that the framework regions of heavy chains can be defined according to three different consensus sequences (analysis not shown). Thus, the disclosure provides polynucleotides comprising at least one of three consensus nucleotide sequences defining the first framework region (FW1) of a heavy chain variable domain; at least one of three consensus nucleotide sequences defining the second framework region (FW2) of a heavy chain variable domain; at least one of three consensus nucleotide sequences defining the third framework region (FW3) of a heavy chain variable domain; and at least one of three consensus nucleotide sequences defining the fourth framework region (FW4) of a heavy chain variable domain

[0231] Accordingly, in one aspect, the disclosure provides a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $X^1X^2X^3X^4X^5X^6SGGX^7X^8X^9X^{10}X^{11}GX^{12}SX^{13}X^{14}LX^{15}C$ (SEQ ID NO: 2251) (HC, FW1, consensus sequence 1) wherein X^1 is selected from E, D, and Q; X^2 is selected from V and A; X^3 is selected from Q, E, and K; X^4 is selected from L and V; X^5 is selected from V and L; X^6 is selected from E and Q; X^7 is selected from G, K, and D; X^8 is selected from L and V; X^9 is selected from V, L, and E; X^{10} is selected from Q, R and K; X^{11} is selected from P, S, and L; X^{12} is selected from G and R; X^{13} is selected from L and R; X^{14} is selected from R and K; and, X^{15} is selected from S and D. In some aspects, the nucleotide sequence encodes a sequence identical to EVQLVESGGGLVQPGGSLRLSC (SEQ ID NO: 2223) except for at least one, two, three, four or five substitutions selected from E1(DQ), V2A, Q3(EK), L4V, V5L, E6Q, G10(KD), L11V, V12(LE), Q13(RK), P14(SL), G16R, L18R, R19K, and S21D. In some aspects, the nucleotide sequence encodes EVQLVESGGGLVQPGGSLRLSC (SEQ ID NO: 2223) , or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%,

89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO:2223.

[0232] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $X^1X^2QLX^3QX^4GX^5X^6X^7X^8X^9X^{10}GX^{11}X^{12}X^{13}X^{14}X^{15}SC$ (SEQ ID NO: 2255) (HC, FW1, consensus sequence 2) wherein X^1 is selected from Q and E; X^2 is selected from V and I; X^3 is selected from V and Q; X^4 is selected from S and P; X^5 is selected from A, S, V, P, T, and G; X^6 is selected from E, G and V; X^7 is selected from V and L; X^8 is selected from K, V, E, and A; X^9 is selected from K, R and Q; X^{10} is selected from P and S; X^{11} is selected from A, E, S, T, and R; X^{12} is selected from S and T; X^{13} is selected from V and L; X^{14} is selected from K and R; and, X^{15} is selected from V, I, L, and M. In some aspects, the nucleotide sequence encodes a sequence identical to QVQLVQSGAEVKKPGASVKVSC (SEQ ID NO: 2227) except for at least one, two, three, four, or five substitution selected from Q1E, V2I, V5Q, S7P, A9(SVPTG), E10(GV), V11L, K12(VEA), K13(RQ), P14S, A16(ESTR), S17T, V18L, K19R, and V20(ILM). In some aspects, the nucleotide sequence encodes QVQLVQSGAEVKKPGASVKVSC (SEQ ID NO: 2227), or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO:2227.

[0233] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $QX^1X^2LX^3X^4X^5GX^6X^7LX^8X^9PX^{10}X^{11}TLX^{12}LTC$ (SEQ ID NO: 2259) (HC, FW1, consensus sequence 3) wherein X^1 is selected from V and L; X^2 is selected from Q and T; X^3 is selected from Q and R; X^4 is selected from E and Q; X^5 is selected from S and W; X^6 is selected from P and A; X^7 is selected from G and A; X^8 is selected from V and L; X^9 is selected from K and R; X^{10} is selected from S and T; X^{11} is selected from Q and E; and, X^{12} is selected from S and T. In some aspects, the nucleotide sequence encodes a sequence identical to QVQLQESGPGLVKPSQTLSLTC (SEQ ID NO: 2231) except for at least one, two, three, four, or five substitutions selected from V2L, Q3T, Q5R, E6Q,

S7W, P9A, G10A, V12L, K13R, S15T, Q16E, and S19T. In some aspects, the nucleotide sequence encodes QVQLQESGPGLVKPSQTLSLTC (SEQ ID NO: 2231), or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO:2231.

[0234] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $WX^1RQX^2PX^3KX^4LX^5X^6X^7X^8$ (SEQ ID NO: 2252) (HC, FW2, consensus sequence 1) wherein X^1 is selected from V, I, and F; X^2 is selected from A, S and T; X^3 is selected from G and E; X^4 is selected from G and R; X^5 is selected from E and D; X^6 is selected from W and L; X^7 is selected from V and I; and, X^8 is selected from A, S, and G. In some aspects, the nucleotide sequence encodes a sequence identical to WVRQAPGKGLEWVA (SEQ ID NO: 2224) except for at least one, two, three, four, or five substitution selected from V2(IF), A5(ST), G7E, G9R, E11D, W12L, V13I, and A14(SG). In some aspects, the nucleotide sequence encodes WVRQAPGKGLEWVA (SEQ ID NO: 2224), or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO:2224.

[0235] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $WX^1X^2QX^3X^4GX^5X^6LX^7WX^8G$ (SEQ ID NO: 2256) (HC, FW2, consensus sequence 2) wherein X^1 is selected from V and I; X^2 is selected from R and K; X^3 is selected from A, M, N, R, K, T, and S; X^4 is selected from P, T, and H; X^5 is selected from Q, K, and R; X^6 is selected from G, R and S; X^7 is selected from E, D, K, Q, and A; and, X^8 is selected from M, I, and V. In some aspects, the nucleotide sequence encodes a sequence identical to WVRQAPGQGLEWMG (SEQ ID NO: 2228) except for at least one, two, three, four, or five substitutions selected from V2I, R3K, A5(MNRKTS), P6(TH), Q8(KR), G9(RS), E11(DKQA), and M13(IV). In some aspects, the nucleotide sequence encodes WVRQAPGQGLEWMG (SEQ ID NO: 2228), or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO:2228.

[0236] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $WX^1RX^2X^3X^4X^5X^6X^7LX^8WX^9X^{10}$ (SEQ ID NO: 2260) (HC, FW2, consensus sequence 3) wherein X^1 is selected from I and V; X^2 is selected from Q and H; X^3 is selected from L, P, S, and H; X^4 is selected from P and S; X^5 is selected from G and E; X^6 is selected from K and R; X^7 is selected from G and A; X^8 is selected from E and Q; X^9 is selected from I and L; and, X^{10} is selected from G and A. In some aspects, the nucleotide sequence encodes a sequence identical to WIRQLPGKGLEWIG (SEQ ID NO: 2232) except for at least one, two, three, four, or five substitution selected from I2V, Q4H, L5(PSH), P6S, G7E, K8R, G9A, E11Q, I13L, and G14A. In some aspects, the nucleotide sequence encodes WIRQLPGKGLEWIG (SEQ ID NO: 2232), or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO: 2232.

[0237] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $X^1X^2X^3X^4SX^5DX^6X^7X^8X^9X^{10}X^{11}X^{12}LX^{13}X^{14}X^{15}X^{16}LX^{17}X^{18}EDTX^{19}X^{20}X^{21}X^{22}C$ (SEQ ID NO: 2253) (HC, FW3, consensus sequence 1) wherein X^1 is selected from R and K; X^2 is selected from F and V; X^3 is selected from T, I, and A; X^4 is selected from L and I; X^5 is selected from V, R, L, and A; X^6 is selected from R, N, T, D, K, and S; X^7 is selected from S, A and V; X^8 is selected from K, R, and E; X^9 is selected from N, S, R, H, and T; X^{10} is selected from T and S; X^{11} is selected from L, A, and F; X^{12} is selected from Y and F; X^{13} is selected from Q and E; X^{14} is selected from M and V; X^{15} is selected from N, D, and S; X^{16} is selected from S, G, and I; X^{17} is selected from R and K; X^{18} is selected from A, S, D, V, and P; X^{19} is selected from A and G; X^{20} is selected from V, M, and L; X^{21} is selected from Y and F; and, X^{22} is selected from Y and F. In some aspects, the nucleotide sequence encodes a sequence identical to RFTLSVDRSKNTLYLQMNSLRAEDTAVYYC (SEQ ID NO: 2225) except for at least one, two, three, four, or five substitutions selected from R1K, F2V, T3(IA), L4I, V6(RLA), R8(NTDKS), S9(AV), K10(RE), N11(SRHT), T12S, L13(AF), Y14F, Q16E,

M17V, N18(DS), S19(GI), R21K, A22(SDVP), A26G, V27(ML), Y28F, and Y29F. In some aspects, the nucleotide sequence encodes RFTLSVDRSKNTLYLQMNSLRAEDTAVYYC (SEQ ID NO: 2225), or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO: 2225.

[0238] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $X^1X^2X^3X^4X^5X^6X^7X^8SX^9X^{10}TX^{11}X^{12}X^{13}X^{14}X^{15}X^{16}X^{17}LX^{18}X^{19}X^{20}DX^{21}X^{22}X^{23}YX^{24}C$ (SEQ ID NO: 2257) (HC, FW3, consensus sequence 2) wherein X^1 is selected from R, Q, and K; X^2 is selected from V, I, F, G, and A; X^3 is selected from T, A, and K; X^4 is selected from M, I, L, and F; X^5 is selected from T and S; X^6 is selected from T, A, R, V, S, E, and L; X^7 is selected from D, E, and N; X^8 is selected from T, K, Q, S, P, R, I, N, and E; X^9 is selected from T, K, S, A, I, and V; X^{10} is selected from S, N, D, and T; X^{11} is selected from A, V, and T; X^{12} is selected from Y, and F; X^{13} is selected from M and L; X^{14} is selected from E, Q, and D; X^{15} is selected from L, M, W, and I; X^{16} is selected from R, S, D, L, K, T, and N; X^{17} is selected from S and R; X^{18} is selected from R, K, Q, and T; X^{19} is selected from S, H, F, A, and P; X^{20} is selected from D, E, and S; X^{21} is selected from T and S; X^{22} is selected from A and G; X^{23} is selected from V, F, T, and M; and, X^{24} is selected from Y, F, and L. In some aspects, the nucleotide sequence encodes a sequence identical to RVTMTTDTSTSTAYMELRSLRSDDTAVYYC (SEQ ID NO: 2229) except for at least one, two, three, four, or five substitutions selected from R1(QK), V2(IFGA), T3(AK), M4(ILF), T5S, T6(ARVSEL), D7(EN), T8(KQSPRINE), T10(KSAIV), S11(NDT), A13(VT), Y14F, M15L, E16(QD), L17(MWI), R18(SDLKTN), S19R, R21(KQT), S22(HFAP), D23(ES), T25S, A26G, V27(FTM), and Y29(FL). In some aspects, the nucleotide sequence encodes RVTMTTDTSTSTAYMELRSLRSDDTAVYYC (SEQ ID NO: 2229), or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO:2229.

[0239] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15,

MAP16 or any combination thereof), wherein the nucleotide sequence encodes $RX^1X^2X^3X^4X^5DX^6SX^7X^8QX^9X^{10}LX^{11}X^{12}X^{13}X^{14}X^{15}X^{16}X^{17}X^{18}DTAX^{19}X^{20}X^{21}C$ (SEQ ID NO: 2261) (HC, FW3, consensus sequence 3) wherein X^1 is selected from V and L; X^2 is selected from T and S; X^3 is selected from I and M; X^4 is selected from S and L; X^5 is selected from V, R, and K; X^6 is selected from T and K; X^7 is selected from K and R; X^8 is selected from K and N; X^9 is selected from F and V; X^{10} is selected from S and V; X^{11} is selected from R, T, K, and M; X^{12} is selected from L, I, M, and V; X^{13} is selected from S, T, and N; X^{14} is selected from S and N; X^{15} is selected from V and M; X^{16} is selected from T and D; X^{17} is selected from A and P; X^{18} is selected from A and V; X^{19} is selected from V and T; X^{20} is selected from Y and W; and, X^{21} is selected from Y, F and W. In some aspects, the nucleotide sequence encodes a sequence identical to RVTISVDTSKKQFSLRLSSVTAADTAVYYC (SEQ ID NO: 2233). except for at least one, two, three, four or five substitutions selected from V2L, T3S, I4M, S5L, V6(RK), T8K, K10R, K11N, F13V, S14V, R16(TKM), L17(IMV), S18(TN), S19N, V20M, T21D, A22P, A23V, V27T, Y28W, and Y29(FW). In some aspects, the nucleotide sequence encodes RVTISVDTSKKQFSLRLSSVTAADTAVYYC (SEQ ID NO: 2233), or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO:2233.

[0240] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $WGX^1GX^2X^3VTVS$ (SEQ ID NO: 2254) (HC, FW4, consensus sequence 1) wherein X^1 is selected from Q, R, and K; X^2 is selected from T, I and A; and, X^3 is selected from L, S, T, M, and P. In some aspects, the nucleotide sequence encodes a sequence identical to WGQGTLVTVS (SEQ ID NO: 2226) except for at least one, two, or three substitutions selected from Q3(RK), T5(IA), and L6(STMP). In some aspects, the nucleotide sequence encodes WGQGTLVTVS (SEQ ID NO: 2226), or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO: 2226.

[0241] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15,

MAP16 or any combination thereof), wherein the nucleotide sequence encodes $WGX^1GTX^2X^3TVS$ (SEQ ID NO: 2258) (HC, FW4, consensus sequence 2) wherein X^1 is selected from R, Q, K, A and S; X^2 is selected from L, M, T, Q, and P; and, X^3 is selected from V and L. In some aspects, the nucleotide sequence encodes a sequence identical to WGRGTLVTVS (SEQ ID NO: 2230) except for at least one or two substitutions selected from R3(QKAS), L6(MTQP), and V7L. In some aspects, the nucleotide sequence encodes WGRGTLVTVS (SEQ ID NO: 2230), or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO:2230.

[0242] Also provided is a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $WX^1X^2GX^3X^4VTVS$ (SEQ ID NO: 2262) (HC, FW4, consensus sequence 3) wherein X^1 is selected from G and D; X^2 is selected from Q and R; X^3 is selected from T and S; and, X^4 is selected from T, L, and M. In some aspects, the nucleotide sequence encodes a sequence identical to WGQGTTVTVS (SEQ ID NO: 2234).except for at least one, two, three or four substitutions selected from G2D, Q3R, T5S, and T6(LM). In some aspects, the nucleotide sequence encodes WGQGTTVTVS (SEQ ID NO: 2234), or a sequence about 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identical to SEQ ID NO:2234.

[0243] In another aspect, the present disclosure provides consensus sequences defining linker sequences. The term "linker" as used herein refers to a polynucleotide encoding a peptide or polypeptide sequence wherein the main function of the expressed peptide or polypeptide is to connect to functional moieties (e.g. a VH domain and VL domain in an scFv). As used herein, the term linker refers interchangeably to the peptide or polypeptide encoded by such polynucleotide.

[0244] Thus, the disclosure provides a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes a sequence of formula $(Gly_xSer)_y$, wherein x and y are integers between 1 and 100. In some aspects, the sequence of formula $(Gly_xSer)_y$ functions as a linker. In some aspects, the

linker comprises the sequence (Gly_4Ser), (Gly_3Ser), (Gly_2Ser), or a combination thereof. In some aspects, the linker comprises the sequence $(Gly_4Ser)_3$. In some aspects, codon-optimized or non-codon-optimized hinge sequences can be used as linkers.

[0245] In some aspects, a polynucleotide disclosed herein, e.g., a polynucleotide comprising a codon-optimized nucleotide sequence encoding a VH domain and a codon-optimized nucleotide sequence encoding a VL domain can further comprise a nucleotide sequence encodes a sequence of formula $(Gly_xSer)_y$, wherein x and y are integers between 1 and 100, interposed between the nucleotide sequence encoding the VH and VL domain. In some aspects, such polynucleotide encodes an scFv.

[0246] In some aspects, two or more linkers (wherein each of the linkers can the same or different) can be linked in tandem. Generally, linkers provided flexibility to the protein product resulting from the expression of polynucleotide disclosed herein. The presence of linkers can maintain structural components in the expressed product (e.g., VH and VL domain in an scFv) at an optimal distance (e.g., so the VH and VL domain interact optimally with an epitope).

[0247] Linkers are not typically cleaved, thus, in some aspects, the linker is a non-cleavable linker. However, in certain aspects, such cleavable can be desirable. Accordingly, in some aspects, a linker can comprise one or more protease-cleavable sites, which can be located within the sequence of the linker or flanking the linker at either end of the linker sequence.

[0248] In some aspects, the linker comprises at least two, at least three, at least four, at least five, at least 10, at least 20, at least 30, at least 40, at least 50, at least 70, at least 80, at least 90, or at least 100 amino acids. The peptide linker can comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81,. 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100 amino acids. In some aspect, a hinge region of an antibody can function as a linker. In some aspects, the hinge region is codon-optimized.

[0249] In some aspects, the disclosure provides a polynucleotide encoding an antibody or an antigen binding portion thereof comprising

(i) a polynucleotide comprising a codon-optimized nucleotide sequence encoding the first framework region (FW1) of a lambda light chain or a kappa light chain variable domain,

(ii) a polynucleotide comprising a codon-optimized nucleotide sequence encoding the second framework region (FW2) of a lambda light chain or a kappa light chain variable domain,

(iii) a polynucleotide comprising a codon-optimized nucleotide sequence encoding the third framework region (FW3) of a lambda light chain or a kappa light chain variable domain,

(iv) a polynucleotide comprising a codon-optimized nucleotide sequence encoding the fourth framework region (FW4) of a lambda light chain or a kappa light chain variable domain, or

(v) any combination thereof.

[0250] In some aspects, the disclosure provides a polynucleotide encoding an antibody or an antigen binding portion thereof comprising

(i) a polynucleotide comprising a codon-optimized nucleotide sequence encoding the first framework region (FW1) of a lambda light chain or a kappa light chain variable domain,

(ii) a polynucleotide comprising a codon-optimized nucleotide sequence encoding the second framework region (FW2) of a lambda light chain or a kappa light chain variable domain,

(iii) a polynucleotide comprising a codon-optimized nucleotide sequence encoding the third framework region (FW3) of a lambda light chain or a kappa light chain variable domain, and

(iv) a polynucleotide comprising a codon-optimized nucleotide sequence encoding the fourth framework region (FW4) of a lambda light chain or a kappa light chain variable domain.

[0251] Also provided is a polynucleotide encoding an antibody or an antigen binding portion thereof comprising

(i) a codon-optimized nucleotide sequence encoding the first framework region (FW1) of a heavy chain variable domain,

(ii) a codon-optimized nucleotide sequence encoding the second framework region (FW2) of a heavy chain variable domain,

(iii) a codon-optimized nucleotide sequence encoding the third framework region (FW3) of a heavy chain variable domain,

(iv) a codon-optimized nucleotide sequence encoding the fourth framework region (FW4) of a heavy chain variable domain, or

(v) any combination thereof.

[0252] Also provided is a polynucleotide encoding an antibody or an antigen binding portion thereof comprising

(i) a codon-optimized nucleotide sequence encoding the first framework region (FW1) of a heavy chain variable domain,

(ii) a codon-optimized nucleotide sequence encoding the second framework region (FW2) of a heavy chain variable domain,

(iii) a codon-optimized nucleotide sequence encoding the third framework region (FW3) of a heavy chain variable domain, and

(iv) a codon-optimized nucleotide sequence encoding the fourth framework region (FW4) of a heavy chain variable domain.

[0253] In some aspects, a polynucleotide comprising codon-optimized nucleotides encoding the FW1-FW4 regions of a light chain also comprises codon-optimized nucleotides encoding the FW1-FW4 regions of a light chain.

[0254] In some aspects, a polypeptide comprising codon-optimized nucleotides encoding the FW1-FW4 regions of a light chain and/or codon-optimized nucleotides encoding the FW1-FW4 regions of a light chain further comprises codon-optimized nucleotides encoding a constant domain (e.g., CL, CH1, CH2, CH3, or a combination thereof).

[0255] The present disclosure also provides a polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the codon-optimized nucleotide sequence encodes a fragment of

(i) the amino acid sequences of any one of the therapeutic antibodies of SEQ ID NO: 1979-2006; or,

(ii) a polypeptide sequence encoded by any polynucleotide sequence disclosed herein, wherein the fragment is about 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 205, 210, 215, 220, 225, 230, 235, 240, 245,

250, 255, 260, 265, 270, 275, 280, 285, 290, 295, 300, 305, 310, 315, 320, 325, 330, 335, 340, 345, 350, 355, 360, 365, 370, 375, 380, 385, 390, 395, 400, 405, 410, 415, 420, 425, 430, 435, 440, 445, 450, 455, 460, 465, 470, 475, 480, 485, 490, 495, 500, 505, 510, 515, 520, 525, 530, 535, 540, 545, or 550 amino acids long.

[0256] The polynucleotides of the present disclosure can be in the form of RNA or in the form of DNA. DNA includes cDNA, and synthetic DNA; and can be double-stranded or single-stranded. In particular aspects, the polynucleotide is an mRNA. In some aspects, the mRNA is a synthetic mRNA. In certain aspects, the polynucleotides are isolated. In certain aspects, the polynucleotides are substantially pure. In some aspects, the polynucleotide comprises at least one nucleotide analogue. In some aspects, the at least one nucleotide analogue is selected from the group consisting of a 2'-O-methoxyethyl-RNA (2'-MOE-RNA) monomer, a 2'-fluoro-DNA monomer, a 2'-O-alkyl-RNA monomer, a 2'-amino-DNA monomer, a locked nucleic acid (LNA) monomer, a cEt monomer, a cMOE monomer, a 5'-Me-LNA monomer, a 2'-(3-hydroxy)propyl-RNA monomer, an arabino nucleic acid (ANA) monomer, a 2'-fluoro-ANA monomer, an anhydrohexitol nucleic acid (HNA) monomer, an intercalating nucleic acid (INA) monomer, and a combination of two or more of said nucleotide analogues. In some aspects, the polynucleotide comprises at least one backbone modification. In some aspects, the at least one backbone modification is a phosphorothioate internucleotide linkage. In some aspects, all of the internucleotide linkages are phosphorothioate internucleotide linkages.

[0257] In certain aspects the polynucleotides comprise the coding sequence for the mature polypeptide fused in the same reading frame to a polynucleotide which aids, for example, in expression and secretion of a polypeptide from a host cell (e.g., a leader sequence which functions as a secretory sequence for controlling transport of a polypeptide from the cell). The polypeptide having a leader sequence is a preprotein and can have the leader sequence cleaved by the host cell to form the mature form of the polypeptide. The polynucleotides can also encode for a proprotein which is the mature protein plus additional 5' amino acid residues.

III. Methods of Making Polynucleotides

[0258] The present disclosure also provides methods for making a polynucleotide disclosed herein or a complement thereof. In some aspect, a codon-optimized nucleotide

sequence (e.g., an mRNA) disclosed herein, and encoding a protein of interest, e.g., an antibody or a functional fragment thereof, or a polypeptide incorporating such codon-optimized nucleotide sequence can be produced using *in vitro* translation (IVT). In other aspects, a codon-optimized nucleotide sequence (e.g., an mRNA) disclosed herein, and encoding a protein of interest, e.g., an antibody or a functional fragment thereof, or a polypeptide incorporating such codon-optimized nucleotide sequence can be constructed by chemical synthesis using an oligonucleotide synthesizer. In other aspects, a codon-optimized nucleotide sequence (e.g., an mRNA) disclosed herein, and encoding a protein of interest, e.g., an antibody or a functional fragment thereof, or a polypeptide incorporating such codon-optimized nucleotide sequence is made by using a host cell. In certain aspects, a codon-optimized nucleotide sequence (e.g., an mRNA) disclosed herein, and encoding a protein of interest, e.g., an antibody or a functional fragment thereof, or a polypeptide incorporating such codon-optimized nucleotide sequence is made by one or more combination of the IVT, chemical synthesis, host cell expression, or any other methods known in the art.

[0259] Naturally occurring nucleosides, non-naturally occurring nucleosides, or combinations thereof, replacing totally or partially naturally occurring nucleosides present in the candidate nucleotide sequence can be incorporated into an codon-optimized nucleotide sequence (e.g., an mRNA) encoding a polypeptide of interest. The resultant mRNAs can then be examined for their ability to produce protein and/or produce a therapeutic outcome.

(a) In Vitro Transcription-enzymatic synthesis

[0260] A codon-optimized nucleotide sequence disclosed herein can be transcribed using an *in vitr*o transcription (IVT) system. The system typically comprises a transcription buffer, nucleotide triphosphates (NTPs), an RNase inhibitor and a polymerase. The NTPs can be selected from, but are not limited to, those described herein including natural and unnatural (modified) NTPs. The polymerase can be selected from, but is not limited to, T7 RNA polymerase, T3 RNA polymerase and mutant polymerases such as, but not limited to, polymerases able to incorporate modified nucleic acids. See U.S. Publ. No. US20130259923, which is herein incorporated by reference in its entirety.

[0261] The IVT system typically comprises a transcription buffer, nucleotide triphosphates (NTPs), an RNase inhibitor and a polymerase. The NTPs can be selected

from, but are not limited to, those described herein including natural and unnatural (modified) NTPs. The polymerase can be selected from, but is not limited to, T7 RNA polymerase, T3 RNA polymerase and mutant polymerases such as, but not limited to, polymerases able to incorporate polynucleotides disclosed herein.

[0262] Any number of RNA polymerases or variants can be used in the synthesis of the polynucleotides of the present invention.

[0263] RNA polymerases can be modified by inserting or deleting amino acids of the RNA polymerase sequence. As a non-limiting example, the RNA polymerase can be modified to exhibit an increased ability to incorporate a 2′-modified nucleotide triphosphate compared to an unmodified RNA polymerase (see International Publication WO2008078180 and U.S. Patent 8,101,385; herein incorporated by reference in their entireties).

[0264] Variants can be obtained by evolving an RNA polymerase, optimizing the RNA polymerase amino acid and/or nucleic acid sequence and/or by using other methods known in the art. As a non-limiting example, T7 RNA polymerase variants can be evolved using the continuous directed evolution system set out by Esvelt et al. (Nature (2011) 472(7344):499-503; herein incorporated by reference in its entirety) where clones of T7 RNA polymerase can encode at least one mutation such as, but not limited to, lysine at position 93 substituted for threonine (K93T), I4M, A7T, E63V, V64D, A65E, D66Y, T76N, C125R, S128R, A136T, N165S, G175R, H176L, Y178H, F182L, L196F, G198V, D208Y, E222K, S228A, Q239R, T243N, G259D, M267I, G280C, H300R, D351A, A354S, E356D, L360P, A383V, Y385C, D388Y, S397R, M401T, N410S, K450R, P451T, G452V, E484A, H523L, H524N, G542V, E565K, K577E, K577M, N601S, S684Y, L699I, K713E, N748D, Q754R, E775K, A827V, D851N or L864F. As another non-limiting example, T7 RNA polymerase variants can encode at least mutation as described in U.S. Pub. Nos. 20100120024 and 20070117112; herein incorporated by reference in their entireties. Variants of RNA polymerase can also include, but are not limited to, substitutional variants, conservative amino acid substitution, insertional variants, deletional variants and/or covalent derivatives.

[0265] In one aspect, the polynucleotide can be designed to be recognized by the wild type or variant RNA polymerases. In doing so, the polynucleotide can be modified to contain sites or regions of sequence changes from the wild type or parent chimeric polynucleotide.

[0266] Polynucleotide or nucleic acid synthesis reactions can be carried out by enzymatic methods utilizing polymerases. Polymerases catalyze the creation of phosphodiester bonds between nucleotides in a polynucleotide or nucleic acid chain. Currently known DNA polymerases can be divided into different families based on amino acid sequence comparison and crystal structure analysis. DNA polymerase I (pol I) or A polymerase family, including the Klenow fragments of E. Coli, Bacillus DNA polymerase I, Thermus aquaticus (Taq) DNA polymerases, and the T7 RNA and DNA polymerases, is among the best studied of these families. Another large family is DNA polymerase α (pol α) or B polymerase family, including all eukaryotic replicating DNA polymerases and polymerases from phages T4 and RB69. Although they employ similar catalytic mechanism, these families of polymerases differ in substrate specificity, substrate analog-incorporating efficiency, degree and rate for primer extension, mode of DNA synthesis, exonuclease activity, and sensitivity against inhibitors.

[0267] DNA polymerases are also selected based on the optimum reaction conditions they require, such as reaction temperature, pH, and template and primer concentrations. Sometimes a combination of more than one DNA polymerases is employed to achieve the desired DNA fragment size and synthesis efficiency. For example, Cheng et al. increase pH, add glycerol and dimethyl sulfoxide, decrease denaturation times, increase extension times, and utilize a secondary thermostable DNA polymerase that possesses a 3′ to 5′ exonuclease activity to effectively amplify long targets from cloned inserts and human genomic DNA. (Cheng et al., PNAS, Vol. 91, 5695-5699 (1994), the contents of which are incorporated herein by reference in their entirety). RNA polymerases from bacteriophage T3, T7, and SP6 have been widely used to prepare RNAs for biochemical and biophysical studies. RNA polymerases, capping enzymes, and poly-A polymerases are disclosed in the co-pending International Publication No. WO2014028429, the contents of which are incorporated herein by reference in their entirety.

[0268] In one aspect, the RNA polymerase which can be used in the synthesis of the polynucleotides described herein is a Syn5 RNA polymerase. (see Zhu et al. Nucleic Acids Research 2013, the contents of which is herein incorporated by reference in its entirety). The Syn5 RNA polymerase was recently characterized from marine cyanophage Syn5 by Zhu et al. where they also identified the promoter sequence (see Zhu et al. Nucleic Acids Research 2013, the contents of which is herein incorporated by reference in its entirety). Zhu et al. found that Syn5 RNA polymerase catalyzed RNA

synthesis over a wider range of temperatures and salinity as compared to T7 RNA polymerase. Additionally, the requirement for the initiating nucleotide at the promoter was found to be less stringent for Syn5 RNA polymerase as compared to the T7 RNA polymerase making Syn5 RNA polymerase promising for RNA synthesis.

[0269] In one aspect, a Syn5 RNA polymerase can be used in the synthesis of the polynucleotides described herein. As a non-limiting example, a Syn5 RNA polymerase can be used in the synthesis of the polynucleotide requiring a precise 3′-termini.

[0270] In one aspect, a Syn5 promoter can be used in the synthesis of the polynucleotides. As a non-limiting example, the Syn5 promoter can be 5′-ATTGGGCACCCGTAAGGG-3′ as described by Zhu et al. (Nucleic Acids Research 2013, the contents of which is herein incorporated by reference in its entirety).

[0271] In one aspect, a Syn5 RNA polymerase can be used in the synthesis of polynucleotides comprising at least one chemical modification described herein and/or known in the art. (see e.g., the incorporation of pseudo-UTP and 5Me-CTP described in Zhu et al. Nucleic Acids Research 2013, the contents of which is herein incorporated by reference in its entirety).

[0272] In one aspect, the polynucleotides described herein can be synthesized using a Syn5 RNA polymerase which has been purified using modified and improved purification procedure described by Zhu et al. (Nucleic Acids Research 2013, the contents of which is herein incorporated by reference in its entirety).

[0273] Various tools in genetic engineering are based on the enzymatic amplification of a target gene which acts as a template. For the study of sequences of individual genes or specific regions of interest and other research needs, it is necessary to generate multiple copies of a target gene from a small sample of polynucleotides or nucleic acids. Such methods can be applied in the manufacture of the polynucleotides of the invention.

[0274] Polymerase chain reaction (PCR) has wide applications in rapid amplification of a target gene, as well as genome mapping and sequencing. The key components for synthesizing DNA comprise target DNA molecules as a template, primers complementary to the ends of target DNA strands, deoxynucleoside triphosphates (dNTPs) as building blocks, and a DNA polymerase. As PCR progresses through denaturation, annealing and extension steps, the newly produced DNA molecules can act as a template for the next circle of replication, achieving exponentially amplification of the target DNA. PCR requires a cycle of heating and cooling for denaturation and annealing. Variations of the

basic PCR include asymmetric PCR [Innis et al., PNAS, vol. 85, 9436-9440 (1988)], inverse PCR [Ochman et al., Genetics, vol. 120(3), 621-623, (1988)], reverse transcription PCR (RT-PCR) (Freeman et al., BioTechniques, vol. 26(1), 112-22, 124-5 (1999), the contents of which are incorporated herein by reference in their entirety and so on). In RT-PCR, a single stranded RNA is the desired target and is converted to a double stranded DNA first by reverse transcriptase.

[0275] A variety of isothermal in vitro nucleic acid amplification techniques have been developed as alternatives or complements of PCR. For example, strand displacement amplification (SDA) is based on the ability of a restriction enzyme to form a nick. (Walker et al., PNAS, vol. 89, 392-396 (1992), the contents of which are incorporated herein by reference in their entirety)). A restriction enzyme recognition sequence is inserted into an annealed primer sequence. Primers are extended by a DNA polymerase and dNTPs to form a duplex. Only one strand of the duplex is cleaved by the restriction enzyme. Each single strand chain is then available as a template for subsequent synthesis. SDA does not require the complicated temperature control cycle of PCR.

[0276] Nucleic acid sequence-based amplification (NASBA), also called transcription mediated amplification (TMA), is also an isothermal amplification method that utilizes a combination of DNA polymerase, reverse transcriptase, RNAse H, and T7 RNA polymerase. [Compton, Nature, vol. 350, 91-92 (1991)] the contents of which are incorporated herein by reference in their entirety. A target RNA is used as a template and a reverse transcriptase synthesizes its complementary DNA strand. RNAse H hydrolyzes the RNA template, making space for a DNA polymerase to synthesize a DNA strand complementary to the first DNA strand which is complementary to the RNA target, forming a DNA duplex. T7 RNA polymerase continuously generates complementary RNA strands of this DNA duplex. These RNA strands act as templates for new cycles of DNA synthesis, resulting in amplification of the target gene.

[0277] Rolling-circle amplification (RCA) amplifies a single stranded circular polynucleotide and involves numerous rounds of isothermal enzymatic synthesis where Φ29 DNA polymerase extends a primer by continuously progressing around the polynucleotide circle to replicate its sequence over and over again. Therefore, a linear copy of the circular template is achieved. A primer can then be annealed to this linear copy and its complementary chain can be synthesized. [See Lizardi et al., Nature Genetics, vol. 19, 225-232 (1998)] the contents of which are incorporated herein by

reference in their entirety. A single stranded circular DNA can also serve as a template for RNA synthesis in the presence of an RNA polymerase. (Daubendiek et al., JACS, vol. 117, 7818-7819 (1995), the contents of which are incorporated herein by reference in their entirety). An inverse rapid amplification of cDNA ends (RACE) RCA is described by Polidoros et al. A messenger RNA (mRNA) is reverse transcribed into cDNA, followed by RNAse H treatment to separate the cDNA. The cDNA is then circularized by CircLigase into a circular DNA. The amplification of the resulting circular DNA is achieved with RCA. (Polidoros et al., BioTechniques, vol. 41, 35-42 (2006), the contents of which are incorporated herein by reference in their entirety).

[0278] Any of the foregoing methods can be utilized in the manufacture of one or more regions of the polynucleotides of the present invention.

[0279] Assembling polynucleotides or nucleic acids by a ligase is also widely used. DNA or RNA ligases promote intermolecular ligation of the 5′ and 3′ ends of polynucleotide chains through the formation of a phosphodiester bond. Ligase chain reaction (LCR) is a promising diagnosing technique based on the principle that two adjacent polynucleotide probes hybridize to one strand of a target gene and couple to each other by a ligase. If a target gene is not present, or if there is a mismatch at the target gene, such as a single-nucleotide polymorphism (SNP), the probes cannot ligase. (Wiedmann et al., PCR Methods and Application, vol.3 (4), s51-s64 (1994), the contents of which are incorporated herein by reference in their entirety). LCR can be combined with various amplification techniques to increase sensitivity of detection or to increase the amount of products if it is used in synthesizing polynucleotides and nucleic acids.

[0280] Several library preparation kits for nucleic acids are now commercially available. They include enzymes and buffers to convert a small amount of nucleic acid samples into an indexed library for downstream applications. For example, DNA fragments can be placed in a NEBNEXT® ULTRATM DNA Library Prep Kit by NEWENGLAND BIOLABS® for end preparation, ligation, size selection, clean-up, PCR amplification and final clean-up.

[0281] Continued development is going on to improvement the amplification techniques. For example, US Pat. 8,367,328 to Asada et al. the contents of which are incorporated herein by reference in their entirety, teaches utilizing a reaction enhancer to increase the efficiency of DNA synthesis reactions by DNA polymerases. The reaction enhancer comprises an acidic substance or cationic complexes of an acidic substance. US Pat.

7.384,739 to Kitabayashi et al. the contents of which are incorporated herein by reference in their entirety, teaches a carboxylate ion-supplying substance that promotes enzymatic DNA synthesis, wherein the carboxylate ion-supplying substance is selected from oxalic acid, malonic acid, esters of oxalic acid, esters of malonic acid, salts of malonic acid, and esters of maleic acid. US Pat. 7,378,262 to Sobek et al. the contents of which are incorporated herein by reference in their entirety, discloses an enzyme composition to increase fidelity of DNA amplifications. The composition comprises one enzyme with 3′ exonuclease activity but no polymerase activity and another enzyme that is a polymerase. Both of the enzymes are thermostable and are reversibly modified to be inactive at lower temperatures.

[0282] US Pat. No. 7,550,264 to Getts et al. teaches multiple round of synthesis of sense RNA molecules are performed by attaching oligodeoxynucleotides tails onto the 3′ end of cDNA molecules and initiating RNA transcription using RNA polymerase, the contents of which are incorporated herein by reference in their entirety. US Pat. Publication No. 2013/0183718 to Rohayem teaches RNA synthesis by RNA-dependent RNA polymerases (RdRp) displaying an RNA polymerase activity on single-stranded DNA templates, the contents of which are incorporated herein by reference in their entirety. Oligonucleotides with non-standard nucleotides can be synthesized with enzymatic polymerization by contacting a template comprising non-standard nucleotides with a mixture of nucleotides that are complementary to the nucleotides of the template as disclosed in US Pat. No. 6,617,106 to Benner, the contents of which are incorporated herein by reference in their entirety.

(b) Chemical synthesis

[0283] Standard methods can be applied to synthesize an isolated polynucleotide sequence encoding an isolated polypeptide of interest. For example, a single DNA or RNA oligomer containing a codon-optimized nucleotide sequence coding for the particular isolated polypeptide can be synthesized. In other aspects, several small oligonucleotides coding for portions of the desired polypeptide can be synthesized and then ligated. In some aspects, the individual oligonucleotides typically contain 5' or 3' overhangs for complementary assembly.

[0284] A polynucleotide disclosed herein (e.g., mRNA) can be chemically synthesized using chemical synthesis methods and potential nucleobase substitutions known in the art.

See, for example, International Publication Nos. WO2014093924, WO2013052523; WO2013039857, WO2012135805, WO2013151671; U.S. Publ. No. US20130115272; or U.S. Pat. Nos. US8999380, US8710200, all of which are herein incorporated by reference in their entireties.

(c) Nucleoside substitutions

[0285] Examples of naturally occurring nucleosides that can be incorporated using IVT or chemical synthesis to generate a codon-optimized nucleotide sequence disclosed herein (e.g., an mRNA) include 2'-O-methylcytidine, 4-thiouridine, 2'-O-methyluridine, 5-methyl-2-thiouridine, 5,2'-O-dimethyluridine, 5-aminomethyl-2-thiouridine, 5,2'-O-dimethylcytidine, 2-methylthio-N6-isopentenyladenosine, 2'-O-methyladenosine, 2'-O-methylguanosine, N6-methyl-N6-threonylcarbamoyladenosine, N6-hydroxynorvalylcarbamoyladenosine, 2-methylthio-N6-hydroxynorvalyl carbamoyl adenosine, 2'-O-ribosyladenosine (phosphate), N6,2'-O-dimethyladenosine, N6,N6,2'-O-trimethyladenosine, 1,2'-O-dimethyladenosine, N6-acetyladenosine, 2-methyladenosine, 2-methylthio-N6-methyladenosine, N2,2'-O-dimethylguanosine, N2,N2,2'-O-trimethylguanosine, 7-cyano-7-deazaguanosine, 7-aminomethyl-7-deazaguanosine, 2'-O-ribosylguanosine (phosphate), N2,7-dimethylguanosine, N2,N2,7-trimethylguanosine, 1,2'-O-dimethylguanosine, peroxywybutosine, hydroxywybutosine, undermodified hydroxywybutosine, methylwyosine, N2,7,2'-O-trimethylguanosine, 1,2'-O-dimethylinosine, 2'-O-methylinosine, 4-demethylwyosine, isowyosine, queuosine, epoxyqueuosine, galactosyl-queuosine, mannosyl-queuosine, archaeosine, and combinations thereof.

[0286] Examples of non-naturally occurring nucleosides that can be incorporated using IVT or chemical synthesis into a codon-optimized nucleotide sequence disclosed herein (e.g., an mRNA) include 5-(1-propynyl)ara-uridine, 2'-O-methyl-5-(1-propynyl)uridine, 2'-O-methyl-5-(1-propynyl)cytidine, 5-(1-propynyl)ara-cytidine, 5-ethynylara-cytidine, 5-ethynylcytidine, 5-vinylarauridine, (Z)-5-(2-bromo-vinyl)ara-uridine, (E)-5-(2-bromo-vinyl)ara-uridine, (Z)-5-(2-bromo-vinyl)uridine, (E)-5-(2-bromo-vinyl)uridine, 5-methoxyuridine, 5-methoxycytidine, 5-formyluridine, 5-cyanouridine, 5-dimethylaminouridine, 5-trideuteromethyl-6-deuterouridine, 5-cyanocytidine, 5-(2-chloro-phenyl)-2-thiocytidine, 5-(4-amino-phenyl)-2-thiocytidine, 5-(2-furanyl)uridine, 5-phenylethynyluridine, N4,2'-O-dimethylcytidine, 3'-ethynylcytidine, 4'-carbocyclic

adenosine, 4'-carbocyclic cytidine, 4'-carbocyclic guanosine, 4'-carbocyclic uridine, 4'-ethynyladenosine, 4'-ethynyluridine, 4'-ethynylcytidine, 4'-ethynylguanosine, 4'-azidouridine, 4'-azidocytidine, 4'-azidoadenosine, 4'-azidoguanosine, 2'-deoxy-2',2'-difluorocytidine, 2'-deoxy-2',2'-difluorouridine, 2'-deoxy-2',2'-difluoroadenosine, 2'-deoxy-2',2'-difluoroguanosine, 2'-deoxy-2'-b-fluorocytidine, 2'-deoxy-2'-b-fluorouridine, 2'-deoxy-2'-b-fluoroadenosine, 2'-deoxy-2'-b-fluoroguanosine, 8-trifluoromethyladenosine, 2'-deoxy-2'-b-chlorouridine, 2'-deoxy-2'-b-bromouridine, 2'-deoxy-2'-b-iodouridine, 2'-deoxy-2'-b-chlorocytidine, 2'-deoxy-2'-b-bromocytidine, 2'-deoxy-2'-b-iodocytidine, 2'-deoxy-2'-b-chloroadenosine, 2'-deoxy-2'-b-bromoadenosine, 2'-deoxy-2'-b-iodoadenosine, 2'-deoxy-2'-b-chloroguanosine, 2'-deoxy-2'-b-bromoguanosine, 2'-deoxy-2'-b-iodoguanosine, 5'-homo-cytidine, 5'-homo-adenosine, 5'-homo-uridine, 5'-homo-guanosine, 2'-deoxy-2'-a-mercaptouridine, 2'-deoxy-2'-a-thiomethoxyuridine, 2'-deoxy-2'-a-azidouridine, 2'-deoxy-2'-a-aminouridine, 2'-deoxy-2'-a-mercaptocytidine, 2'-deoxy-2'-a-thiomethoxycytidine, 2'-deoxy-2'-a-azidocytidine, 2'-deoxy-2'-a-aminocytidine, 2'-deoxy-2'-a-mercaptoadenosine, 2'-deoxy-2'-a-thiomethoxyadenosine, 2'-deoxy-2'-a-azidoadenosine, 2'-deoxy-2'-a-aminoadenosine, 2'-deoxy-2'-a-mercaptoguanosine, 2'-deoxy-2'-a-thiomethoxyguanosine, 2'-deoxy-2'-a-azidoguanosine, 2'-deoxy-2'-a-aminoguanosine, 2'-deoxy-2'-b-mercaptouridine, 2'-deoxy-2'-b-thiomethoxyuridine, 2'-deoxy-2'-b-azidouridine, 2'-deoxy-2'-b-aminouridine, 2'-deoxy-2'-b-mercaptocytidine, 2'-deoxy-2'-b-thiomethoxycytidine, 2'-deoxy-2'-b-azidocytidine, 2'-deoxy-2'-b-aminocytidine, 2'-deoxy-2'-b-mercaptoadenosine, 2'-deoxy-2'-b-thiomethoxyadenosine, 2'-deoxy-2'-b-azidoadenosine, 2'-deoxy-2'-b-aminoadenosine, 2'-deoxy-2'-b-mercaptoguanosine, 2'-deoxy-2'-b-thiomethoxyguanosine, 2'-deoxy-2'-b-azidoguanosine, 2'-deoxy-2'-b-aminoguanosine, 2'-b-trifluoromethyladenosine, 2'-b-trifluoromethylcytidine, 2'-b-trifluoromethylguanosine, 2'-b-trifluoromethyluridine, 2'-a-trifluoromethyladenosine, 2'-a-trifluoromethylcytidine, 2'-a-trifluoromethylguanosine, 2'-a-trifluoromethyluridine, 2'-b-ethynyladenosine, 2'-b-ethynylcytidine, 2'-b-ethynylguanosine, 2'-b-ethynyluridine, 2'-a-ethynyladenosine, 2'-a-ethynylcytidine, 2'-a-ethynylguanosine, 2'-a-ethynyluridine, (E)-5-(2-bromo-vinyl)cytidine, 2-trifluoromethyladenosine, 2-mercaptoadenosine, 2-aminoadenosine, 2-azidoadenosine, 2-fluoroadenosine, 2-chloroadenosine, 2-bromoadenosine, 2-iodoadenosine, formycin A, formycin B, oxoformycin, pyrrolosine, 9-deazaadenosine, 9-deazaguanosine, 3-

deazaadenosine, 3-deaza-3-fluoroadenosine, 3-deaza-3-chloroadenosine, 3-deaza-3-bromoadenosine, 3-deaza-3-iodoadenosine, 1-deazaadenosine, or combinations thereof.

[0287] In some aspects, a codon-optimized nucleotide sequence disclosed herein comprises at least one nucleotide analogue. In some aspects, at least one nucleotide analogue introduced by using IVT or chemical synthesis is selected from the group consisting of a 2'-O-methoxyethyl-RNA (2'-MOE-RNA) monomer, a 2'-fluoro-DNA monomer, a 2'-O-alkyl-RNA monomer, a 2'-amino-DNA monomer, a locked nucleic acid (LNA) monomer, a cEt monomer, a cMOE monomer, a 5'-Me-LNA monomer, a 2'-(3-hydroxy)propyl-RNA monomer, an arabino nucleic acid (ANA) monomer, a 2'-fluoro-ANA monomer, an anhydrohexitol nucleic acid (HNA) monomer, an intercalating nucleic acid (INA) monomer, and a combination of two or more of said nucleotide analogues. In some aspects, the optimized nucleic acid molecule comprises at least one backbone modification, for example, a phosphorothioate internucleotide linkage.

[0288] In some aspects, a codon-optimized nucleotide sequence disclosed herein comprises at least one nucleoside analogue introduced by using IVT or chemical synthesis selected from the group consisting of 2-pseudouridine, 5-methoxyuridine, 2-thiouridine, 4-thiouridine, N1-methyl-pseudouridine, 5-aza-uridine, 2-thio-5-aza-uridine, 4-thio-pseudouridine, 2-thio-pseudouridine, 5-hydroxyuridine, 4-methoxy-pseudouridine, 4-methoxy-2-thio-pseudouridine, 3-methyluridine, 5-carboxymethyl-uridine, 1-carboxymethyl-pseudouridine, 5-propynyl-uridine, 1-propynyl-pseudouridine, 2-methoxy-4-thio-uridine, 5-taurinomethyluridine, 1-taurinomethyl-pseudouridine, 5-taurinomethyl-2-thio-uridine, 1-taurinomethyl-4-thio-uridine, 5-methyl-uridine, 2-methoxyuridine, 1-methyl-pseudouridine, 4-thio-1-methyl-pseudouridine, 2-thio-1-methyl-pseudouridine, 1-methyl-1-deaza-pseudouridine, 2-thio-1-methyl-1-deaza-pseudouridine, 2-thio-dihydrouridine, 1-ethyl-pseudouridine.

[0289] In some aspects, a codon-optimized nucleotide sequence disclosed herein comprises at least one nucleoside analogue introduced by using IVT or chemical synthesis selected from the group consisting of 2-aminopurine, 2,6-diaminopurine, 7-deaza-adenine, 7-deaza-8-aza-adenine, 7-deaza-2-aminopurine, 7-deaza-8-aza-2-aminopurine, 7-deaza-2,6-diaminopurine, 7-deaza-8-aza-2,6-diaminopurine, 1-methyladenosine, N6-methyladenosine, N6-isopentenyladenosine, N6-(cis-hydroxyisopentenyl)adenosine, 2-methylthio-N6-(cis-hydroxyisopentenyl)adenosine, N6-

glycinylcarbamoyladenosine, N6-threonylcarbamoyladenosine, 2-methylthio-N6-threonyl carbamoyladenosine, N6,N6-dimethyladenosine, or 7-methyladenine.

[0290] In some aspects, a codon-optimized nucleotide sequence disclosed herein comprises at least one nucleoside analogue introduced by using IVT or chemical synthesis selected from the group consisting of inosine, 1-methyl-inosine, wyosine, wybutosine, 7-deaza-guanosine, 7-deaza-8-aza-guanosine, 6-thio-guanosine, 6-thio-7-deaza-guanosine, 6-thio-7-deaza-8-aza-guanosine, 7-methyl-guanosine, 6-thio-7-methyl-guanosine, 7-methylinosine, 6-methoxy-guanosine, 1-methylguanosine, N2-methylguanosine, N2,N2-dimethylguanosine, 8-oxo-guanosine, 7-methyl-8-oxo-guanosine, and 1-methyl-6-thio-guanosine.

[0291] In some aspects, a codon-optimized nucleotide sequence disclosed herein comprises at least one nucleoside analogue introduced by using IVT or chemical synthesis selected from the group consisting of 5-methylcytidine, 5-aza-cytidine, pseudoisocytidine, 3-methyl-cytidine, N4-acetylcytidine, 5-formylcytidine, N4-methylcytidine, 5-hydroxymethylcytidine, 1-methyl-pseudoisocytidine, pyrrolo-cytidine, pyrrolo-pseudoisocytidine, 2-thio-cytidine, 2-thio-5-methyl-cytidine, 4-thio-pseudoisocytidine, 4-thio-1-methyl-pseudoisocytidine, 4-thio-1-methyl-1-deaza-pseudoisocytidine, 1-methyl-1-deaza-pseudoisocytidine, zebularine, 5-aza-zebularine, 5-methyl-zebularine, 5-aza-2-thio-zebularine, 2-thio-zebularine, 2-methoxy-cytidine, 2-methoxy-5-methyl-cytidine, and 4-methoxy-pseudoisocytidine.

[0292] In some aspects, at least 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95%, or 100% of the uridine nucleosides in an nucleotide sequence disclosed herein (e.g., a candidate nucleotide sequence or a codon-optimized nucleotide sequence) have been replaced with a nucleoside selected from the group consisting of pseudouridine, 5-methoxyuridine, 2-thiouridine, 4-thiouridine, N1-methyl-pseudouridine, 5-aza-uridine, 2-thio-5-aza-uridine, 4-thio-pseudouridine, 2-thio-pseudouridine, 5-hydroxyuridine, 4-methoxy-pseudouridine, 4-methoxy-2-thio-pseudouridine, 3-methyluridine, 5-carboxymethyl-uridine, 1-carboxymethyl-pseudouridine, 5-propynyl-uridine, 1-propynyl-pseudouridine, 2-methoxy-4-thio-uridine, 5-taurinomethyluridine, 1-taurinomethyl-pseudouridine, 5-taurinomethyl-2-thio-uridine, 1-taurinomethyl-4-thio-uridine, 5-methyl-uridine, 2-methoxyuridine, 1-methyl-pseudouridine, 4-thio-1-methyl-pseudouridine, 2-thio-1-

methyl-pseudouridine, 1-methyl-1-deaza-pseudouridine, 2-thio-1-methyl-1-deaza-pseudouridine, 2-thio-dihydrouridine, or 1-ethyl-pseudouridine.

[0293] In some aspects, at least 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95%, or 100% of the adenosine nucleosides in a nucleotide sequence disclosed herein (e.g., a candidate nucleotide sequence or a codon-optimized nucleotide sequence) have been replaced with a nucleoside selected from the group consisting of 2-aminopurine, 2,6-diaminopurine, 7-deaza-adenine, 7-deaza-8-aza-adenine, 7-deaza-2-aminopurine, 7-deaza-8-aza-2-aminopurine, 7-deaza-2,6-diaminopurine, 7-deaza-8-aza-2,6-diaminopurine, 1-methyladenosine, N6-methyladenosine, N6-isopentenyladenosine, N6-(cis-hydroxyisopentenyl)adenosine, 2-methylthio-N6-(cis-hydroxyisopentenyl)adenosine, N6-glycinylcarbamoyladenosine, N6-threonylcarbamoyladenosine, 2-methylthio-N6-threonyl carbamoyladenosine, N6,N6-dimethyladenosine, or 7-methyladenine.

[0294] In some aspects, at least 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95%, or 100% of the guanosine nucleosides in a nucleotide sequence disclosed herein (e.g., a candidate nucleotide sequence or a codon-optimized nucleotide sequence) have been replaced with a nucleoside selected from the group consisting of inosine, 1-methyl-inosine, wyosine, wybutosine, 7-deaza-guanosine, 7-deaza-8-aza-guanosine, 6-thio-guanosine, 6-thio-7-deaza-guanosine, 6-thio-7-deaza-8-aza-guanosine, 7-methyl-guanosine, 6-thio-7-methyl-guanosine, 7-methylinosine, 6-methoxy-guanosine, 1-methylguanosine, N2-methylguanosine, N2,N2-dimethylguanosine, 8-oxo-guanosine, 7-methyl-8-oxo-guanosine, or 1-methyl-6-thio-guanosine.

[0295] In some aspects, at least 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95%, or 100% of the uridine nucleosides in a nucleotide sequence disclosed herein (e.g., a candidate nucleotide sequence or a codon-optimized nucleotide sequence) have been replaced with a nucleoside selected from the group consisting of 5-methylcytidine, 5-aza-cytidine, pseudoisocytidine, 3-methyl-cytidine, N4-acetylcytidine, 5-formylcytidine, N4-methylcytidine, 5-hydroxymethylcytidine, 1-methyl-pseudoisocytidine, pyrrolo-cytidine, pyrrolo-pseudoisocytidine, 2-thio-cytidine, 2-thio-5-methyl-cytidine, 4-thio-pseudoisocytidine, 4-thio-1-methyl-pseudoisocytidine, 4-thio-1-methyl-1-deaza-pseudoisocytidine, 1-methyl-1-deaza-pseudoisocytidine, zebularine, 5-aza-zebularine, 5-methyl-zebularine, 5-aza-2-

thio-zebularine, 2-thio-zebularine, 2-methoxy-cytidine, 2-methoxy-5-methyl-cytidine, 4-methoxy-pseudoisocytidine, or 4-methoxy-1-methyl-pseudoisocytidine.

[0296] In some aspects, a polynucleotide disclosed herein comprises a codon-optimized nucleotide sequence produced by IVT or chemical synthesis wherein

(i) at least one uridine in a candidate nucleotide sequence has been replaced with 2- pseudouridine, 5-methoxyuridine, 2-thiouridine, 4-thiouridine, N1-methyl-pseudouridine, 5-aza-uridine, 2-thio-5-aza-uridine, 4-thio-pseudouridine, 2-thio-pseudouridine, 5-hydroxyuridine, 4-methoxy-pseudouridine, 4-methoxy-2-thio-pseudouridine, 3-methyluridine, 5-carboxymethyl-uridine, 1-carboxymethyl-pseudouridine, 5-propynyl-uridine, 1-propynyl-pseudouridine, 2-methoxy-4-thio-uridine, 5-taurinomethyluridine, 1-taurinomethyl-pseudouridine, 5-taurinomethyl-2-thio-uridine, 1-taurinomethyl-4-thio-uridine, 5-methyl-uridine, 2-methoxyuridine, 1-methyl-pseudouridine, 4-thio-1-methyl-pseudouridine, 2-thio-1-methyl-pseudouridine, 1-methyl-1-deaza-pseudouridine, 2-thio-1-methyl-1-deaza-pseudouridine, 1-ethyl-pseudouridine, or 2-thio-dihydrouridine; and/or,

(ii) at least one adenosine in a candidate nucleotide sequence has been replaced with 2-aminopurine, 2,6-diaminopurine, 7-deaza-adenine, 7-deaza-8-aza-adenine, 7-deaza-2-aminopurine, 7-deaza-8-aza-2-aminopurine, 7-deaza-2,6-diaminopurine, 7-deaza-8-aza-2,6-diaminopurine, 1-methyladenosine, N6-methyladenosine, N6-isopentenyladenosine, N6-(cis-hydroxyisopentenyl)adenosine, 2-methylthio-N6-(cis-hydroxyisopentenyl) adenosine, N6-glycinylcarbamoyladenosine, N6-threonylcarbamoyladenosine, 2-methylthio-N6-threonyl carbamoyladenosine, N6,N6-dimethyladenosine, or 7-methyladenine; and/or,

(iii) at least one guanosine in a candidate nucleotide sequence has been replaced with inosine, 1-methyl-inosine, wyosine, wybutosine, 7-deaza-guanosine, 7-deaza-8-aza-guanosine, 6-thio-guanosine, 6-thio-7-deaza-guanosine, 6-thio-7-deaza-8-aza-guanosine, 7-methyl-guanosine, 6-thio-7-methyl-guanosine, 7-methylinosine, 6-methoxy-guanosine, 1-methylguanosine, N2-methylguanosine, N2,N2-dimethylguanosine, 8-oxo-guanosine, 7-methyl-8-oxo-guanosine, or 1-methyl-6-thio-guanosine; and/or,

(iv) at least one cytidine in a candidate nucleotide sequence has been replaced with 5-methylcytidine, 5-aza-cytidine, pseudoisocytidine, 3-methyl-cytidine, N4-acetylcytidine, 5-formylcytidine, N4-methylcytidine, 5-hydroxymethylcytidine, 1-methyl-pseudoisocytidine, pyrrolo-cytidine, pyrrolo-pseudoisocytidine, 2-thio-cytidine, 2-thio-5-

methyl-cytidine, 4-thio-pseudoisocytidine, 4-thio-1-methyl-pseudoisocytidine, 4-thio-1-methyl-1-deaza-pseudoisocytidine, 1-methyl-1-deaza-pseudoisocytidine, zebularine, 5-aza-zebularine, 5-methyl-zebularine, 5-aza-2-thio-zebularine, 2-thio-zebularine, 2-methoxy-cytidine, 2-methoxy-5-methyl-cytidine, or 4-methoxy-pseudoisocytidine.

[0297] In some aspects, a polynucleotide disclosed herein has been codon-optimized optimized, for example, by replacing by IVT or chemical synthesis in a candidate nucleotide sequence:

(i) 25% of uridines with 4-thiouridine; 50% of uridines with 4-thiouridine; or,

(ii) 100% of uridines with 4-thiouridine;

(iii) 25% of uridines with 2-thiouridine (s2U) and 25% of cytidines with 5-methylcytidine (m5C);

(iv) 50% of uridines with 2-thiouridine (s2U);

(v) 100% of uridines with pseudouridine (Ψ);

(vi) 100% of uridines with pseudouridine (Ψ) and 100% of cytidines with 5-methylcytidine (5mC);

(vii) 25% of uridines with 5-methoxyuridine (5moU) and 50% of cytidines with 5-methylcytidine (5mC);

(viii) 25% of uridines with 5-methoxyuridine (5moU) and 100% of cytidines with 5-methylcytidine (5mC);

(ix) 100% of uridines with 5-methoxyuridine (5moU);

(x) 100% of uridines with 5-methoxyuridine (5moU) and 100% of cytidines with 5-methylcytidine (5mC);

(xi) 100% of uridines with N1-methylpseudouridine (1mΨ);

(xii) 100% of uridines with 1-ethyl-pseudouridine; or,

(xiii) 100% of uridines with N1-methylpseudouridine (1mΨ) and 100% of cytidines with 5-methylcytidine (5mC).

IV. Characterization of Codon-optimized Nucleotide Sequences Encoding Antibodies or Functional Fragments Thereof

[0298] In some aspects of the present disclosure, the codon-optimized nucleotide sequences (e.g., mRNAs) disclosed herein can be tested to determine whether at least one nucleotide sequence property (e.g., stability when exposed to nucleases) or expression

property has been improved with respect to the non-codon-optimized nucleotide sequence.

[0299] The term "expression property" refers to a property of a nucleotide sequence *in vivo* (e.g., translation efficacy of a synthetic mRNA after administration to a subject in need thereof) or *in vitro* (e.g., translation efficacy of a synthetic mRNA tested in an *in vitro* model system). Expression properties include but are not limited to the amount of protein produced by a therapeutic mRNA after administration, and the amount of soluble or otherwise functional protein produced. In some aspects, codon-optimized nucleotide sequences disclosed herein can be evaluated according to the viability of the cells expressing an antibody or functional fragment thereof encoded by an codon-optimized nucleotide sequence disclosed herein (e.g., a mRNA).

[0300] In a particular aspect, a plurality of codon-optimized nucleotide sequences disclosed herein (e.g., mRNAs) containing codon substitutions with respect to the non-optimized candidate nucleic acid sequence can be characterized functionally to measure a property of interest, for example an expression property in an *in vitro* model system, or *in vivo* in a target tissue or cell. Examples of expression properties include but are not limited to, expression levels of an antibody or functional fragment thereof, soluble expression of an antibody or functional fragment thereof, or expression of an antibody or functional fragment thereof in biologically or chemically active form.

a. Optimization of Nucleotide Sequence Intrinsic Properties

[0301] In some aspects of the present disclosure, the desired property optimized is an intrinsic property of the nucleotide sequence (e.g., an mRNA) encoding an antibody or a recombinant protein comprising a functional fragment thereof. For example, the nucleotide sequence (e.g., an mRNA) can be optimized for *in vivo* or *in vitro* stability. In some aspects, the nucleotide sequence can be optimized for expression in a particular target tissue or cell. In some aspects, the nucleotide sequence is optimized to increase its plasma half by preventing its degradation by endo and exonucleases.

[0302] In other aspects, the nucleotide sequence is optimized to increase its resistance to hydrolysis in solution, for example, to lengthen the time that the codon-optimized nucleotide sequence (e.g., an mRNA) or a pharmaceutical composition comprising the codon-optimized nucleic acid sequence can be stored under aqueous conditions with minimal degradation.

[0303] In other aspects, the codon-optimized nucleotide sequence (e.g., an mRNA) can be optimized to increase its resistance to hydrolysis in dry storage conditions, for example, to lengthen the time that the codon-optimized nucleotide sequence can be stored after lyophilization with minimal degradation.

b. Nucleotide Sequences Codon-Optimized for Protein Expression

[0304] In some aspects of the present disclosure, the desired property optimized is the level of expression of an antibody or a recombinant protein comprising a functional fragment thereof encoded by a codon-optimized nucleotide sequence (e.g., an mRNA) disclosed herein. Protein expression levels can be measured using one or more expression systems. In some aspects, expression can be measured in cell culture systems, e.g., CHO cells or HEK293 cells. In some aspects, expression can be measured using *in vitro* expression systems prepared from extracts of living cells, e.g., rabbit reticulocyte lysates, or *in vitro* expression systems prepared by assembly of purified individual components. In other aspects, the protein expression is measured in an *in vivo* system, e.g., mouse, rabbit, monkey, etc.

[0305] In some aspects, protein expression in solution form can be desirable. Accordingly, in some aspects, a candidate sequence can be codon-optimized to yield a codon-optimized nucleotide sequence having optimized levels of expressed proteins in soluble form. Levels of protein expression and other properties such as solubility, levels of aggregation, and the presence of truncation products (i.e., fragments due to proteolysis, hydrolysis, or defective translation) can be measured according to methods known in the art, for example, using electrophoresis (e.g., native or SDS-PAGE) or chromatographic methods (e.g., HPLC, size exclusion chromatography, etc.).

c. Optimization of Target Tissue or Target Cell Viability

[0306] In some aspects, the expression of heterologous therapeutic proteins encoded by a nucleotide sequence (e.g., an mRNA) can have deleterious effects in the target tissue or cell, reducing protein yield, or reducing the quality of the expressed product (e.g., due to the presence of protein fragments or precipitation of the expressed protein in inclusion bodies), or causing toxicity. Heterologous protein expression can also be deleterious to cells transfected with a nucleotide sequence (e.g., an mRNA) for autologous or heterologous transplantation. Accordingly, in some aspects of the present disclosure the

codon-optimized nucleotide sequence (e.g., an mRNA) disclosed herein can be used to increase the viability of target cells expressing the protein encoded by the codon-optimized nucleotide sequence. Changes in cell or tissue viability, toxicity, and other physiological reaction can be measured according to methods known in the art.

V. Vectors, Cells, Methods of Manufacture, and Pharmaceutical Compositions

[0307] The present disclosure also provides a vector or set of vectors comprising a polynucleotide comprising a codon-optimized nucleotide sequence encoding an antibody or a functional fragment thereof disclosed herein or a complement thereof.

[0308] The term "vector" means a construct, which is capable of delivering, and in some aspects, expressing, one or more gene(s) or sequence(s) of interest in a host cell. Examples of vectors include, but are not limited to, viral vectors, naked DNA or RNA expression vectors, plasmid, cosmid or phage vectors, DNA or RNA expression vectors associated with cationic condensing agents, DNA or RNA expression vectors encapsulated in liposomes, and certain eukaryotic cells, such as producer cells.

[0309] Once assembled (by synthesis, site-directed mutagenesis or another method), the polynucleotides disclosed herein (e.g., DNAs or RNAs) encoding an antibody or functional fragment thereof can be inserted into an expression vector and operatively linked to an expression control sequence appropriate for expression of the protein in a desired host.

[0310] A transcriptional unit in a vector disclosed herein generally comprises an assembly of (1) a genetic element or elements having a regulatory role in gene expression, for example, transcriptional promoters or enhancers, (2) a structural or coding sequence which is transcribed into mRNA and translated into protein (e.g., a codon-optimized nucleotide sequence encoding an antibody or functional fragment thereof), and (3) appropriate transcription and translation initiation and termination sequences. Such regulatory elements can include an operator sequence to control transcription. The ability to replicate in a host, usually conferred by an origin of replication, and a selection gene to facilitate recognition of transformants can additionally be incorporated.

[0311] DNA regions are operatively linked when they are functionally related to each other. For example, DNA for a signal peptide (secretory leader) is operatively linked to DNA for a polypeptide if it is expressed as a precursor which participates in the secretion of the polypeptide; a promoter is operatively linked to a coding sequence if it controls the

transcription of the sequence; or a ribosome binding site is operatively linked to a coding sequence if it is positioned so as to permit translation. Structural elements intended for use in yeast expression systems include a leader sequence enabling extracellular secretion of translated protein by a host cell. Alternatively, where recombinant protein is expressed without a leader or transport sequence, it can include an N-terminal methionine residue. This residue can optionally be subsequently cleaved from the expressed recombinant protein to provide a final product.

[0312] RNAs, e.g., mRNAs, can further comprises additional elements for translation. In one aspect, 5' untranslated regions, 3' untranslated regions, microRNA binding sites, 5' cap, polyadenylation sites, IRES regions, or any combination thereof.

Flanking Regions: Untranslated Regions (UTRs)

[0313] Untranslated regions (UTRs) useful for the invention can be transcribed but not translated. 5'UTRs can start at the transcription start site and continue to the start codon but may not include the start codon; whereas, 3 'UTRs can start immediately following the stop codon and continues until the transcriptional termination signal. There is growing body of evidence about the regulatory roles played by the UTRs in terms of stability of the nucleic acid molecule and translation. The regulatory features of a UTR can be incorporated into the polynucleotides, primary constructs and/or mRNA of the present invention to enhance the stability of the molecule. The specific features can also be incorporated to ensure controlled down-regulation of the transcript in case they are misdirected to undesired organs sites.

5' UTR and Translation Initiation

[0314] Natural 5'UTRs bear features which play roles in for translation initiation. They harbor signatures like Kozak sequences which are commonly known to be involved in the process by which the ribosome initiates translation of many genes. Kozak sequences have the consensus CCR(A/G)CCAUGG, where R is a purine (adenine or guanine) three bases upstream of the start codon (AUG), which is followed by another 'G'. 5'UTR also have been known to form secondary structures which are involved in elongation factor binding.

[0315] In one aspect, the polynucleotides disclosed herein includes a 5'UTR so that the proteins encoded by the polynucleotides are expressed at specific target organs, show

enhanced stability and exhibit increased protein production. Likewise, use of 5' UTR for a tissue-specific expression is possible.

[0316] Other non-UTR sequences can be incorporated into the 5' (or 3' UTR) UTRs. For example, introns or portions of introns sequences can be incorporated into the flanking regions of the polynucleotides (e.g., mRNA) of the invention. Incorporation of intronic sequences can increase protein production as well as mRNA levels.

[0317] The 5 'UTR that is useful for the present invention can be a structured UTR such as, but not limited to, 5 'UTRs to control translation.

3' UTR and the AU Rich Elements

[0318] In some aspects, the polynucleotides described herein include a 3 'UTR. 3' UTRs can have stretches of Adenosines and Uridines embedded in them. These AU rich signatures are particularly prevalent in genes with high rates of turnover. Based on their sequence features and functional properties, the AU rich elements (AREs) can be separated into three classes (Chen et al, 1995): Class I AREs contain several dispersed copies of an AUUUA motif within U-rich regions. C-Myc and MyoD contain class I AREs. Class II AREs possess two or more overlapping UUAUUUA(U/A)(U/A) nonamers. Any one of the AU rich elements or any combination thereof can be included in the polynucleotides described herein. In one aspect, 3' UTR AU rich elements (AREs) is used to modulate the stability of polynucleotides (e.g., mRNA).

microRNA Binding Sites

[0319] In other aspects, the polynucleotides (e.g., mRNA) of the invention includes a microRNA binding site or microRNA. microRNAs (or miRNA) are 19-25 nucleotide long noncoding RNAs that bind to a UTR of nucleic acid molecules and modulate gene expression. The polynucleotides (e.g., mRNA) of the invention can comprise one or more microRNA target sequences, microRNA sequences, microRNA binding sites, or microRNA seeds.

5' Capping

[0320] In other aspects, the polynucleotides (e.g., mRNA) comprises a 5' cap. The 5' cap structure of an mRNA is involved in nuclear export, increasing mRNA stability and binds the mRNA Cap Binding Protein (CBP), which is responsible for mRNA stability in the

cell and translation competency through the association of CBP with poly(A) binding protein to form the mature cyclic mRNA species. The cap further assists the removal of 5' proximal introns removal during mRNA splicing.

[0321] Endogenous mRNA molecules can be 5 '-end capped generating a 5'-ppp-5'-triphosphate linkage between a terminal guanosine cap residue and the 5 '-terminal transcribed sense nucleotide of the mRNA molecule. This 5'-guanylate cap can then be methylated to generate an N7-methyl-guanylate residue. The ribose sugars of the terminal and/or anteterminal transcribed nucleotides of the 5' end of the mRNA can optionally also be 2'-0-methylated. 5'-decapping through hydrolysis and cleavage of the guanylate cap structure can target a nucleic acid molecule, such as an mRNA molecule, for degradation.

[0322] In some aspects, a 5' cap for the invention can comprise a non-hydrolyzable cap structure preventing decapping and thus increasing mRNA half-life. Because cap structure hydrolysis requires cleavage of 5'-ppp-5' phosphorodiester linkages, modified nucleotides can be used during the capping reaction.

IRES Sequences

[0323] In certain aspects, the polynucleotides (e.g., mRNA) further comprise an internal ribosome entry site (IRES). First identified as a feature Picorna virus RNA, IRES plays an important role in initiating protein synthesis in absence of the 5' cap structure. An IRES can act as the sole ribosome binding site, or can serve as one of multiple ribosome binding sites of an mRNA. Polynucleotides (e.g., mRNA) containing more than one functional ribosome binding site can encode several peptides or polypeptides that are translated independently by the ribosomes ("multicistronic nucleic acid molecules"). When polynucleotides (e.g., mRNA) are provided with an IRES, further optionally provided is a second translatable region.

Poly-A tails

[0324] In further aspects, the polynucleotides (e.g., mRNAs) of the invention comprises a poly A tail. During RNA processing, a long chain of adenine nucleotides (poly- A tail) can be added to a polynucleotide such as an mRNA molecules in order to increase stability. Immediately after transcription, the 3' end of the transcript can be cleaved to free a 3' hydroxyl. Then poly-A polymerase adds a chain of adenine nucleotides to the RNA.

The process, called polyadenylation, adds a poly-A tail that can be between 100 and 250 residues long.

[0325] In one aspect, the length of a poly-A tail is greater than 30 nucleotides in length. In another aspect, the poly-A tail is greater than 35 nucleotides in length (e.g., at least or greater than about 35, 40, 45, 50, 55, 60, 70, 80, 90, 100, 120, 140, 160, 180, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1,000, 1,100, 1,200, 1,300, 1,400, 1,500, 1,600, 1,700, 1,800, 1,900, 2,000, 2,500, and 3,000 nucleotides). In some aspects, the polynucleotide (e.g., mRNA) includes from about 30 to about 3,000 nucleotides (e.g., from 30 to 50, from 30 to 100, from 30 to 250, from 30 to 500, from 30 to 750, from 30 to 1,000, from 30 to 1,500, from 30 to 2,000, from 30 to 2,500, from 50 to 100, from 50 to 250, from 50 to 500, from 50 to 750, from 50 to 1,000, from 50 to 1,500, from 50 to 2,000, from 50 to 2,500, from 50 to 3,000, from 100 to 500, from 100 to 750, from 100 to 1,000, from 100 to 1,500, from 100 to 2,000, from 100 to 2,500, from 100 to 3,000, from 500 to 750, from 500 to 1,000, from 500 to 1,500, from 500 to 2,000, from 500 to 2,500, from 500 to 3,000, from 1,000 to 1,500, from 1,000 to 2,000, from 1,000 to 2,500, from 1,000 to 3,000, from 1,500 to 2,000, from 1,500 to 2,500, from 1,500 to 3,000, from 2,000 to 3,000, from 2,000 to 2,500, and from 2,500 to 3,000).

[0326] In one aspect, the poly-A tail is designed relative to the length of the overall polynucleotides. This design can be based on the length of the coding region, the length of a particular feature or region (such as the first or flanking regions), or based on the length of the ultimate product expressed from the polynucleotides.

[0327] Proper assembly of the components of a vector can be confirmed by nucleotide sequencing, restriction mapping, and expression of the biologically active polypeptide encoded by the vector (e.g., an antibody of functional fragment thereof) in a suitable host. As is well known in the art, in order to obtain high expression levels of a transfected gene in a target tissue or target cell, the gene must be operatively linked to transcriptional and translational expression control sequences that are functional in the chosen expression host.

[0328] The present disclosure also provides a cell comprising any polynucleotide disclosed herein (e.g., a codon-optimized nucleotide sequence encoding an antibody or functional fragment thereof) or a complement thereof, or the vector or set of vectors disclosed above.

[0329] In some aspects, the cell is an autologous cell, e.g., a cell from a patient to which a codon-optimized nucleotide sequence encoding an antibody or functional fragment thereof is administered, either *in vivo* or *ex vivo*. In some aspects, the cell is a heterologous cell. In some aspects, the heterologous cell can be cell from another patient which has been transfected with a codon-optimized nucleotide sequence encoding an antibody or functional fragment thereof disclosed herein. In some aspects, the heterologous cell can express the antibody or functional fragment thereof transiently. In other aspects, the heterologous cells have been stably transfected. In some aspects, the cells express the antibody or functional fragment thereof constitutively. In other aspects, expression of the antibody or functional fragment thereof is inducible. In some aspects, the cell is a cultured human or animal cell.

[0330] Various mammalian or insect cell culture systems can also be advantageously employed to express codon-optimized nucleotide sequences encoding an antibody or functional fragments disclosed herein (e.g., mRNAs). Expression of the recombinant antibody or functional fragment in mammalian cell model can be used to determine the level of functionality of the optimized nucleotide sequence, e.g., it translational efficacy, and therefore to evaluate whether the codon-optimized nucleotide sequence is suitable for *in vivo* administration to a target tissue or cell in a subject in need thereof.

[0331] The present disclosure also provides a method of expressing a polypeptide comprising a codon-optimized nucleotide sequence encoding an antibody or functional fragment thereof in an expression system comprising contacting an effective amount of (i) the polynucleotide or a complement thereof or (ii) a vector or set of vectors disclosed herein with a cell, wherein the polypeptide encoded by the polynucleotide is expressed in the cell. In some aspects, the polypeptide is expressed *in vitro*. In some aspects, the polypeptide is expressed *in vivo*. In some aspect, a method for expressing or producing a protein encoded a polynucleotide disclosed herein is conducted using an *in vitro* translation system.

[0332] The term "expression system" as used herein refers to any *in vivo*, *in vitro*, or *ex vivo* biological system that is used to produce one or more proteins encoded by a polynucleotide disclosed herein (e.g., a synthetic therapeutic mRNA). In particular aspects of the present disclosure, the term expression system encompasses tissues or cells of a subject to whom a codon-optimized nucleic acid sequence presented in this disclosures (e.g., a synthetic therapeutic mRNA) has been administered.

[0333] Examples of suitable mammalian model cell lines for *in vitro* expression include HEK-293 and HEK-293T, the COS-7 lines of monkey kidney cells, described by Gluzman (Cell 23:175, 1981), and other cell lines including, for example, L cells, C127, 3T3, Chinese hamster ovary (CHO), NSO, HeLa and BHK cell lines.

[0334] Mammalian expression vectors can comprise nontranscribed elements such as an origin of replication, a suitable promoter and enhancer linked to the gene to be expressed, and other 5' or 3' flanking nontranscribed sequences, and 5' or 3' nontranslated sequences, such as necessary ribosome binding sites, a polyadenylation site, splice donor and acceptor sites, and transcriptional termination sequences. Baculovirus systems for production of heterologous proteins in insect cells are reviewed by Luckow and Summers, BioTechnology 6:47 (1988).

[0335] Also provided is a pharmaceutical composition comprising

(i) a polynucleotide comprising a codon-optimized nucleotide sequence encoding an antibody or functional fragment thereof disclosed herein or a complement thereof,

(ii) a vector or set of vectors disclosed herein,

(iii) a cell disclosed herein,

and a pharmaceutically acceptable vehicle or excipient.

[0336] The term "pharmaceutical composition" refers to a preparation which is in such form as to permit the biological activity of the active ingredient to be effective, and which contains no additional components which are unacceptably toxic to a subject to which the composition would be administered. Such composition can be sterile.

VI. Methods of Treatment

[0337] The present disclosure also provides methods to treat a disease or condition in a subject in need thereof comprising administering a therapeutically effective amount of

(i) a polynucleotide comprising a codon-optimized nucleotide sequence encoding an antibody or functional fragment thereof disclosed herein or a complement thereof, or

(ii) a vector or set of vectors disclosed herein,

(iii) a cell disclosed herein,

(iv) a pharmaceutical composition disclosed herein, or

(v) a combination thereof.

[0338] The term "subject" refers to any animal (*e.g.*, a mammal), including, but not limited to humans, non-human primates, rodents, and the like, which is to be the recipient of a particular treatment. Typically, the terms "subject" and "patient" are used interchangeably herein in reference to a human subject.

[0339] An "effective amount" of (i) a polynucleotide disclosed herein or a complement thereof, (ii) a vector or set of vectors disclosed herein, (iii) a cell disclosed herein, (iv) a pharmaceutical composition disclosed, or (v) a combination thereof, is an amount sufficient to carry out a specifically stated purpose, e.g., preventing, treating, alleviating the symptoms, or curing a disease or condition. An "effective amount" can be determined empirically and in a routine manner, in relation to the stated purpose.

[0340] The term "therapeutically effective amount" refers to an amount of (i) a polynucleotide disclosed herein or a complement thereof, (ii) a vector or set of vectors disclosed herein, (iii) a cell disclosed herein, (iv) a pharmaceutical composition disclosed, or (v) a combination thereof, or other drug effective to "treat" a disease or disorder in a subject or mammal.

[0341] Terms such as "treating" or "treatment" or "to treat" or "alleviating" or "to alleviate" refer to both (1) therapeutic measures that cure, slow down, lessen symptoms of, and/or halt progression of a diagnosed pathologic condition or disorder and (2) prophylactic or preventative measures that prevent and/or slow the development of a targeted pathologic condition or disorder. Thus, those in need of treatment include those already with the disorder; those prone to have the disorder; and those in whom the disorder is to be prevented.

[0342] The methods of treatment disclosed herein comprise administering codon-optimized polynucleotides encoding antibodies or antigen binding fragments thereof comprising codon-optimized nucleic acids corresponding, e.g., to the sequences disclosed in TABLE 4. For example, the polynucleotide can be a codon-optimized mRNA encoding the heavy chain of any of the antibodies disclosed in TABLE 4 (SEQ ID NO:1979-2083) or a functional fragment thereof, the light chain of any of the antibodies disclosed in TABLE 4 (SEQ ID NO:2083-2188) or a functional fragment thereof, or combinations of both (e.g., a full antibody comprising a codon-optimized nucleic acid encoding the heavy chain, and a codon-optimized nucleic acid encoding the light chain).

[0343] The composition disclosed herein, e.g., (i) a polynucleotide disclosed herein or a complement thereof, (ii) a vector or set of vectors disclosed herein, (iii) a cell disclosed

herein, (iv) a pharmaceutical composition disclosed, or (v) a combination thereof, wherein the composition results in the *in vivo* expression of an antibody or antigen-binding fragment thereof, can be used to treat a disease or condition mediated by the antigen targeted by the antibody or antigen-binding fragment thereof.

[0344] For example, a composition disclosed herein, e.g., (i) a polynucleotide disclosed herein or a complement thereof, (ii) a vector or set of vectors disclosed herein, (iii) a cell disclosed herein, (iv) a pharmaceutical composition disclosed, or (v) a combination thereof, resulting in the *in vivo* expression of an antibody disclosed in TABLE 6 or a functional fragment thereof, can be used to treat a disease or condition mediated by the target antigen disclosed in TABLE 6. For example, diseases and conditions known in the art to be mediated by TNF-alpha could be treated by the administration of an mRNA comprising a codon-optimized nucleotide sequence encoding adalimumab (e.g., encoding both heavy chain and light chain; encoding either the heavy chain or the light chain; or encoding an antigen-binding molecule comprising a codon-optimized nucleotide sequence encoding an antigen-binding region of adalimumab, such as a VH region, VL region, or one or more CDRs from adalimumab).

TABLE 6: Therapeutic antibodies, their heavy chain and light chain sequences, and their target antigens.

Heavy Chain SEQ ID	Light Chain SEQ ID	INN Antibody Name	Target Antigen	Clinical Trial Phase
1979	2084	Abagovomab	CA-125	3
1980	2085	Abciximab	Integrins alpha-IIb/beta-3 & alpha-V/beta-3	4
1981	2086	Adalimumab	TNF-alpha	4
1982	2087	Alemtuzumab	CAMPATH-1 antigen	4
1983	2088	Alirocumab	NARP-1	2
1984	2089	Amatuximab	Mesothelin	2
1985	2090	Anrukinzumab	Interleukin-13	2
1986	2091	Arcitumomab	carcinoembryonic antigen (CEA)	4
1987	2092	Basiliximab	Interleukin-2 receptor	4
1988	2093	Bavituximab	Phosphatidylserine	2
1989	2094	Benralizumab	CD125	2
1990	2095	Bevacizumab	VEGF A	4
1991	2096	Bezlotoxumab	Clostridium difficile	2
1992	2097	Bimagrumab	ACVR2B	2
1993	2098	Bococizumab	PCSK9	3
1994	2099	Brentuximab	TNFR superfamily member 8	4
1995	2100	Briakinumab	IL-12, IL-23	3
1996	2101	Brodalumab	IL-17	2
1997	2102	Canakinumab	Interleukin-1 beta	4
1998	2103	Cantuzumab	MUC1	2

1999	2104	Carlumab	MCP-1	2
2000	2105	Cetuximab	Epidermal growth factor receptor erbB1	4
2001	2106	Cixutumumab	IGF-1 receptor	2
2002	2107	Clivatuzumab	MUC1	2
2003	2108	Conatumumab	TRAIL-R2	2
2004	2109	Crenezumab	1-40-β-amyloid	2
2005	2110	Dacetuzumab	CD40	2
2006	2111	Daclizumab	Interleukin-2 receptor	4
2007	2112	Dalotuzumab	insulin-like growth factor I receptor	2
2008	2113	Denosumab	TNF ligand superfamily member 11	4
2009	2114	Drozitumab	DR5	2
2010	2115	Dupilumab	IL4	2
2011	2116	Dusigitumab	ILGF2	2
2012	2117	Eculizumab	Complement C5	4
2013	2118	Elotuzumab	SLAMF7	3
2014	2119	Enokizumab	IL9	2
2015	2120	Epratuzumab	CD22	3
2016	2121	Etaracizumab	integrin αvβ3	2
2017	2122	Evolocumab	PCSK9	3
2018	2123	Farletuzumab	folate receptor 1	3
2019	2124	Fasinumab	HNGF	2
2020	2125	Fezakinumab	IL-22	2
2021	2126	Ficlatuzumab	HGF	2
2022	2127	Figitumumab	IGF-1 receptor	3
2023	2128	Fresolimumab	TGF-β	2
2024	2129	Fulranumab	NGF	2
2025	2130	Ganitumab	IGF-I	3
2026	2131	Gantenerumab	beta amyloid	2
2027	2132	Gevokizumab	IL-1β	2
2028	2133	Girentuximab	carbonic anhydrase 9 (CA-IX)	3
2029	2134	Glembatumumab	GPNMB	2
2030	2135	Ibalizumab	CD4	2
2031	2136	Ibritumomab	B-lymphocyte antigen CD20	4
2032	2137	Icrucumab	VEGFR-1	2
2033	2138	Inotuzumab	CD22	3
2034	2139	Intetumumab	CD51	2
2035	2140	Itolizumab	CD6	3
2036	2141	Ixekizumab	IL-17A	3
2037	2142	Lebrikizumab	IL-13	2
2038	2143	Lorvotuzumab	CD56	2
2039	2144	Mavrilimumab	GMCSF receptor α-chain	2
2040	2145	Mepolizumab	IL-5	3
2041	2146	Milatuzumab	CD74	2
2042	2147	Mogamulizumab	CCR4	2
2043	2148	Motavizumab	respiratory syncytial virus	3
2044	2149	Naptumomab	5T4	2
2045	2150	Necitumumab	EGFR	3
2046	2151	Nivolumab	IgG4	2
2047	2152	Obinutuzumab	B-lymphocyte antigen CD20	4
2048	2153	Ocrelizumab	CD20	3
2049	2154	Olaratumab	PDGF-R α	2
2050	2155	Omalizumab	Ig epsilon chain C region	4
2051	2156	Otelixizumab	CD3	3
2052	2157	Oxelumab	OX-40	2
2053	2158	Pateclizumab	LTA	2
2054	2159	Pembrolizumab	Programmed cell death protein 1	4

2055	2160	Pertuzumab	Receptor protein-tyrosine kinase erbB-2	4
2056	2161	Ponezumab	human beta-amyloid	2
2057	2162	Ramucirumab	VEGF receptor 2	4
2058	2163	Rilotumumab	HGF	2
2059	2164	Rituximab	B-lymphocyte antigen CD20	4
2060	2165	Robatumumab	CD221	2
2061	2166	Romosozumab	Sclerostin	3
2062	2167	Rontalizumab	IFN-α	2
2063	2168	Samalizumab	CD200	2
2064	2169	Sarilumab	IL6	3
2065	2170	Secukinumab	IL-17A	3
2066	2171	Sifalimumab	IFN-α	2
2067	2172	Siltuximab	Interleukin-6	4
2068	2173	Sirukumab	IL-6	2
2069	2174	Solanezumab	beta amyloid	3
2070	2175	Tabalumab	BAFF	3
2071	2176	Tanezumab	NGF	3
2072	2177	Tenatumomab	Tenascin C	2
2073	2178	Teplizumab	CD3	3
2074	2179	Tigatuzumab	TRAIL-R2	3
2075	2180	Tildrakizumab	IL23	3
2076	2181	Tocilizumab	Interleukin-6 receptor alpha subunit	4
2077	2182	Tositumomab	B-lymphocyte antigen CD20	4
2078	2183	Tralokinumab	IL-13	2
2079	2184	Trastuzumab	Receptor protein-tyrosine kinase erbB-2	4
2080	2185	Urelumab	4-1BB	2
2081	2186	Ustekinumab	Interleukin-12 Interleukin-23	4
2082	2187	Vedolizumab	Integrin alpha-4/beta-7	4
2083	2188	Veltuzumab	CD20	2

VII. Optimization of Nucleotide Sequences Encoding Therapeutic Antibodies

[0345] The polynucleotides disclosed herein comprise a nucleotide sequence that is not a wild type sequence, i.e., it comprises a nucleotide sequence that has been codon-optimized. These optimized nucleic acid sequences have at least one optimized property with respect to the candidate nucleic acid sequence.

[0346] In some aspects, the nucleotide sequence has been optimized according to a method comprising (i) modifying at least one subsequence in a candidate nucleic acid sequence to generate a ramp subsequence; (ii) substituting at least one codon in a candidate nucleic acid sequence with an alternative codon to increase or decrease uridine content to generate a uridine-modified sequence; (iii) substituting at least one codon in a candidate nucleic acid sequence or the uridine-modified sequence with a fast recharging codon; (iv) substituting at least one codon in a candidate nucleic acid sequence with an alternative codon having a higher codon frequency in the synonymous codon set; (v) substituting at least one natural nucleobase in a candidate nucleic acid sequence with an

alternative synthetic nucleobase; (vi) substituting at least one internucleoside linkage in a candidate nucleic acid sequence with a non-natural internucleoside linkage; or, (vii) a combination thereof, wherein the resulting optimized nucleic acid sequence has at least one optimized property with respect to the candidate nucleic acid sequence.

[0347] In some aspects, the codon optimization method is multiparametric and comprises one, two, three, four, five or six optimization methods selected from the group consisting of (i) modifying at least one subsequence in a candidate nucleic acid sequence to generate a ramp subsequence; (ii) substituting at least one codon in a candidate nucleic acid sequence with an alternative codon to increase or decrease uridine content to generate a uridine-modified sequence; (iii) substituting at least one codon in a candidate nucleic acid sequence or the uridine-modified sequence with a fast recharging codon; (iv) substituting at least one codon in a candidate nucleic acid sequence with an alternative codon having a higher codon frequency in the synonymous codon set; (v) substituting at least one natural nucleobase in a candidate nucleic acid sequence with an alternative synthetic nucleobase; and (vi) substituting at least one internucleoside linkage in a candidate nucleic acid sequence with a non-natural internucleoside linkage. In preferred aspects the substitutions are to the polynucleotide, as above-described, and the encoded antibody sequence is as described herein, for example (i) the amino acid sequence of any one of SEQ ID NOS:1979-2188 or a functional fragment thereof, (ii) a sequence corresponding to any one of the consensus sequences disclosed herein or a combination thereof.

[0348] In some aspects, the multiparametric method comprises replacing at least 5%, at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% of the codons in the candidate nucleic acid sequence. In some aspects, the candidate nucleic acid sequence is SEQ ID NOS: 1979-2188, or a fragment thereof. In some aspects, the fragment comprises (a) one, two, or three VH-CDRs from SEQ ID NOS: 1979-2083; (b) one, two, or three VL-CDRs from SEQ ID NOS: 2084-2188; (c) one, two, three, or four VH framework (FW) regions from SEQ ID NOS: 1979-2083; (d) one, two, three, or four VL framework (FW) regions from SEQ ID NOS: 2084-2188; (e) a VH domain from SEQ ID NOS: 1979-2083; (f) a VL domain from SEQ ID NOS: 2084-2188; (g) a CL domain from SEQ ID NOS: 2084-2188; (h) a CH1 domain from SEQ ID NOS: 1979-2083; (i) a CH2 domain from SEQ ID NOS:

1979-2083; (j) a CH3 domain from SEQ ID NOS: 1979-2083; or, (k) a combination thereof.

[0349] In some particular aspects, codon optimization is conducted by substituting codons in a candidate sequence with alternative codons according to a codon substitution map (e.g., the map in TABLE 1 or any one or more maps included in TABLE 2). In some aspects, the codon substitution map is a limited codon set, e.g., a codon set wherein less than the native number of codons is used to encode the 20 natural amino acids, a subset of the 20 natural amino acids, or an expanded set of amino acids including, for example, non-natural amino acids.

[0350] A codon set can be optimized to generate a codon substitution map by reducing the codon number, by replacing natural codons with codons having unnatural bases, expanding the codon number to incorporate non-natural amino acids, or even introducing codons that have lengths different than 3. For example, 4 base codons are disclosed in Taira et al. (2005) J. Biosci. Bioeng. 99:473-6; and 5 base codons are disclosed in Hohsaka et al. (2001) Nucl. Acids Res. 29:3646-3651), both of which are herein incorporated by reference in their entireties.

[0351] The genetic code is highly similar among all organisms and can be expressed in a simple table with 64 entries which would encode the 20 standard amino acids involved in protein translation plus start and stop codons. The genetic code is degenerate, i.e., in general, more than one codon specifies each amino acid. For example, the amino acid leucine is specified by the UUA, UUG, CUU, CUC, CUA, or CUG codons, while the amino acid serine is specified by UCA, UCG, UCC, UCU, AGU, or AGC codons (difference in the first, second, or third position). Native genetic codes comprise 62 codons encoding naturally occurring amino acids. Thus, in some aspects of the methods disclosed herein codon substitution maps comprising less than 62 codons to encode 20 amino acids, and can comprise 61, 60, 59, 58, 57, 56, 55, 54, 53, 52, 51, 50, 49, 48, 47, 46, 45, 44, 43, 42, 41, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, or 20 codons.

[0352] In some aspects, the codon substitution map comprises less than 20 codons. For example, if a protein contains less than 20 types of amino acids, such protein could be encoded by a codon substitution map with less than 20 codons. Accordingly, in some aspects, a codon substitution map comprises as many codons as different types of amino acids are present in the protein encoded by the candidate nucleic acid sequence.

[0353] In some aspects, at least one amino acid selected from the group consisting of Ala, Arg, Asn, Asp, Cys, Gln, Glu, Gly, His, Ile, Leu, Lys, Phe, Pro, Ser, Thr, Tyr, and Val, i.e., amino acids which are naturally encoded by more than one codon, is encoded with less codons than the naturally occurring number of synonymous codons. For example, in some aspects, Ala can be encoded in the codon-optimized nucleic acid sequence by 3, 2 or 1 codons; Cys can be encoded in the codon-optimized nucleic acid sequence by 1 codon; Asp can be encoded in the codon-optimized nucleic acid sequence by 1 codon; Glu can be encoded in the codon-optimized nucleic acid sequence by 1 codon; Phe can be encoded in the codon-optimized nucleic acid sequence by 1 codon; Gly can be encoded in the codon-optimized nucleic acid sequence by 3 codons, 2 codons or 1 codon; His can be encoded in the codon-optimized nucleic acid sequence by 1 codon; Ile can be encoded in the codon-optimized nucleic acid sequence by 2 codons or 1 codon; Lys can be encoded in the codon-optimized nucleic acid sequence by 1 codon; Leu can be encoded in the codon-optimized nucleic acid sequence by 5 codons, 4 codons, 3 codons, 2 codons or 1 codon; Asn can be encoded in the codon-optimized nucleic acid sequence by 1 codon; Pro can be encoded in the codon-optimized nucleic acid sequence by 3 codons, 2 codons, or 1 codon; Gln can be encoded in the codon-optimized nucleic acid sequence by 1 codon; Arg can be encoded in the codon-optimized nucleic acid sequence by 5 codons, 4 codons, 3 codons, 2 codons, or 1 codon; Ser can be encoded in the codon-optimized nucleic acid sequence by 5 codons, 4 codons, 3 codons, 2 codons, or 1 codon; Thr can be encoded in the codon-optimized nucleic acid sequence by 3 codons, 2 codons, or 1 codon; Val can be encoded in the codon-optimized nucleic acid sequence by 3 codons, 2 codons, or 1 codon; and, Tyr can be encoded in the codon-optimized nucleic acid sequence by 1 codon.

[0354] In some aspects, at least one amino acid selected from the group consisting of Ala, Arg, Asn, Asp, Cys, Gln, Glu, Gly, His, Ile, Leu, Lys, Phe, Pro, Ser, Thr, Tyr, and Val, i.e., amino acids which are naturally encoded by more than one codon, is encoded by a single codon in the codon substitution map.

[0355] In some specific aspects, the codon-optimized nucleic acid sequence is a DNA and the codon substitution map consists of 20 codons, wherein each codon encodes one of 20 amino acids. In some aspects, the codon-optimized nucleic acid sequence is a DNA and the codon substitution map comprises at least one codon selected from the group consisting of GCT, GCC, GCA, and GCG; at least a codon selected from the group consisting of CGT, CGC, CGA, CGG, AGA, and AGG; at least a codon selected from

AAT or ACC; at least a codon selected from GAT or GAC; at least a codon selected from TGT or TGC; at least a codon selected from CAA or CAG; at least a codon selected from GAA or GAG; at least a codon selected from the group consisting of GGT, GGC, GGA, and GGG; at least a codon selected from CAT or CAC; at least a codon selected from the group consisting of ATT, ATC, and ATA; at least a codon selected from the group consisting of TTA, TTG, CTT, CTC, CTA, and CTG; at least a codon selected from AAA or AAG; an ATG codon; at least a codon selected from TTT or TTC; at least a codon selected from the group consisting of CCT, CCC, CCA, and CCG; at least a codon selected from the group consisting of TCT, TCC, TCA, TCG, AGT, and AGC; at least a codon selected from the group consisting of ACT, ACC, ACA, and ACG; a TGG codon; at least a codon selected from TAT or TAC; and, at least a codon selected from the group consisting of GTT, GTC, GTA, and GTG.

[0356] In other aspects, the codon-optimized nucleic acid sequence is an RNA (e.g., an mRNA) and the codon substitution map consists of 20 codons, wherein each codon encodes one of 20 amino acids. In some aspects, the codon-optimized nucleic acid sequence is an RNA and the codon substitution map comprises at least one codon selected from the group consisting of GCU, GCC, GCA, and GCG; at least a codon selected from the group consisting of CGU, CGC, CGA, CGG, AGA, and AGG; at least a codon selected from AAU or ACC; at least a codon selected from GAU or GAC; at least a codon selected from UGU or UGC; at least a codon selected from CAA or CAG; at least a codon selected from GAA or GAG; at least a codon selected from the group consisting of GGU, GGC, GGA, and GGG; at least a codon selected from CAU or CAC; at least a codon selected from the group consisting of AUU, AUC, and AUA; at least a codon selected from the group consisting of UUA, UUG, CUU, CUC, CUA, and CUG; at least a codon selected from AAA or AAG; an AUG codon; at least a codon selected from UUU or UUC; at least a codon selected from the group consisting of CCU, CCC, CCA, and CCG; at least a codon selected from the group consisting of UCU, UCC, UCA, UCG, AGU, and AGC; at least a codon selected from the group consisting of ACU, ACC, ACA, and ACG; a UGG codon; at least a codon selected from UAU or UAC; and, at least a codon selected from the group consisting of GUU, GUC, GUA, and GUG.

[0357] In some specific aspects, the codon substitution map has been optimized for *in vivo* expression of an optimized nucleic acid sequence (e.g., a synthetic mRNA) following administration to a certain tissue or cell. Thus, in some aspects, the optimized

property with respect to the candidate nucleic acid sequence is optimized *in vivo* expression following administration to a certain tissue or cell in a subject in need thereof.

[0358] In some aspects, the codon substitution map comprises at least one codon consisting of more than 3 nucleobases, for example, 4 nucleobases or 5 nucleobases. In some aspects, the optimized codon set comprises at least one codon encoding an unnatural amino acid (i.e., a non-canonical amino acid). See, e.g., Liu et al. (1997) Proc. Natl. Acad Sci. USA 94:10092-10097; Link et al. (2003) Curr. Opin. Biotechnol. 14:603-609; Sakamoto et al. (2002) Nucl. Acids Res. 30:4692-4699; Zhang et al. (2013) Curr. Opin. Struct. Biol. 23:581-587; Ma (2003) Chem. Today, 65; Dougherty (2000) Curr Opin Chem Biol. 6:645; Kitamura et al. (2005) Chem. Int. Ed. 44: 1549; Ooi et al. (2007) Aldrichimica Acta 40:77; Rutjes et al. (2005) J. Org. Biol. Chem. 3:3435; Rutjes et al. (2000) J. Chem. Soc., Perkin Trans. 1:4197; Vignola (2003) Am. Chem. Soc. 125:450; Dalko (2004) Chem. Int. Ed. 43:5138; Lelais (2004) Biopolymers 76:206; and Seebach et al. (2004) Chem. & Biodiv. 1:1111, all of which are herein incorporated by reference in their entireties.

[0359] In some aspects, the codon substitution map comprises at least one codon comprising an unnatural nucleobase. In some aspects, the unnatural nucleobase is an adenosine analog. In other aspects, the unnatural nucleobase in a cytidine analog. In other aspects, the unnatural nucleobase is a thymidine analog. In other aspects, the unnatural nucleobase is a guanidine analog. In yet other aspects, the unnatural nucleobase is a uridine analog.

[0360] In some specific aspects, the codon substitution map comprises at least one codon comprising a nucleobase selected from the group consisting of 5-trifluoromethyl-cytosine, 1-methyl-pseudo-uracil, 5-hydroxymethyl-cytosine, 5-bromo-cytosine, 5-methoxy-uracil, 1-ethyl-pseudo-uracil, or 5-methyl-cytosine. See, for example, International Publication Nos. WO2014093924A1 and WO2013052523 A1, which are herein incorporated by reference in their entireties.

[0361] In some specific aspects, at least one codon in the codon substitution map has the second highest, the third highest, the fourth highest, the fifth highest or the sixth highest frequency in the synonymous codon set. In some specific aspects, at least one codon in the codon substitution map has the second lowest, the third lowest, the fourth lowest, the fifth lowest, or the sixth lowest frequency in the synonymous codon set.

[0362] See also, U.S. Publ. No. US20110082055, Int'l. Publ. No. WO2000018778. See also Pechmann & Frydman, Nature Structural & Molecular Biology 20:237-244 (2013); does Reis et al., Nucleic Acids Research 32: 5036–5044 (2004); Kim et al., Science 348:444-448 (2015), all of which are incorporated herein by reference in their entireties.

[0363] All patents and publications referred to herein are expressly incorporated by reference in their entireties.

VIII. Embodiments

[0364] E1. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes TVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQES VTEQDSKDSTYSLSX[1]TLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNR (SEQ ID NO: 2200),

wherein X^1 is selected from N and S.

[0365] E2. The polynucleotide according to embodiment E1, wherein the nucleotide sequence encodes SEQ ID NO:2189.

[0366] E3. The polynucleotide according to any one of embodiments E1 or E2, wherein the nucleotide sequence encodes a kappa light chain constant domain of an antibody or a fragment thereof.

[0367] E4. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes PKAAPSVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTT PSKQSNNKYAASSYLSLTPEQWKSHX[2]SYSCQVTHEGSTVEKTVAPX[3]ECS (SEQ ID NO: 2201),

wherein X^2 is selected from R and K, and X^3 is selected from T and A.

[0368] E5. The polynucleotide according to embodiment E4, wherein the nucleotide sequence encodes SEQ ID NO: 2190.

[0369] E6. The polynucleotide according to any one of embodiments E4 or E5, wherein the nucleotide sequence encodes a lambda light chain constant domain of an antibody or a fragment thereof.

[0370] E7. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $SX^4GPSVX^5PLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGVHTFPAVLQSSG$ $LYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKX^6X^7$ (SEQ ID NO: 2202) wherein X^4 is an optional ASTK sequence, X^5 is selected from F and L, X^6 is selected from K and R, and X^7 is selected from V and A.

[0371] E8. The polynucleotide according to embodiment E7, wherein the nucleotide sequence encodes SEQ ID NO: 2191.

[0372] E9. The polynucleotide according to any one of embodiments E7 or E8, wherein the nucleotide sequence encodes a CH1 domain of an IgG1 antibody or a fragment thereof.

[0373] E10. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $APEX^8X^9GX^{10}PSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNX^{11}YVDGV$ $EVHNAKTKPREEQYX^{12}STYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK$ TISKAK (SEQ ID NO: 2203)

wherein X^8 is selected from L and A, X^9 is selected from L and A, X^{10} is selected from G and A, and X^{11} is selected from V and W, and X^{12} is selected from N and A.

[0374] E11. The polynucleotide according to embodiment E10, wherein the nucleotide sequence encodes SEQ ID NO: 2192.

[0375] E12. The polynucleotide according to any one of embodiments E10 or E11, wherein the nucleotide sequence encodes a CH2 domain of an IgG1 antibody or a fragment thereof.

[0376] E13. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or

any combination thereof), wherein the nucleotide sequence encodes GQPREPQVYTLPPSRX[13]EX[14]TKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYK TTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG (SEQ ID NO: 2204)

wherein X^{13} is selected from E and D, and X^{14} is selected from M and L.

[0377] E14. The polynucleotide according to embodiment E13, wherein the nucleotide sequence encodes SEQ ID NO: 2193.

[0378] E15. The polynucleotide according to any one of embodiments E13 or E14, wherein the nucleotide sequence encodes a CH3 domain of an IgG1 antibody or a fragment thereof.

[0379] E16. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes SASTKGPSVFPLAPCSRSTSESTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPA VLQSSGLYSLSSVVTVX[15]SSNFGTQTYTCNVDHKPSNTKVDKTV (SEQ ID NO: 2205)

wherein X^{15} is selected from P and T.

[0380] E17. The polynucleotide according to embodiment E16, wherein the nucleotide sequence encodes SEQ ID NO: 2194.

[0381] E18. The polynucleotide according to any one of embodiments E16 or E17, wherein the nucleotide sequence encodes a CH1 domain of an IgG2 antibody or a fragment thereof.

[0382] E19. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes APPVAGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFNWYVDGX[16]EV HNAKTKPREEQFNSTFRVVSVLTVVHQDWLNGKEYKCKVSNKGLPX[17]X[18]IEKTI SKTK (SEQ ID NO: 2206)

wherein X^{16} is selected from V and M, X^{17} is selected from A and S; and X^{18} is selected from P and S.

[0383] E20. The polynucleotide according to embodiment E19, wherein the nucleotide sequence encodes SEQ ID NO: 2195.

[0384] E21. The polynucleotide according to any one of embodiments E19 or E20, wherein the nucleotide sequence encodes a CH2 domain of an IgG2 antibody or a fragment thereof.

[0385] E22. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes GQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTP PMLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG (SEQ ID NO: 2196).

[0386] E23. The polynucleotide according to embodiment E22, wherein the nucleotide sequence encodes a CH3 domain of an IgG2 antibody or a fragment thereof.

[0387] E24. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes SASTKGPSVFPLAPCSRSTSESTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPA VLQSSGLYSLSSVVTVPSSSLGTKTYTCNVDHKPSNTKVDKRV (SEQ ID NO: 2197).

[0388] E25. The polynucleotide according to embodiment E24, wherein the nucleotide sequence encodes a CH1 domain of an IgG4 antibody or a fragment thereof.

[0389] E26. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes APEFLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSQEDPEVQFNWYVDGVEVH NAKTKPREEQFNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKGLPSSIEKTISKA K (SEQ ID NO: 2198).

[0390] E27. The polynucleotide according to embodiment E26, wherein the nucleotide sequence encodes a CH2 domain of an IgG4 antibody or a fragment thereof.

[0391] E28. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes GQPREPQVYTLPPSQEEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTP PVLDSDGSFFLYSRLTVDKSRWQEGNVFSCSVMHEALHNHYTQKSLSLSLG (SEQ ID NO: 2199).

[0392] E29. The polynucleotide according to embodiment E28, wherein the nucleotide sequence encodes a CH3 domain of an IgG4 antibody or a fragment thereof.

[0393] E30. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $X^1X^2X^3LTQX^4X^5X^6VSX^7X^8X^9GX^{10}X^{11}X^{12}X^{13}X^{14}X^{15}C$ (SEQ ID NO: 2235) wherein

X^1 is selected from Q, D, E and S;

X^2 is selected from S, I, A, and Y;

X^3 is selected from V, Q, A, and E;

X^4 is selected from P and D;

X^5 is selected from P, N, and A;

X^6 is selected from S and A;

X^7 is selected from G, T, A, and V;

X^8 is selected from A and S;

X^9 is selected from P and L;

X^{10} is selected from Q, K, and S;

X^{11} is selected from R, K, T, and S;

X^{12} is selected from V, I, and A;

X^{13} is selected from T, K, and R;

X^{14} is selected from I and L; and,

X^{15} is selected from S at T.

[0394] E31. The polynucleotide according to embodiment E30, wherein the nucleotide sequence encodes a sequence identical to QSVLTQPPSVSGAPGQRVTISC (SEQ ID NO: 2207) except for at least one substitution selected from Q1(DES), S2(IAY),

V3(QAE), P7D, P8(NA), S9A, G12(TAV), A13S, P14L, Q16(KS), R17(KTS), V18(IA), T19(KR), I20L, and S21T.

[0395] E32. The polynucleotide according to embodiment E31, wherein the nucleotide sequence encodes QSVLTQPPSVSGAPGQRVTISC (SEQ ID NO: 2207).

[0396] E33. The polynucleotide according to any one of embodiments E30 to E32, wherein the nucleotide sequence encodes the first framework region (FW1) of a lambda light chain variable domain.

[0397] E34. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $WYQX^1X^2X^3GX^4X^5PX^6X^7X^8I$ (SEQ ID NO: 2236)

wherein

X^1 is selected from Q and L;

X^2 is selected from L,Y, H, and K;

X^3 is selected from P and E;

X^4 is selected from T, R, K, and Q;

X^5 is selected from A and S;

X^6 is selected from K, T, V and I;

X^7 is selected from L and T; and

X^8 is selected from L, M, and V.

[0398] E35. The polynucleotide according to embodiment E34, wherein the nucleotide sequence encodes a sequence identical to WYQQLPGTAPKLLI (SEQ ID NO: 2208) except for at least one substitution selected from Q4L, L5(YHK), P6E, T8(RKQ), A9S, K11(TVI), L12T, and L13(MV).

[0399] E36. The polynucleotide according to embodiment E34, wherein the nucleotide sequence encodes WYQQLPGTAPKLL (SEQ ID NO: 2208).

[0400] E37. The polynucleotide according to any one of embodiments E34 to E36, wherein the nucleotide sequence encodes the second framework region (FW2) of a lambda light chain variable domain.

[0401] E38. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or

any combination thereof), wherein the nucleotide sequence encodes RFSGSX^{1}SX^{2}X^{3}X^{4}AX^{5}LX^{6}IX^{7}X^{8}X^{9}X^{10}X^{11}X^{12}DEAX^{13}YX^{14}C (SEQ ID NO: 2237) wherein

X^1 is selected from K, N, S, and I;

X^2 is selected from G and S;

X^3 is selected from T and N;

X^4 is selected from S and T;

X^5 is selected from S, T, and F;

X^6 is selected from A, T, and G;

X^7 is selected from T, H, and S;

X^8 is selected from G, N, and R;

X^9 is selected from L, V, and A;

X^{10} is selected from Q, E, and A;

X^{11} is selected from A, T, and I;

X^{12} is selected from E and G;

X^{13} is selected from D and I; and,

X^{14} is selected from Y and F.

[0402] E39. The polynucleotide according to embodiment E38, wherein the nucleotide sequence encodes a sequence identical to RFSGSKSGTSASLAITGLQAEDEADYYC (SEQ ID NO: 2209) except for at least one substitution selected from K6(NSI), G8S, T9N, S10T, S12(TF), A14(TG), T16(HS), G17(NR), L18(VA), Q19(EA), A20(TI), E21G, D25I, and Y27F.

[0403] E40. The polynucleotide according to embodiment E39, wherein the nucleotide sequence encodes RFSGSKSGTSASLAITGLQAEDEADYYC (SEQ ID NO: 2209).

[0404] E41. The polynucleotide according to any one of embodiments E38 to E40, wherein the nucleotide sequence encodes the third framework region (FW3) of a lambda light chain variable domain.

[0405] E42. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes FGX^{1}GTX^{2}X^{3}TVL (SEQ ID NO:2238)

wherein

X^1 is selected from G and T;

X^2 is selected from K and Q; and

X^3 is selected from L and V.

[0406] E43. The polynucleotide according to embodiment E42, wherein the nucleotide sequence encodes a sequence identical to FGGGTKLTVL (SEQ ID NO: 2210) except for at least one substitution selected from G3T, K6Q, and L7V.

[0407] E44. The polynucleotide according to embodiment E43, wherein the nucleotide sequence encodes FGGGTKLTVL (SEQ ID NO: 2210).

[0408] E45. The polynucleotide according to any one of embodiments E42 to E44, wherein the nucleotide sequence encodes the fourth framework region (FW4) of a lambda light chain variable domain.

[0409] E46. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $X^1X^2QX^3TQX^4X^5SX^6X^7SASX^8CDRVTX^9X^{10}C$ (SEQ ID NO: 2239)

wherein

X^1 is selected from D and A;

X^2 is selected from I and V;

X^3 is selected from M, L, and V;

X^4 is selected from S and F;

X^5 is selected from P and T;

X^6 is selected from S and T;

X^7 is selected from L and V;

X^8 is selected from V, I, and A;

X^9 is selected from I and M; and,

X^{10} is selected from T and S.

[0410] E47. The polynucleotide according to embodiment E46, wherein the nucleotide sequence encodes a sequence identical to DIQMTQSPSSLSASVCDRVTITC (SEQ ID NO: 2211) except for at least one substitution selected from D1A, I2V, M4(LV), S7F, P8T, S10T, L11V, V15(IA), I21M, and T22S.

[0411] E48. The polynucleotide according to embodiment E47, wherein the nucleotide sequence encodes DIQMTQSPSSLSASVCDRVTITC (SEQ ID NO: 2211).

[0412] E49. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $DX^1X^2X^3TQX^4PX^5SX^6X^7X^8X^9X^{10}GX^{11}X^{12}X^{13}X^{14}X^{15}X^{16}C$ (SEQ ID NO: 2243) wherein

X^1 is selected from I and V;

X^2 is selected from V, L, and Q;

X^3 is selected from M and L;

X^4 is selected from S and T;

X^5 is selected from L and D;

X^6 is selected from L and V;

X^7 is selected from P, S and A;

X^8 is selected from V and M;

X^9 is selected from T and S;

X^{10} is selected from P and L;

X^{11} is selected from E and Q;

X^{12} is selected from P and R;

X^{13} is selected from A and V;

X^{14} is selected from S and T;

X^{15} is selected from I, M, and L; and

X^{16} is selected from S and N.

[0413] E50. The polynucleotide according to embodiment E49, wherein the nucleotide sequence encodes a sequence identical to DIVMTQSPLSLPVTPGEPASISC (SEQ ID NO: 2215) except for at least one substitution selected from I2V, V3(LQ), M4L, S7T, L9D, L11V, P12(SA), V13M, T14S, P15L, E17Q, P18R, A19V, S20T, I21(ML), and S22N.

[0414] E51. The polynucleotide according to embodiment E50, wherein the nucleotide sequence encodes DIVMTQSPLSLPVTPGEPASISC (SEQ ID NO: 2215).

[0415] E52. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or

any combination thereof), wherein the nucleotide sequence encodes $X^1X^2VX^3TQSPX^4TLSX^5SPGERATLSC$ (SEQ ID NO: 2247)

wherein

X^1 is selected from E and D;

X^2 is selected from I and T;

X^3 is selected from L and M;

X^4 is selected from G and A; and,

X^5 is selected from L and V.

[0416] E53. The polynucleotide according to embodiment E52, wherein the nucleotide sequence encodes a sequence identical to EIVLTQSPGTLSLSPGERATLSC (SEQ ID NO: 2219) except for at least one substitution selected from E1D, I2T, L4M, G9A, and L13V.

[0417] E54. The polynucleotide according to embodiment E53, wherein the nucleotide sequence encodes EIVLTQSPGTLSLSPGERATLSC (SEQ ID NO: 2219).

[0418] E55. The polynucleotide according to any one of embodiments E46 to E54, wherein the nucleotide sequence encodes the first framework region (FW1) of a kappa light chain variable domain.

[0419] E56. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $WX^1X^2X^3X^4PX^5KX^6X^7X^8X^9X^{10}IX^{11}$ (SEQ ID NO: 2240)

wherein

X^1 is selected from Y and F;

X^2 is selected from Q and L;

X^3 is selected from Q and H;

X^4 is selected from K and I;

X^5 is selected from G and E;

X^6 is selected from A and V;

X^7 is selected from P and V;

X^8 is selected from K and Q;

X^9 is selected from L, T, S, R, P, and V;

X^{10} is selected from L and W; and,

X^{11} is selected from Y and S.

[0420] E57. The polynucleotide according to embodiment E56, wherein the nucleotide sequence encodes a sequence identical to WYQQKPGKAPKLLIY (SEQ ID NO: 2212) except for at least one substitution selected from Y2F, Q3L, Q4H, K5I, G7E, A9V, P10V, K11Q, L12(TSRPV), L13W, and Y15S.

[0421] E58. The polynucleotide according to embodiment E57, wherein the nucleotide sequence encodes WYQQKPGKAPKLLIY (SEQ ID NO: 2212).

[0422] E59. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes
$WX^1X^2QX^3X^4GQX^5PX^6X^7LIX^8$ (SEQ ID NO: 2244)
wherein

X^1 is selected from Y, F, and W;

X^2 is selected from L and Q;

X^3 is selected from K and R;

X^4 is selected from P and S;

X^5 is selected from S and P;

X^6 is selected from Q, K, R, and N;

X^7 is selected from L and R; and,

X^8 is selected from Y and W.

[0423] E60. The polynucleotide according to embodiment E59, wherein the nucleotide sequence encodes a sequence identical to WYLQKPGQSPQLLIY (SEQ ID NO: 2216) except for at least one substitution selected from Y2(FW), L3Q, K5R, P6S, S9P,Q11(KRN), L12R, and Y15W.

[0424] E61. The polynucleotide according to embodiment E60, wherein the nucleotide sequence encodes WYLQKPGQSPQLLIY (SEQ ID NO: 2216).

[0425] E62. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes
$WX^1X^2QX^3PGQAPRX^4LIX^5$ (SEQ ID NO: 2248)
wherein

X^1 is selected from Y and F;

X^2 is selected from Q and R;

X^3 is selected from K and R;

X^4 is selected from L and P; and

X^5 is selected from Y, R, and K.

[0426] E63. The polynucleotide according to embodiment E62, wherein the nucleotide sequence encodes a sequence identical to WYQQKPGQAPRLLIY (SEQ ID NO: 2220) except for at least one substitution selected from Y2F, Q3R, K5R, L12P, and Y15(RK).

[0427] E64. The polynucleotide according to embodiment E63, wherein the nucleotide sequence encodes WYQQKPGQAPRLLIY (SEQ ID NO: 2220).

[0428] E65. The polynucleotide according to any one of embodiments E56 to E64, wherein the nucleotide sequence encodes the second framework region (FW2) of a kappa light chain variable domain.

[0429] E66. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $RFSGSX^1SGX^2X^3X^4X^5X^6TISSLX^7X^8X^9DX^{10}AX^{11}YX^{12}C$ (SEQ ID NO: 2241) wherein

X^1 is selected from G and R;

X^2 is selected from T and Q;

X^3 is selected from D, E, and Y;

X^4 is selected from F and Y;

X^5 is selected from T and S;

X^6 is selected from L and F;

X^7 is selected from Q and E;

X^8 is selected from P, Q, A, and S;

X^9 is selected from E and D;

X^{10} is selected from F, I, S, L, V, and T;

X^{11} is selected from T, S, and V; and,

X^{12} is selected from Y and F.

[0430] E67. The polynucleotide according to embodiment E66, wherein the nucleotide sequence encodes a sequence identical to RFSGSGSGTDFTLTISSLQPEDFATYYC

(SEQ ID NO: 2213) except for at least one substitution selected from G6R, T9Q, D10(EY), F11Y, T12S, L13F, Q19E, P20(QAS), E21D, F23(ISLVT), T25(SV), and Y27F.

[0431] E68. The polynucleotide according to embodiment E67, wherein the nucleotide sequence encodes RFSGSGSGTDFTLTISSLQPEDFATYYC (SEQ ID NO: 2213).

[0432] E69. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $RFSGSGSX^1TX^2FTLX^3ISX^4X^5X^6AX^7DVX^8X^9X^{10}X^{11}C$ (SEQ ID NO: 2245) wherein

X^1 is selected from G and A;

X^2 is selected from D and A;

X^3 is selected from K, R, and T;

X^4 is selected from R and S;

X^5 is selected from V and L;

X^6 is selected from E and Q;

X^7 is selected from E and Q;

X^8 is selected from G and A;

X^9 is selected from V, D, and F;

X^{10} is selected from Y and W; and,

X^{11} is selected from Y, F, and W.

[0433] E70. The polynucleotide according to embodiment E69, wherein the nucleotide sequence encodes a sequence identical to RFSGSGSGTDFTLKISRVEAEDVGVYYC (SEQ ID NO: 2217) except for at least one substitution selected from G8A, D10A, K14(RT), R17S, V18L, E19Q, E21Q, G24A, V25(DF), Y26W, and Y27(FW).

[0434] E71. The polynucleotide according to embodiment E70, wherein the nucleotide sequence encodes RFSGSGSGTDFTLKISRVEAEDVGVYYC (SEQ ID NO: 2217).

[0435] E72. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $RFSGSGSGTX^1X^2TLTISX^3LX^4X^5EDFAX^6X^7YC$ (SEQ ID NO: 2249)

wherein

X^1 is selected from D and E;

X^2 is selected from F and S;

X^3 is selected from R and S;

X^4 is selected from E and Q;

X^5 is selected from P and S;

X^6 is selected from V and T; and,

X^7 is selected from Y and F.

[0436] E73. The polynucleotide according to embodiment E72, wherein the nucleotide sequence encodes a sequence identical to RFSGSGSGTDFTLTISRLEPEDFAVYYC (SEQ ID NO: 2221) except for at least one substitution selected from D10E, F11S, R17S, E19Q, P20S, V25T, and Y26F.

[0437] E74. The polynucleotide according to embodiment E73, wherein the nucleotide sequence encodes RFSGSGSGTDFTLTISRLEPEDFAVYYC (SEQ ID NO: 2221).

[0438] E75. The polynucleotide according to any one of embodiments E66 to E74, wherein the nucleotide sequence encodes the third framework region (FW3) of a kappa light chain variable domain.

[0439] E76. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes

$X^1GX^2GTX^3X^4X^5X^6X^7$ (SEQ ID NO: 2242)

wherein

X^1 is selected from F and L;

X^2 is selected from Q, G, and S;

X^3 is selected from K and R;

X^4 is selected from V and L;

X^5 is selected from E, D, and Q;

X^6 is selected from I and V; and,

X^7 is selected from K and T.

[0440] E77. The polynucleotide according to embodiment E76, wherein the nucleotide sequence encodes a sequence identical to FGQGTKVEIK (SEQ ID NO: 2214) except for at least one substitution selected from F1L, Q3(GS), K6R, V7L, E8(DQ), I9V, and K10T.

[0441] E78. The polynucleotide according to embodiment E77, wherein the nucleotide sequence encodes FGQGTKVEIK (SEQ ID NO: 2214).

[0442] E79. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $FGX^1GTX^2X^3X^4X^5K$ (SEQ ID NO: 2246)

wherein

X^1 is selected from Q, A, P, and G;

X^2 is selected from K and R;

X^3 is selected from V and L;

X^4 is selected from E and Q; and

X^5 is selected from I and L.

[0443] E80. The polynucleotide according to embodiment E79, wherein the nucleotide sequence encodes a sequence identical to FGQGTKVEIK (SEQ ID NO: 2218) except for at least one substitution selected from Q3(APG), K6R, V7L, E8Q, and I9L.

[0444] E81. The polynucleotide according to embodiment E80, wherein the nucleotide sequence encodes FGQGTKVEIK (SEQ ID NO: 2218).

[0445] E82. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $FX^1X^2GTX^3X^4X^5IK$ (SEQ ID NO: 2250)

wherein

X^1 is selected from G and C;

X^2 is selected from Q, G, and P;

X^3 is selected from K and R;

X^4 is selected from V, L, and A; and,

X^5 is selected from E and D.

[0446] E83. The polynucleotide according to embodiment E82, wherein the nucleotide sequence encodes a sequence identical to FGQGTKVEIK (SEQ ID NO: 2222) except for at least one substitution selected from G2C, Q3(GP), K6R, V7(LA), and E8D.

[0447] E84. The polynucleotide according to embodiment E83, wherein the nucleotide sequence encodes FGQGTKVEIK (SEQ ID NO: 2222).

[0448] E85. The polynucleotide according to any one of embodiments E76 to E84, wherein the nucleotide sequence encodes the fourth framework region (FW4) of a kappa light chain variable domain.

[0449] E86. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $X^1X^2X^3X^4X^5X^6SGGX^7X^8X^9X^{10}X^{11}GX^{12}SX^{13}X^{14}LX^{15}C$ (SEQ ID NO: 2251)
wherein

X^1 is selected from E, D, and Q;

X^2 is selected from V and A;

X^3 is selected from Q, E, and K;

X^4 is selected from L and V;

X^5 is selected from V and L;

X^6 is selected from E and Q;

X^7 is selected from G, K, and D;

X^8 is selected from L and V;

X^9 is selected from V, L, and E;

X^{10} is selected from Q, R and K;

X^{11} is selected from P, S, and L;

X^{12} is selected from G and R;

X^{13} is selected from L and R;

X^{14} is selected from R and K; and,

X^{15} is selected from S and D.

[0450] E87. The polynucleotide according to embodiment E86, wherein the nucleotide sequence encodes a sequence identical to EVQLVESGGGLVQPGGSLRLSC (SEQ ID NO: 2223) except for at least one substitution selected from E1(DQ), V2A, Q3(EK), L4V, V5L, E6Q, G10(KD), L11V, V12(LE), Q13(RK), P14(SL), G16R, L18R, R19K, and S21D.

[0451] E88. The polynucleotide according to embodiment E87, wherein the nucleotide sequence encodes EVQLVESGGGLVQPGGSLRLSC (SEQ ID NO: 2223).

[0452] E89. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $X^1X^2QLX^3QX^4GX^5X^6X^7X^8X^9X^{10}GX^{11}X^{12}X^{13}X^{14}X^{15}SC$ (SEQ ID NO: 2255)

wherein

X^1 is selected from Q and E;

X^2 is selected from V and I;

X^3 is selected from V and Q;

X^4 is selected from S and P;

X^5 is selected from A, S, V, P, T, and G;

X^6 is selected from E, G and V;

X^7 is selected from V and L;

X^8 is selected from K, V, E, and A;

X^9 is selected from K, R and Q;

X^{10} is selected from P and S;

X^{11} is selected from A, E, S, T, and R;

X^{12} is selected from S and T;

X^{13} is selected from V and L;

X^{14} is selected from K and R; and,

X^{15} is selected from V, I, L, and M.

[0453] E90. The polynucleotide according to embodiment E89, wherein the nucleotide sequence encodes a sequence identical to QVQLVQSGAEVKKPGASVKVSC (SEQ ID NO: 2227) except for at least one substitution selected from Q1E, V2I, V5Q, S7P, A9(SVPTG), E10(GV), V11L, K12(VEA), K13(RQ), P14S, A16(ESTR), S17T, V18L, K19R, and V20(ILM).

[0454] E91. The polynucleotide according to embodiment E90, wherein the nucleotide sequence encodes QVQLVQSGAEVKKPGASVKVSC (SEQ ID NO: 2227).

[0455] E92. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $QX^1X^2LX^3X^4X^5GX^6X^7LX^8X^9PX^{10}X^{11}TLX^{12}LTC$ (SEQ ID NO: 2259)

wherein

X^1 is selected from V and L;

X^2 is selected from Q and T;

X^3 is selected from Q and R;

X^4 is selected from E and Q;

X^5 is selected from S and W;

X^6 is selected from P and A;

X^7 is selected from G and A;

X^8 is selected from V and L;

X^9 is selected from K and R;

X^{10} is selected from S and T;

X^{11} is selected from Q and E; and,

X^{12} is selected from S and T.

[0456] E93. The polynucleotide according to embodiment E92, wherein the nucleotide sequence encodes a sequence identical to QVQLQESGPGLVKPSQTLSLTC (SEQ ID NO: 2231). except for at least one substitution selected from V2L, Q3T, Q5R, E6Q, S7W, P9A, G10A, V12L, K13R, S15T, Q16E, and S19T.

[0457] E94. The polynucleotide according to embodiment E92, wherein the nucleotide sequence encodes QVQLQESGPGLVKPSQTLSLTC (SEQ ID NO: 2231).

[0458] E95. The polynucleotide according to any one of embodiments E86 to E94, wherein the nucleotide sequence encodes the first framework region (FW1) of a heavy chain variable domain.

[0459] E96. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes

$WX^1RQX^2PX^3KX^4LX^5X^6X^7X^8$ (SEQ ID NO: 2252)

wherein

X^1 is selected from V, I, and F;

X^2 is selected from A, S and T;

X^3 is selected from G and E;

X^4 is selected from G and R;

X^5 is selected from E and D;

X^6 is selected from W and L;

X^7 is selected from V and I; and,

X^8 is selected from A, S, and G.

[0460] E97. The polynucleotide according to embodiment E96, wherein the nucleotide sequence encodes a sequence identical to WVRQAPGKGLEWVA (SEQ ID NO: 2224) except for at least one substitution selected from V2(IF), A5(ST), G7E, G9R, E11D, W12L, V13I, and A14(SG).

[0461] E98. The polynucleotide according to embodiment E97, wherein the nucleotide sequence encodes WVRQAPGKGLEWVA (SEQ ID NO: 2224).

[0462] E99. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $WX^1X^2QX^3X^4GX^5X^6LX^7WX^8G$ (SEQ ID NO: 2256)

wherein

X^1 is selected from V and I;

X^2 is selected from R and K;

X^3 is selected from A, M, N, R, K, T, and S;

X^4 is selected from P, T, and H;

X^5 is selected from Q, K, and R;

X^6 is selected from G, R and S;

X^7 is selected from E, D, K, Q, and A; and,

X^8 is selected from M, I, and V.

[0463] E100. The polynucleotide according to embodiment E99, wherein the nucleotide sequence encodes a sequence identical to WVRQAPGQGLEWMG (SEQ ID NO: 2228) except for at least one substitution selected from V2I, R3K, A5(MNRKTS), P6(TH), Q8(KR), G9(RS), E11(DKQA), and M13(IV).

[0464] E101. The polynucleotide according to embodiment E100, wherein the nucleotide sequence encodes WVRQAPGQGLEWMG (SEQ ID NO: 2228).

[0465] E102. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or

any combination thereof), wherein the nucleotide sequence encodes $WX^1RX^2X^3X^4X^5X^6X^7LX^8WX^9X^{10}$ (SEQ ID NO: 2260)

wherein

X^1 is selected from I and V;

X^2 is selected from Q and H;

X^3 is selected from L, P, S, and H;

X^4 is selected from P and S;

X^5 is selected from G and E;

X^6 is selected from K and R;

X^7 is selected from G and A;

X^8 is selected from E and Q

X^9 is selected from I and L; and,

X^{10} is selected from G and A.

[0466] E103. The polynucleotide according to embodiment E102, wherein the nucleotide sequence encodes a sequence identical to WIRQLPGKGLEWIG (SEQ ID NO: 2232) except for at least one substitution selected from I2V, Q4H, L5(PSH), P6S, G7E, K8R, G9A, E11Q, I13L, and G14A.

[0467] E104. The polynucleotide according to embodiment E103, wherein the nucleotide sequence encodes WIRQLPGKGLEWIG (SEQ ID NO: 2232).

[0468] E105. The polynucleotide according to any one of embodiments E96 to E104, wherein the nucleotide sequence encodes the second framework region (FW2) of a heavy chain variable domain.

[0469] E106. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $X^1X^2X^3X^4SX^5DX^6X^7X^8X^9X^{10}X^{11}X^{12}LX^{13}X^{14}X^{15}X^{16}LX^{17}X^{18}EDTX^{19}X^{20}X^{21}X^{22}C$ (SEQ ID NO: 2253)

wherein

X^1 is selected from R and K;

X^2 is selected from F and V;

X^3 is selected from T, I, and A;

X^4 is selected from L and I;

X^5 is selected from V, R, L, and A;

X^6 is selected from R, N, T, D, K, and S;

X^7 is selected from S, A and V;

X^8 is selected from K, R, and E;

X^9 is selected from N, S, R, H, and T;

X^{10} is selected from T and S;

X^{11} is selected from L, A, and F;

X^{12} is selected from Y and F;

X^{13} is selected from Q and E;

X^{14} is selected from M and V;

X^{15} is selected from N, D, and S;

X^{16} is selected from S, G, and I;

X^{17} is selected from R and K;

X^{18} is selected from A, S, D, V, and P;

X^{19} is selected from A and G;

X^{20} is selected from V, M, and L;

X^{21} is selected from Y and F; and,

X^{22} is selected from Y and F.

[0470] E107. The polynucleotide according to embodiment E106, wherein the nucleotide sequence encodes a sequence identical to RFTLSVDRSKNTLYLQMNSLRAEDTAVYYC (SEQ ID NO: 2225) except for at least one substitution selected from R1K, F2V, T3(IA), L4I, V6(RLA), R8(NTDKS), S9(AV), K10(RE), N11(SRHT), T12S, L13(AF), Y14F, Q16E, M17V, N18(DS), S19(GI), R21K, A22(SDVP), A26G, V27(ML), Y28F, and Y29F.

[0471] E108. The polynucleotide according to embodiment E107, wherein the nucleotide sequence encodes RFTLSVDRSKNTLYLQMNSLRAEDTAVYYC (SEQ ID NO: 2225).

[0472] E109. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $X^1X^2X^3X^4X^5X^6X^7X^8SX^9X^{10}TX^{11}X^{12}X^{13}X^{14}X^{15}X^{16}X^{17}LX^{18}X^{19}X^{20}DX^{21}X^{22}X^{23}YX^{24}C$ (SEQ ID NO: 2257)

wherein

X^1 is selected from R, Q, and K;

X^2 is selected from V, I, F, G, and A;

X^3 is selected from T, A, and K;

X^4 is selected from M, I, L, and F;

X^5 is selected from T and S;

X^6 is selected from T, A, R, V, S, E, and L;

X^7 is selected from D, E, and N;

X^8 is selected from T, K, Q, S, P, R, I, N, and E;

X^9 is selected from T, K, S, A, I, and V;

X^{10} is selected from S, N, D, and T;

X^{11} is selected from A, V, and T;

X^{12} is selected from Y, and F;

X^{13} is selected from M and L;

X^{14} is selected from E, Q, and D;

X^{15} is selected from L, M, W, and I;

X^{16} is selected from R, S, D, L, K, T, and N;

X^{17} is selected from S and R;

X^{18} is selected from R, K, Q, and T;

X^{19} is selected from S, H, F, A, and P;

X^{20} is selected from D, E, and S;

X^{21} is selected from T and S;

X^{22} is selected from A and G;

X^{23} is selected from V, F, T, and M; and,

X^{24} is selected from Y, F, and L.

[0473] E110. The polynucleotide according to embodiment E109, wherein the nucleotide sequence encodes a sequence identical to RVTMTTDTSTSTAYMELRSLRSDDTAVYYC (SEQ ID NO: 2229) except for at least one substitution selected from R1(QK), V2(IFGA), T3(AK), M4(ILF), T5S, T6(ARVSEL), D7(EN), T8(KQSPRINE), T10(KSAIV), S11(NDT), A13(VT), Y14F, M15L, E16(QD), L17(MWI), R18(SDLKTN), S19R, R21(KQT), S22(HFAP), D23(ES), T25S, A26G, V27(FTM), and Y29(FL).

[0474] E111. The polynucleotide according to embodiment E110, wherein the nucleotide sequence encodes RVTMTTDTSTSTAYMELRSLRSDDTAVYYC (SEQ ID NO: 2229).

[0475] E112. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $RX^1X^2X^3X^4X^5DX^6SX^7X^8QX^9X^{10}LX^{11}X^{12}X^{13}X^{14}X^{15}X^{16}X^{17}X^{18}DTAX^{19}X^{20}X^{21}C$ (SEQ ID NO: 2261)

wherein

X^1 is selected from V and L;

X^2 is selected from T and S;

X^3 is selected from I and M;

X^4 is selected from S and L;

X^5 is selected from V, R, and K;

X^6 is selected from T and K;

X^7 is selected from K and R;

X^8 is selected from K and N;

X^9 is selected from F and V;

X^{10} is selected from S and V;

X^{11} is selected from R, T, K, and M;

X^{12} is selected from L, I, M, and V;

X^{13} is selected from S, T, and N;

X^{14} is selected from S and N;

X^{15} is selected from V and M;

X^{16} is selected from T and D;

X^{17} is selected from A and P;

X^{18} is selected from A and V;

X^{19} is selected from V and T;

X^{20} is selected from Y and W; and,

X^{21} is selected from Y, F and W.

[0476] E113. The polynucleotide according to embodiment E112, wherein the nucleotide sequence encodes a sequence identical to

RVTISVDTSKKQFSLRLSSVTAADTAVYYC (SEQ ID NO: 2233). except for at least one substitution selected from V2L, T3S, I4M, S5L, V6(RK), T8K, K10R, K11N, F13V, S14V, R16(TKM), L17(IMV), S18(TN), S19N, V20M, T21D, A22P, A23V, V27T, Y28W, and Y29(FW).

[0477] E114. The polynucleotide according to embodiment E113, wherein the nucleotide sequence encodes RVTISVDTSKKQFSLRLSSVTAADTAVYYC (SEQ ID NO: 2233).

[0478] E115. The polynucleotide according to any one of embodiments E106 to E114, wherein the nucleotide sequence encodes the third framework region (FW3) of a heavy chain variable domain.

[0479] E116. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $WGX^1GX^2X^3VTVS$ (SEQ ID NO: 2254)

wherein

X^1 is selected from Q, R, and K;

X^2 is selected from T, I and A; and,

X^3 is selected from L, S, T, M, and P.

[0480] E117. The polynucleotide according to embodiment E116, wherein the nucleotide sequence encodes a sequence identical to WGQGTLVTVS (SEQ ID NO: 2226) except for at least one substitution selected from Q3(RK), T5(IA), and L6(STMP).

[0481] E118. The polynucleotide according to embodiment E117, wherein the nucleotide sequence encodes WGQGTLVTVS (SEQ ID NO: 2226).

[0482] E119. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $WGX^1GTX^2X^3TVS$ (SEQ ID NO: 2258)

wherein

X^1 is selected from R, Q, K, A and S;

X^2 is selected from L, M, T, Q, and P; and,

X^3 is selected from V and L.

[0483] E120. The polynucleotide according to embodiment E119, wherein the nucleotide sequence encodes a sequence identical to WGRGTLVTVS (SEQ ID NO: 2230) except for at least one substitution selected from R3(QKAS), L6(MTQP), and V7L.

[0484] E121. The polynucleotide according to embodiment E120, wherein the nucleotide sequence encodes WGRGTLVTVS (SEQ ID NO: 2230).

[0485] E122. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes $WX^1X^2GX^3X^4VTVS$ (SEQ ID NO: 2262)

wherein

X^1 is selected from G and D;

X^2 is selected from Q and R;

X^3 is selected from T and S; and,

X^4 is selected from T, L, and M.

[0486] E123. The polynucleotide according to embodiment E122, wherein the nucleotide sequence encodes a sequence identical to WGQGTTVTVS (SEQ ID NO: 2234), except for at least one substitution selected from G2D, Q3R, T5S, and T6(LM).

[0487] E124. The polynucleotide according to embodiment E123, wherein the nucleotide sequence encodes WGQGTTVTVS (SEQ ID NO: 2234).

[0488] E125. The polynucleotide according to any one of embodiments E116 to E124, wherein the nucleotide sequence encodes the fourth framework region (FW4) of a heavy chain variable domain.

[0489] E126. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes a sequence of formula $(Gly_xSer)_y$, wherein x and y are integers between E1 and E100.

[0490] E127. The polynucleotide of any one of embodiments E1 to E125, further comprising a nucleotide sequence which encodes a sequence of formula $(Gly_xSer)_y$, wherein x and y are integers between E1 and E100.

[0491] E128. The polynucleotide according to embodiment E126 or E127, wherein the sequence of formula $(Gly_xSer)_y$ is a linker.

[0492] E129. The polynucleotide according to embodiment E128, wherein the linker comprises the sequence (Gly_4Ser), (Gly_3Ser), (Gly_2Ser), or a combination thereof.

[0493] E130. The polynucleotide according to embodiment E128, wherein the linker comprises the sequence $(Gly_4Ser)_3$.

[0494] E131. The polynucleotide according to embodiment E128, wherein the linker is interposed between a VH domain and a VL domain.

[0495] E132. The polynucleotide according to any one of embodiment E128, which encodes an scFv.

[0496] E133. A polynucleotide encoding an antibody or an antigen binding portion thereof comprising the polynucleotide of any one of embodiments E30 – E33 and E46 – E55, the polynucleotide of any one of embodiments E34 – E37 and E56 – E65, the polynucleotide of any one of embodiments E38 – E41 and E66 – E75, the polynucleotide of any one of embodiments E42 – E45 and E76 – E85, or any combination thereof.

[0497] E134. The polynucleotide according to embodiment E133, comprising the polynucleotide of any one of embodiments E30 – E33 and E46 – E55, the polynucleotide of any one of embodiments E34 – E37 and E56 – E65, the polynucleotide of any one of embodiments E38 – E41 and E66 – E75, and the polynucleotide of any one of embodiments E42 – E45 and E76 – E85.

[0498] E135. A polynucleotide encoding an antibody or an antigen binding portion thereof comprising the polynucleotide of any one of embodiments E86 – E95, the polynucleotide of any one of embodiments E96 – E105, the polynucleotide of any one of embodiments E106 – E115, the polynucleotide of any one of embodiments E116 – E125, or any combination thereof.

[0499] E136. The polynucleotide according to embodiment E135, comprising the polynucleotide of any one of embodiments E86 – E95, the polynucleotide of any one of embodiments E96 – E105, the polynucleotide of any one of embodiments E106 – E115, and the polynucleotide of any one of embodiments E116 – E125.

[0500] E137. The polynucleotide according to embodiment E135, further comprising the polynucleotide of embodiment E133 or E134.

[0501] E138. The polynucleotide according to any one of embodiments E133 to E137, further comprising the polynucleotide according to any one of embodiments E1 to E30.

[0502] E139. A polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%,

97%, 98% or 99% identical to (i) any one of the polynucleotides of SEQ ID NOS:1-88, or (ii) a subsequence of any one of the polynucleotides of SEQ ID NOS: 89-1978 encoding an Ig constant domain, wherein the nucleotide subsequence encodes an Immunoglobulin (Ig) polypeptide that has a significant match to a corresponding sequence of CDD domain CD00098 (FIG. 1).

[0503] E140. The polynucleotide of embodiment E139, wherein the Ig polypeptide comprises an Ig constant domain of an antibody or a fragment thereof.

[0504] E141. The polynucleotide according to embodiment E140, wherein the Ig constant domain is an CL, CH1, CH2, or CH3 constant domain from an IgG.

[0505] E142. A polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to (i) any one of the polynucleotides of SEQ ID NOS:1-8, or 45-52, or (ii) a subsequence of any one of the polynucleotides of SEQ ID NOS:1034-1978 encoding an Ig light chain constant domain (CL), wherein the nucleotide subsequence encodes an Ig polypeptide that has a significant match to a corresponding sequence of CDD domain CD07699 (FIG. 2).

[0506] E143. The polynucleotide of embodiment E142, wherein the Ig polypeptide comprises a light chain constant region of an antibody or a fragment thereof.

[0507] E144. A polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to (i) any one of the polynucleotides of SEQ ID NOS:9-12, 21-24, 33-36, 53-56, 65-68, or 77-80, or (ii) a subsequence of any one of the polynucleotides of SEQ ID NOS: 89-1033 encoding a CH1 constant domain, wherein the nucleotide subsequence encodes an Ig polypeptide that has a significant match to a corresponding sequence of CDD domain CD04985 (FIG. 3).

[0508] E145. The polynucleotide of embodiment E144, wherein the Ig polypeptide comprises a heavy chain CH1 constant domain of an antibody or a fragment thereof.

[0509] E146. A polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to (i) any one of the polynucleotides of SEQ ID NO:13-16, 25-28, 37-40, 57-60, 69-72, or 81-84, or (ii) a subsequence of any one of the polynucleotides of SEQ ID NOS: 89-1033 encoding a CH2 constant domain, wherein the

nucleotide subsequence encodes an Ig polypeptide that has a significant match to a corresponding sequence of CDD domain CD04986 (FIG. 4).

[0510] E147. The polynucleotide of embodiment E146, wherein the Ig polypeptide comprises a heavy chain CH2 constant domain of an antibody or a fragment thereof.

[0511] E148. A polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to (i) any one of the polynucleotides of SEQ ID NO:17-20, 29-32, 41-44, 61-64, 73-76, or 85-88, or (ii) a subsequence of any one of the polynucleotides of SEQ ID NOS: 89-1033 encoding a CH3 constant region, wherein the nucleotide subsequence encodes an Ig polypeptide that has a significant match to a corresponding sequence of CDD domain CD07696 (FIG. 5).

[0512] E149. The polynucleotide of embodiment E148, wherein the Ig polypeptide comprises a heavy chain CH3 constant domain of an antibody or a fragment thereof.

[0513] E150. A polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to a subsequence of any one of the polynucleotides of SEQ ID NOS: 89-1978 encoding a variable domain, wherein the nucleotide subsequence encodes an Ig polypeptide that has a significant match to a corresponding sequence of CDD domain CD00099 (FIG. 6).

[0514] E151. The polynucleotide of embodiment E150, wherein the Ig polypeptide comprises a variable domain of an antibody or a fragment thereof.

[0515] E152. A polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to a subsequence of any one of the polynucleotides of SEQ ID NOS: 89-1033 encoding a VH domain, wherein the nucleotide subsequence encodes an Ig polypeptide that has a significant match to a corresponding sequence of CDD domain CD04981 (FIG. 7).

[0516] E153. The polynucleotide of embodiment E152, wherein the Ig polypeptide comprises a VH domain of an antibody or a fragment thereof.

[0517] E154. A polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical a subsequence of any one of the polynucleotides of SEQ ID NOS: 1034-1978 encoding a VL domain, wherein the nucleotide subsequence encodes an

Ig polypeptide that has a significant match to a corresponding sequence of CDD domain CD04980 (FIG. 8) or CD04984 (FIG. 9).

[0518] E155. The polynucleotide of embodiment E154, wherein the Ig polypeptide comprises a VL kappa domain or a VL lambda domain of an antibody or a fragment thereof.

[0519] E156. A polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to any one of the polynucleotides of SEQ ID NOS:89-1033, wherein the nucleotide sequence encodes an Ig polypeptide that has non-overlapping significant matches to CDD domains CD04981/CD4984, CD04985, and CD04986.

[0520] E157. The polynucleotide of embodiment E156, wherein the Ig polypeptide comprises the heavy chain of an antibody or a fragment thereof.

[0521] E158. A polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to any one of the polynucleotides of SEQ ID NO:1034-1978, wherein the nucleotide sequence encodes an Ig polypeptide that has non-overlapping significant matches to CD04980 and CD07699.

[0522] E159. The polynucleotide of embodiment E158, wherein the Ig polypeptide comprises the light chain of an antibody or a fragment thereof.

[0523] E160. A polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to any one of the polynucleotides of SEQ ID NOS: 1-4, or 45-48, wherein the nucleotide sequence encodes a CL kappa domain or a functional fragment thereof from a therapeutic antibody.

[0524] E161. The polynucleotide according to embodiment E160, wherein the CL kappa domain comprises

TVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQES VTEQDSKDSTYSLSX[1]TLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNR (SEQ ID NO: 2200),

wherein X^1 is selected from N and S.

[0525] E162. A polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to any one of the polynucleotides of SEQ ID NOS: 5-8, or

49-52, wherein the nucleotide sequence encodes a CL lambda domain or a functional fragment thereof from a therapeutic antibody.

[0526] E163. The polynucleotide according to embodiment E162, wherein the CL lambda domain comprises PKAAPSVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTT PSKQSNNKYAASSYLSLTPEQWKSHX^{2}SYSCQVTHEGSTVEKTVAPX^{3}ECS (SEQ ID NO: 2201),

wherein X^2 is selected from R and K, and X^3 is selected from T and A.

[0527] E164. A polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to any one of the polynucleotides of SEQ ID NOS: 9-12, 21-24, 33-36, 53-56, 65-68, or 77-80, wherein the nucleotide sequence encodes a CH1 domain or a functional fragment thereof from a therapeutic antibody.

[0528] E165. The polynucleotide according to embodiment E3, wherein the nucleotide sequence is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO: 9-12, or 53-56 and the CH1 domain is an IgG1 CH1 domain.

[0529] E166. The polynucleotide according to embodiment E165, wherein the IgG1 CH1 domain comprises SX^{4}GPSVX^{5}PLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGVHTFPAVLQSSG LYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKX^{6}X^7 (SEQ ID NO: 2202) wherein X^4 is an optional ASTK sequence, X^5 is selected from F and L, X^6 is selected from K and R, and X^7 is selected from V and A.

[0530] E167. The polynucleotide according to embodiment E164, wherein the nucleotide sequence is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO: 21-24, or 65-68 and the CH1 domain is an IgG2 CH1 domain.

[0531] E168. The polynucleotide according to embodiment E167, wherein the IgG2 CH1 domain comprises SASTKGPSVFPLAPCSRSTSESTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPA VLQSSGLYSLSSVVTVX^{15}SSNFGTQTYTCNVDHKPSNTKVDKTV (SEQ ID NO: 2205)

wherein X^{15} is selected from P and T.

[0532] E169. The polynucleotide according to embodiment E164, wherein the nucleotide sequence is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO: 33-36, or 77-80 and the CH1 domain is an IgG4 CH1 domain.

[0533] E170. The polynucleotide according to embodiment E169, wherein the IgG4 CH1 domain comprises SASTKGPSVFPLAPCSRSTSESTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPA VLQSSGLYSLSSVVTVPSSSLGTKTYTCNVDHKPSNTKVDKRV (SEQ ID NO: 2197).

[0534] E171. A polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to any one of the polynucleotides of SEQ ID NOS: 13-16, 25-28, 37-40, 57-60, 69-72, or 81-84, wherein the nucleotide sequence encodes a CH2 domain or a functional fragment thereof from a therapeutic antibody.

[0535] E172. The polynucleotide according to embodiment E171, wherein the nucleotide sequence is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO: 13-16, or 57-60 and the CH2 domain is an IgG1 CH2 domain.

[0536] E173. The polynucleotide according to embodiment E172, wherein the IgG1 CH2 domain comprises APEX^{8}X^{9}GX^{10}PSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNX^{11}YVDGV EVHNAKTKPREEQYX^{12}STYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK TISKAK (SEQ ID NO: 2203)

wherein X^{8} is selected from L and A, X^{9} is selected from L and A, X^{10} is selected from G and A, and X^{11} is selected from V and W, and X^{12} is selected from N and A.

[0537] E174. The polynucleotide according to embodiment E171, wherein the nucleotide sequence is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO: 25-28, or 69-72 and the CH2 domain is an IgG2 CH2 domain.

[0538] E175. The polynucleotide according to embodiment E174, wherein the IgG2 CH2 domain comprises APPVAGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFNWYVDGX^{16}EV

HNAKTKPREEQFNSTFRVVSVLTVVHQDWLNGKEYKCKVSNKGLPX[17]X[18]IEKTI SKTK (SEQ ID NO: 2206)

wherein X^{16} is selected from V and M, X^{17} is selected from A and S; and X^{18} is selected from P and S.

[0539] E176. The polynucleotide according to embodiment E171, wherein the nucleotide sequence is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO: 37-40, or 81-84 and the CH2 domain is an IgG4 CH2 domain.

[0540] E177. The polynucleotide according to embodiment E176, wherein the IgG4 CH2 domain comprises APEFLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSQEDPEVQFNWYVDGVEVH NAKTKPREEQFNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKGLPSSIEKTISKA K (SEQ ID NO: 2198).

[0541] E178. A polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to any one of the polynucleotides of SEQ ID NOS: 17-20, 29-32, 41-44, 61-64, 73-76, or 85-88, wherein the nucleotide sequence encodes a CH3 domain or a functional fragment thereof from a therapeutic antibody.

[0542] E179. The polynucleotide according to embodiment E178, wherein the nucleotide sequence is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO: 17-20, or 61-64 and the CH3 domain is an IgG1 CH3 domain.

[0543] E180. The polynucleotide according to embodiment E178, wherein the IgG1 CH3 domain comprises GQPREPQVYTLPPSRX[13]EX[14]TKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYK TTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG (SEQ ID NO: 2204)

wherein X^{13} is selected from E and D, and X^{14} is selected from M and L.

[0544] E181. The polynucleotide according to embodiment E178, wherein the nucleotide sequence is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO: 29-32, or 73-76 and the CH3 domain is an IgG2 CH3 domain.

[0545] E182. The polynucleotide according to embodiment E181, wherein the IgG2 CH3 domain comprises GQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTP PMLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG (SEQ ID NO: 2196).

[0546] E183. The polynucleotide according to embodiment E178, wherein the nucleotide sequence is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO: 41-44, or 85-88 and the CH3 domain is an IgG4 CH3 domain.

[0547] E184. The polynucleotide according to embodiment E183, wherein the IgG4 CH3 domain comprises GQPREPQVYTLPPSQEEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTP PVLDSDGSFFLYSRLTVDKSRWQEGNVFSCSVMHEALHNHYTQKSLSLSLG (SEQ ID NO: 2199).

[0548] E185. A polynucleotide comprising a nucleotide sequence at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identical to a subsequence from any one of the polynucleotides of SEQ ID NOS: 89-1978, , wherein said subsequence encodes

(a) one, two, or three VH-CDRs from a therapeutic antibody;

(b) one, two, or three VL-CDRs from a therapeutic antibody;

(c) one, two, three, or four VH framework (FW) regions from a therapeutic antibody;

(d) one, two, three, or four VL framework (FW) regions from a therapeutic antibody;

(e) a VH domain from a therapeutic antibody;

(f) a VL domain from a therapeutic antibody;

(g) a CL domain of a therapeutic antibody;

(h) a CH1 domain of a therapeutic antibody;

(i) a CH2 domain of a therapeutic antibody;

(j) a CH3 domain of a therapeutic antibody; or,

(k) a combination thereof.

[0549] E186. The polynucleotide according to embodiment E185, wherein the subsequence encoding one, two, three, or four VH framework (FW) regions from a

therapeutic antibody comprises the polynucleotide of any one of embodiments E86 – E95, embodiments E96 – E105, embodiments E106 – E115, embodiments E116 – E125, or any combinations thereof.

[0550] E187. The polynucleotide according to embodiment E185, wherein the subsequence encoding one, two, three, or four VL framework (FW) regions from a therapeutic antibody comprises the polynucleotide of any one of embodiments E30 – E33 and E46 – E55, embodiments E34 – E37 and E56 – E65, embodiments E38 – E41 and E66-E75, embodiments E42-E45 and E76-E85, or any combinations thereof.

[0551] E188. The polynucleotide according to embodiment E185, wherein the subsequence encoding a CL domain of a therapeutic antibody comprises the polynucleotide of any one of embodiments E1 to E6.

[0552] E189. The polynucleotide according to embodiment E185, wherein the subsequence encoding a CH1 domain of a therapeutic antibody comprises the polynucleotide of any one of embodiments E7 to E9, E16-E18, and E24 – E25.

[0553] E190. The polynucleotide according to embodiment E185, wherein the subsequence encoding a CH2 domain of a therapeutic antibody comprises the polynucleotide of any one of embodiments E10 – E12, E19 – E21, and E26 – E27.

[0554] E191. The polynucleotide according to embodiment E185, wherein the subsequence encoding a CH3 domain of a therapeutic antibody comprises the polynucleotide of any one of embodiments E13 – E15, E22 – E23, and E28 – E29.

[0555] E192. The polynucleotide according to embodiment E185, further comprising a nucleotide sequence encoding a linker sequence.

[0556] E193. The polynucleotide according to embodiment E185, which encodes an scFv.

[0557] E194. The polynucleotide according to any one of embodiments E160 to E193, wherein the therapeutic antibody is selected from the group consisting of abagovomab, abciximab, adalimumab, alemtuzumab, alirocumab, amatuximab, anrukinzumab, arcitumomab, basiliximab, bavituximab, benralizumab, bevacizumab, bezlotoxumab, bimagrumab, bococizumab, brentuximab, briakinumab, brodalumab, canakinumab, cantuzumab, carlumab, cetuximab, cixutumumab, clivatuzumab, conatumumab, crenezumab, dacetuzumab, daclizumab, dalotuzumab, denosumab, drozitumab, dupilumab, dusigitumab, eculizumab, elotuzumab, enokizumab, epratuzumab, etaracizumab, evolocumab, farletuzumab, fasinumab, fezakinumab, ficlatuzumab,

figitumumab, fresolimumab, fulranumab, ganitumab, gantenerumab, gevokizumab, girentuximab, glembatumumab, ibalizumab, ibritumomab, icrucumab, inotuzumab, intetumumab, itolizumab, ixekizumab, lebrikizumab, lorvotuzumab, mavrilimumab, mepolizumab, milatuzumab, mogamulizumab, motavizumab, naptumomab, necitumumab, nivolumab, obinutuzumab, ocrelizumab, olaratumab, omalizumab, otelixizumab, oxelumab, pateclizumab, pembrolizumab, pertuzumab, ponezumab, ramucirumab, rilotumumab, rituximab, robatumumab, romosozumab, rontalizumab, samalizumab, sarilumab, secukinumab, sifalimumab, siltuximab, sirukumab, solanezumab, tabalumab, tanezumab, tenatumomab, teplizumab, tigatuzumab, tildrakizumab, tocilizumab, tositumomab, tralokinumab, trastuzumab, urelumab, ustekinumab, vedolizumab, and veltuzumab, and functional fragments thereof.

[0558] E195. The polynucleotide according to embodiment E194, wherein the functional fragment is an antigen-binding fragment.

[0559] E196. The polynucleotide according to embodiment E195, wherein the functional fragment is a non-antigen-binding fragment.

[0560] E197. The polynucleotide according to embodiment E196, wherein the non-antigen-binding fragment is an Fc fragment.

[0561] E198. A polynucleotide comprising a nucleotide sequence encoding an antibody or a fragment thereof, wherein Ala is encoded by GCC, GCG or GCT; Cys is encoded by TGC or TGT; Asp is encoded by GAC; Glu is encoded by GAG or GAA; Phe is encoded by TTC; Gly is encoded by GGC, GGT, or GGG; His is encoded by CAC; Ile is encoded by ATC or ATT; Lys is encoded by AAG; Leu is encoded by CTG, CTC or TTG; Met is encoded by ATG; Asn is encoded by AAC; Pro is encoded by CCC, CCA or CCG; Gln is encoded by CAG or CAA, Arg is encoded by CGG, AGG, CGC, CGT, AGA, CGA, Ser is encoded by AGC, TCC or TCT, Thr is encoded by ACC, ACG or ACT, Val is encoded by GTG, GTC or GTT, Trp is encoded by TGG, and Tyr is encoded by TAC, wherein the nucleotide sequence encodes abagovomab, abciximab, adalimumab, alemtuzumab, alirocumab, amatuximab, anrukinzumab, arcitumomab, basiliximab, bavituximab, benralizumab, bevacizumab, bezlotoxumab, bimagrumab, bococizumab, brentuximab, briakinumab, brodalumab, canakinumab, cantuzumab, carlumab, cetuximab, cixutumumab, clivatuzumab, conatumumab, crenezumab, dacetuzumab, daclizumab, dalotuzumab, denosumab, drozitumab, dupilumab, dusigitumab, eculizumab, elotuzumab, enokizumab, epratuzumab, etaracizumab, evolocumab, farletuzumab, fasinumab,

fezakinumab, ficlatuzumab, figitumumab, fresolimumab, fulranumab, ganitumab, gantenerumab, gevokizumab, girentuximab, glembatumumab, ibalizumab, ibritumomab, icrucumab, inotuzumab, intetumumab, itolizumab, ixekizumab, lebrikizumab, lorvotuzumab, mavrilimumab, mepolizumab, milatuzumab, mogamulizumab, motavizumab, naptumomab, necitumumab, nivolumab, obinutuzumab, ocrelizumab, olaratumab, omalizumab, otelixizumab, oxelumab, pateclizumab, pembrolizumab, pertuzumab, ponezumab, ramucirumab, rilotumumab, rituximab, robatumumab, romosozumab, rontalizumab, samalizumab, sarilumab, secukinumab, sifalimumab, siltuximab, sirukumab, solanezumab, tabalumab, tanezumab, tenatumomab, teplizumab, tigatuzumab, tildrakizumab, tocilizumab, tositumomab, tralokinumab, trastuzumab, urelumab, ustekinumab, vedolizumab, veltuzumab, or an antigen binding fragment thereof.

[0562] E199. The polynucleotide according to embodiment E198, wherein the nucleotide sequence is selected from SEQ ID NOS: 89-1978, and subsequences thereof encoding antigen binding fragments.

[0563] E200. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein the nucleotide sequence encodes a fragment of

(i) the sequences of SEQ ID NO: 1979-2006; or,

(ii) a polypeptide sequence encoded by the nucleotide of any one of embodiments E1 to E199, and

wherein the fragment is about 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 205, 210, 215, 220, 225, 230, 235, 240, 245, 250, 255, 260, 265, 270, 275, 280, 285, 290, 295, 300, 305, 310, 315, 320, 325, 330, 335, 340, 345, 350, 355, 360, 365, 370, 375, 380, 385, 390, 395, 400, 405, 410, 415, 420, 425, 430, 435, 440, 445, 450, 455, 460, 465, 470, 475, 480, 485, 490, 495, 500, 505, 510, 515, 520, 525, 530, 535, 540, 545, or 550 amino acids long.

[0564] E201. The polynucleotide of any one of embodiments E1 to E200, wherein the nucleotide sequence is not a wild type sequence.

[0565] E202. The polynucleotide of any one of embodiments E1 to E201, wherein the nucleotide sequence has been optimized according to a multiparametric method comprising:

(i) modifying at least one subsequence in a candidate nucleic acid sequence to generate a ramp subsequence;

(ii) substituting at least one codon in a candidate nucleic acid sequence with an alternative codon to increase or decrease uridine content to generate a uridine-modified sequence;

(iii) substituting at least one codon in a candidate nucleic acid sequence or the uridine-modified sequence with a fast recharging codon;

(iv) substituting at least one codon in a candidate nucleic acid sequence with an alternative codon having a higher codon frequency in the synonymous codon set;

(v) substituting at least one natural nucleobase in a candidate nucleic acid sequence with an alternative synthetic nucleobase;

(vi) substituting at least one internucleoside linkage in a candidate nucleic acid sequence with a non-natural internucleoside linkage; or,

(vii) a combination thereof.

wherein the resulting optimized nucleic acid sequence has at least one optimized property with respect to the candidate nucleic acid sequence.

[0566] E203. The polynucleotide of embodiment E202, wherein the multiparametric method comprises one, two, three, four, five or six optimization methods selected from the group consisting of (i) modifying at least one subsequence in a candidate nucleic acid sequence to generate a ramp subsequence; (ii) substituting at least one codon in a candidate nucleic acid sequence with an alternative codon to increase or decrease uridine content to generate a uridine-modified sequence; (iii) substituting at least one codon in a candidate nucleic acid sequence or the uridine-modified sequence with a fast recharging codon; (iv) substituting at least one codon in a candidate nucleic acid sequence with an alternative codon having a higher codon frequency in the synonymous codon set; (v) substituting at least one natural nucleobase in a candidate nucleic acid sequence with an alternative synthetic nucleobase; and (vi) substituting at least one internucleoside linkage in a candidate nucleic acid sequence with a non-natural internucleoside linkage.

[0567] E204. The polynucleotide of embodiment E202 or E203, wherein the multiparametric method comprises replacing at least 5%, at least 10%, at least 15%, at

least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% of the codons in the candidate nucleic acid sequence.

[0568] E205. The polynucleotide according to any one of embodiments E201 to E204, wherein the candidate nucleic acid sequence is SEQ ID NOS: 1979-2188, or a fragment thereof.

[0569] E206. The polynucleotide according to embodiment E205, wherein the fragment comprises

(a) one, two, or three VH-CDRs from SEQ ID NOS: 1979-2083;

(b) one, two, or three VL-CDRs from SEQ ID NOS: 2084-2188;

(c) one, two, three, or four VH framework (FW) regions from SEQ ID NOS: 1979-2083;

(d) one, two, three, or four VL framework (FW) regions from SEQ ID NOS: 2084-2188;

(e) a VH domain from SEQ ID NOS: 1979-2083;

(f) a VL domain from SEQ ID NOS: 2084-2188;

(g) a CL domain from SEQ ID NOS: 2084-2188;

(h) a CH1 domain from SEQ ID NOS: 1979-2083;

(i) a CH2 domain from SEQ ID NOS: 1979-2083;

(j) a CH3 domain from SEQ ID NOS: 1979-2083; or,

(k) a combination thereof.

[0570] E207. The polynucleotide according to any one of embodiments E1 to E206, wherein the polynucleotide is a DNA.

[0571] E208. The polynucleotide according to any one of embodiments E1 to E206, wherein the polynucleotide is an RNA.

[0572] E209. The polynucleotide according to embodiment E208, wherein the RNA is mRNA.

[0573] E210. The polynucleotide according to embodiment E209, wherein the mRNA is synthetic.

[0574] E211. The polynucleotide according to any one of embodiments E1 to E210, wherein the polynucleotide comprises at least one nucleotide analogue.

[0575] E212. The polynucleotide according to embodiment E211, wherein the at least one nucleotide analogue is selected from the group consisting of a 2'-O-methoxyethyl-RNA (2'-MOE-RNA) monomer, a 2'-fluoro-DNA monomer, a 2'-O-alkyl-RNA monomer, a 2'-amino-DNA monomer, a locked nucleic acid (LNA) monomer, a cEt monomer, a cMOE monomer, a 5'-Me-LNA monomer, a 2'-(3-hydroxy)propyl-RNA monomer, an arabino nucleic acid (ANA) monomer, a 2'-fluoro-ANA monomer, an anhydrohexitol nucleic acid (HNA) monomer, an intercalating nucleic acid (INA) monomer, and a combination of two or more of said nucleotide analogues.

[0576] E213. The polynucleotide according to any one of embodiments E1 to E212, wherein said polynucleotide comprises at least one backbone modification.

[0577] E214. The polynucleotide according to embodiment E213, wherein the at least one backbone modification is a phosphorothioate internucleotide linkage.

[0578] E215. The polynucleotide according to embodiment E214, wherein all of the internucleotide linkages are phosphorothioate internucleotide linkages.

[0579] E216. The polynucleotide according to any one of embodiments E1 to E215, wherein

(i) at least one uridine has been replaced with 2-pseudouridine, 5-methoxyuridine, 2-thiouridine, 4-thiouridine, N1-methylpseudouridine, 5-aza-uridine, 2-thio-5-aza-uridine, 4-thio-pseudouridine, 2-thio-pseudouridine, 5-hydroxyuridine, 4-methoxy-pseudouridine, 4-methoxy-2-thio-pseudouridine, 3-methyluridine, 5-carboxymethyl-uridine, 1-carboxymethyl-pseudouridine, 5-propynyl-uridine, 1-propynyl-pseudouridine, 2-methoxy-4-thio-uridine, 5-taurinomethyluridine, 1-taurinomethyl-pseudouridine, 5-taurinomethyl-2-thio-uridine, 1-taurinomethyl-4-thio-uridine, 5-methyl-uridine, 2-methoxyuridine, 1-methyl-pseudouridine, 4-thio-1-methyl-pseudouridine, 2-thio-1-methyl-pseudouridine, 1-methyl-1-deaza-pseudouridine, 2-thio-1-methyl-1-deaza-pseudouridine, 1-ethyl-pseudouridine, or 2-thio-dihydrouridine; and/or,

(ii) at least one adenosine has been replaced with 2-aminopurine, 2,6-diaminopurine, 7-deaza-adenine, 7-deaza-8-aza-adenine, 7-deaza-2-aminopurine, 7-deaza-8-aza-2-aminopurine, 7-deaza-2,6-diaminopurine, 7-deaza-8-aza-2,6-diaminopurine, 1-methyladenosine, N6-methyladenosine, N6-isopentenyladenosine, N6-(cis-hydroxyisopentenyl)adenosine, 2-methylthio-N6-(cis-hydroxyisopentenyl)adenosine, N6-glycinylcarbamoyladenosine, N6-threonylcarbamoyladenosine, 2-methylthio-N6-threonyl carbamoyladenosine, N6,N6-dimethyladenosine, or 7-methyladenine; and/or,

(iii) at least one guanosine has been replaced with inosine, 1-methyl-inosine, wyosine, wybutosine, 7-deaza-guanosine, 7-deaza-8-aza-guanosine, 6-thio-guanosine, 6-thio-7-deaza-guanosine, 6-thio-7-deaza-8-aza-guanosine, 7-methyl-guanosine, 6-thio-7-methyl-guanosine, 7-methylinosine, 6-methoxy-guanosine, 1-methylguanosine, N2-methylguanosine, N2,N2-dimethylguanosine, 8-oxo-guanosine, 7-methyl-8-oxo-guanosine, or 1-methyl-6-thio-guanosine; and/or,

(iv) at least one cytidine has been replaced with 5-methylcytidine, 5-aza-cytidine, pseudoisocytidine, 3-methyl-cytidine, N4-acetylcytidine, 5-formylcytidine, N4-methylcytidine, 5-hydroxymethylcytidine, 1-methyl-pseudoisocytidine, pyrrolo-cytidine, pyrrolo-pseudoisocytidine, 2-thio-cytidine, 2-thio-5-methyl-cytidine, 4-thio-pseudoisocytidine, 4-thio-1-methyl-pseudoisocytidine, 4-thio-1-methyl-1-deaza-pseudoisocytidine, 1-methyl-1-deaza-pseudoisocytidine, zebularine, 5-aza-zebularine, 5-methyl-zebularine, 5-aza-2-thio-zebularine, 2-thio-zebularine, 2-methoxy-cytidine, 2-methoxy-5-methyl-cytidine, or 4-methoxy-pseudoisocytidine.

[0580] E217. The polynucleotide according to any one of embodiments 1 to 216, wherein

(i) 25% of uridines have been replaced with 4-thiouridine;

(ii) 50% of uridines have been replaced with 4-thiouridine;

(iii) 100% of uridines have been replaced with 4-thiouridine;

(iv) 25% of uridines have been replaced with 2-thiouridine (s2U) and 25% of cytidines have been replaced with 5-methylcytidine (m5C);

(v) 50% of uridines have been replaced with 2-thiouridine (s2U);

(vi) 100% of uridines have been replaced with pseudouridine (Ψ);

(vii) 100% of uridines have been replaced with pseudouridine (Ψ) and 100% of cytidines have been replaced with 5-methylcytidine (5mC);

(viii) 25% of uridines have been replaced with 5-methoxyuridine (5moU) and 50% of cytidines have been replaced with 5-methylcytidine (5mC);

(ix) 25% of uridines have been replaced with 5-methoxyuridine (5moU) and 100% of cytidines have been replaced with 5-methylcytidine (5mC);

(x) 100% of uridines have been replaced with 5-methoxyuridine (5moU);

(xi) 100% of uridines have been replaced with 5-methoxyuridine (5moU) and 100% of cytidines have been replaced with 5-methylcytidine (5mC);

(xii) 100% of uridines have been replaced with N1-methylpseudouridine (1mΨ);

(xiii) 100% of uridines have been replaced with 1-ethyl-pseudouridine; or,

(xiv) 100% of uridines have been replaced with N1-methylpseudouridine (1mΨ) and 100% of cytidines have been replaced with 5-methylcytidine (5mC).

[0581] E218. A vector or set of vectors comprising a polynucleotide according to any one of embodiments E1 to E217 or a complement thereof.

[0582] E219. A method for making a polynucleotide according to any one of embodiments E1 to E217 or a complement thereof comprising chemically synthesizing said polynucleotide.

[0583] E220. A method for producing a protein encoded a polynucleotide according to any one of embodiments E1 to E217, wherein the expression is conducted using an in vitro translation system.

[0584] E221. A cell comprising the polynucleotide according to any one of embodiments E1 to E217 or a complement thereof, or the vector or set of vectors according to embodiment E218.

[0585] E222. The cell according to embodiment E221, wherein the cell is an autologous cell or a heterologous cell.

[0586] E223. A pharmaceutical composition comprising (i) a polynucleotide according to any one of embodiments E1 to E217 or a complement thereof, (ii) a vector or set of vectors according to embodiment E218, or (iii) a cell according to embodiment E221 or embodiment E222, and a pharmaceutically acceptable vehicle or excipient.

[0587] E224. A method of expressing a polypeptide comprising contacting an effective amount of (i) a polynucleotide according to any one of embodiments E1 to E217 or a complement thereof or (ii) a vector or set of vectors according to embodiment E218 in a cell, wherein the polypeptide encoded by the polynucleotide is expressed.

[0588] E225. The method of embodiment E224, wherein the polypeptide is expressed in vitro.

[0589] E226. The method of embodiment E224, wherein the polypeptide is expressed in vivo.

[0590] E227. A method to treat a disease or condition in a subject in need thereof comprising administering a therapeutically effective amount of (i) a polynucleotide according to any one of embodiments E1 to E217 or a complement thereof, (ii) a vector or set of vectors according to embodiment E218, (iii) a cell according to embodiment

E221 or embodiment E222, (iv) a pharmaceutical composition according to embodiment E223, or (v) a combination thereof.

* * *

[0591] It is to be appreciated that the Detailed Description section, and not the Summary and Abstract sections, is intended to be used to interpret the claims. The Summary and Abstract sections can set forth one or more but not all exemplary aspects of the present invention as contemplated by the inventor(s), and thus, are not intended to limit the present invention and the appended claims in any way.

[0592] The present invention has been described above with the aid of functional building blocks illustrating the implementation of specified functions and relationships thereof. The boundaries of these functional building blocks have been arbitrarily defined herein for the convenience of the description. Alternate boundaries can be defined so long as the specified functions and relationships thereof are appropriately performed.

[0593] The foregoing description of the specific aspects will so fully reveal the general nature of the invention that others can, by applying knowledge within the skill of the art, readily modify and/or adapt for various applications such specific aspects, without undue experimentation, without departing from the general concept of the present invention. Therefore, such adaptations and modifications are intended to be within the meaning and range of equivalents of the disclosed aspects, based on the teaching and guidance presented herein. It is to be understood that the phraseology or terminology herein is for the purpose of description and not of limitation, such that the terminology or phraseology of the present specification is to be interpreted by the skilled artisan in light of the teachings and guidance.

[0594] The breadth and scope of the present invention should not be limited by any of the above-described exemplary aspects, but should be defined only in accordance with the following claims and their equivalents.

WHAT IS CLAIMED IS:

1. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein

(a) the nucleotide sequence encodes TVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGN SQESVTEQDSKDSTYSLSX[1]TLTLSKADYEKHKVYACEVTHQGLSSPVTKS FNR (SEQ ID NO: 2200), wherein X^1 is selected from N and S, and wherein the nucleotide sequence encodes a kappa light chain constant domain of an antibody or a fragment thereof; or,

(b) the nucleotide sequence encodes PKAAPSVTLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAG VETTTPSKQSNNKYAASSYLSLTPEQWKSHX[2]SYSCQVTHEGSTVEKTVA PX[3]ECS (SEQ ID NO: 2201), wherein X^2 is selected from R and K, and X^3 is selected from T and A, and wherein the nucleotide sequence encodes a lambda light chain constant domain of an antibody or a fragment thereof.

2. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein

(a) the nucleotide sequence encodes SX[4]GPSVX[5]PLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGVHTFPAV LQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKX[6]X[7] (SEQ ID NO: 2202) wherein X^4 is an optional ASTK sequence, X^5 is selected from F and L, X^6 is selected from K and R, and X^7 is selected from V and A, and wherein the nucleotide sequence encodes a CH1 domain of an IgG1 antibody or a fragment thereof; and/or,

(b) the nucleotide sequence encodes APEX[8]X[9]GX[10]PSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNX[11] YVDGVEVHNAKTKPREEQYX[12]STYRVVSVLTVLHQDWLNGKEYKCKVS

NKALPAPIEKTISKAK (SEQ ID NO: 2203) wherein X^8 is selected from L and A, X^9 is selected from L and A, X^{10} is selected from G and A, and X^{11} is selected from V and W, and X^{12} is selected from N and A, and wherein the nucleotide sequence encodes a CH2 domain of an IgG1 antibody or a fragment thereof; and/or,

(c) the nucleotide sequence encodes GQPREPQVYTLPPSRX^{13}EX^{14}TKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG (SEQ ID NO: 2204) wherein X^{13} is selected from E and D, and X^{14} is selected from M and L, and wherein the nucleotide sequence encodes a CH3 domain of an IgG1 antibody or a fragment thereof; and/or,

(d) the nucleotide sequence encodes SASTKGPSVFPLAPCSRSTSESTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVX^{15}SSNFGTQTYTCNVDHKPSNTKVDKTV (SEQ ID NO: 2205) wherein X^{15} is selected from P and T, and wherein the nucleotide sequence encodes a CH1 domain of an IgG2 antibody or a fragment thereof; and/or,

(e) the nucleotide sequence encodes APPVAGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFNWYVDGX^{16}EVHNAKTKPREEQFNSTFRVVSVLTVVHQDWLNGKEYKCKVSNKGLP$X^{17}X^{18}$IEKTISKTK (SEQ ID NO: 2206) wherein X^{16} is selected from V and M, X^{17} is selected from A and S; and X^{18} is selected from P and S, and wherein the nucleotide sequence encodes a CH2 domain of an IgG2 antibody or a fragment thereof; and/or,

(f) the nucleotide sequence encodes GQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPMLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG (SEQ ID NO: 2196), wherein the nucleotide sequence encodes a CH3 domain of an IgG2 antibody or a fragment thereof; and/or,

(g) the nucleotide sequence encodes SASTKGPSVFPLAPCSRSTSESTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTKTYTCNVDHKPSNTKVDKRV

(SEQ ID NO: 2197), wherein the nucleotide sequence encodes a CH1 domain of an IgG4 antibody or a fragment thereof; and/or,

(h) the nucleotide sequence encodes APEFLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSQEDPEVQFNWYVD GVEVHNAKTKPREEQFNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKGL PSSIEKTISKAK (SEQ ID NO: 2198), wherein the nucleotide sequence encodes a CH2 domain of an IgG4 antibody or a fragment thereof; and/or,

(i) the nucleotide sequence encodes GQPREPQVYTLPPSQEEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENN YKTTPPVLDSDGSFFLYSRLTVDKSRWQEGNVFSCSVMHEALHNHYTQK SLSLSLG (SEQ ID NO: 2199), wherein the nucleotide sequence encodes a CH3 domain of an IgG4 antibody or a fragment thereof.

3. A polynucleotide comprising a nucleotide sequence codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof), wherein

(a) the nucleotide sequence encodes $X^1X^2X^3LTQX^4X^5X^6VSX^7X^8X^9GX^{10}X^{11}X^{12}X^{13}X^{14}X^{15}C$ (SEQ ID NO: 2235) wherein X^1 is selected from Q, D, E and S; X^2 is selected from S, I, A, and Y; X^3 is selected from V, Q, A, and E; X^4 is selected from P and D; X^5 is selected from P, N, and A; X^6 is selected from S and A; X^7 is selected from G, T, A, and V; X^8 is selected from A and S; X^9 is selected from P and L; X^{10} is selected from Q, K, and S; X^{11} is selected from R, K, T, and S; X^{12} is selected from V, I, and A; X^{13} is selected from T, K, and R; X^{14} is selected from I and L; and, X^{15} is selected from S at T, wherein the nucleotide sequence encodes the first framework region (FW1) of a lambda light chain variable domain; and/or,

(b) the nucleotide sequence encodes $X^1X^2QX^3TQX^4X^5SX^6X^7SASX^8CDRVTX^9X^{10}C$ (SEQ ID NO: 2239) wherein X^1 is selected from D and A; X^2 is selected from I and V; X^3 is selected from M, L, and V; X^4 is selected from S and F; X^5 is selected from P and T; X^6 is selected from S and T; X^7 is selected from L and V; X^8 is selected from V, I, and A; X^9 is selected from I and M; and, X^{10} is selected

from T and S, wherein the nucleotide sequence encodes the first framework region (FW1) of a kappa light chain variable domain; and/or,

(c) the nucleotide sequence encodes $DX^1X^2X^3TQX^4PX^5SX^6X^7X^8X^9X^{10}GX^{11}X^{12}X^{13}X^{14}X^{15}X^{16}C$ (SEQ ID NO: 2243) wherein X^1 is selected from I and V; X^2 is selected from V, L, and Q; X^3 is selected from M and L; X^4 is selected from S and T; X^5 is selected from L and D; X^6 is selected from L and V; X^7 is selected from P, S and A; X^8 is selected from V and M; X^9 is selected from T and S; X^{10} is selected from P and L; X^{11} is selected from E and Q; X^{12} is selected from P and R; X^{13} is selected from A and V; X^{14} is selected from S and T; X^{15} is selected from I, M, and L; and X^{16} is selected from S and N, wherein the nucleotide sequence encodes the first framework region (FW1) of a kappa light chain variable domain; and/or,

(d) the nucleotide sequence encodes $X^1X^2VX^3TQSPX^4TLSX^5SPGERATLSC$ (SEQ ID NO: 2247) wherein X^1 is selected from E and D; X^2 is selected from I and T; X^3 is selected from L and M; X^4 is selected from G and A; and, X^5 is selected from L and V, wherein the nucleotide sequence encodes the first framework region (FW1) of a kappa light chain variable domain; and/or,

(e) the nucleotide sequence encodes $X^1X^2X^3X^4X^5X^6SGGX^7X^8X^9X^{10}X^{11}GX^{12}SX^{13}X^{14}LX^{15}C$ (SEQ ID NO: 2251) wherein X^1 is selected from E, D, and Q; X^2 is selected from V and A; X^3 is selected from Q, E, and K; X^4 is selected from L and V; X^5 is selected from V and L; X^6 is selected from E and Q; X^7 is selected from G, K, and D; X^8 is selected from L and V; X^9 is selected from V, L, and E; X^{10} is selected from Q, R and K; X^{11} is selected from P, S, and L; X^{12} is selected from G and R; X^{13} is selected from L and R; X^{14} is selected from R and K; and, X^{15} is selected from S and D, wherein the nucleotide sequence encodes the first framework region (FW1) of a heavy chain variable domain; and/or,

(f) the nucleotide sequence encodes $X^1X^2QLX^3QX^4GX^5X^6X^7X^8X^9X^{10}GX^{11}X^{12}X^{13}X^{14}X^{15}SC$ (SEQ ID NO: 2255) wherein X^1 is selected from Q and E; X^2 is selected from V and I; X^3 is selected from V and Q; X^4 is selected from S and P; X^5 is selected from A, S, V, P, T, and G; X^6 is selected from E, G and V; X^7 is selected from V and L; X^8 is selected from K, V, E, and A; X^9 is selected from K, R and Q; X^{10} is selected from P and

S; X^{11} is selected from A, E, S, T, and R; X^{12} is selected from S and T; X^{13} is selected from V and L; X^{14} is selected from K and R; and, X^{15} is selected from V, I, L, and M, wherein the nucleotide sequence encodes the first framework region (FW1) of a heavy chain variable domain; and/or,

(g) the nucleotide sequence encodes $QX^1X^2LX^3X^4X^5GX^6X^7LX^8X^9PX^{10}X^{11}TLX^{12}LTC$ (SEQ ID NO: 2259) wherein X^1 is selected from V and L; X^2 is selected from Q and T; X^3 is selected from Q and R; X^4 is selected from E and Q; X^5 is selected from S and W; X^6 is selected from P and A; X^7 is selected from G and A; X^8 is selected from V and L; X^9 is selected from K and R; X^{10} is selected from S and T; X^{11} is selected from Q and E; and, X^{12} is selected from S and T, wherein the nucleotide sequence encodes the first framework region (FW1) of a heavy chain variable domain; and/or,

(h) the nucleotide sequence encodes $WYQX^1X^2X^3GX^4X^5PX^6X^7X^8I$ (SEQ ID NO: 2236) wherein X^1 is selected from Q and L; X^2 is selected from L,Y, H, and K; X^3 is selected from P and E; X^4 is selected from T, R, K, and Q; X^5 is selected from A and S; X^6 is selected from K, T, V and I; X^7 is selected from L and T; and X^8 is selected from L, M, and V, wherein the nucleotide sequence encodes the second framework region (FW2) of a lambda light chain variable domain; and/or,

(i) the nucleotide sequence encodes $WX^1RQX^2PX^3KX^4LX^5X^6X^7X^8$ (SEQ ID NO: 2252) wherein X^1 is selected from V, I, and F; X^2 is selected from A, S and T; X^3 is selected from G and E; X^4 is selected from G and R; X^5 is selected from E and D; X^6 is selected from W and L; X^7 is selected from V and I; and, X^8 is selected from A, S, and G, wherein the nucleotide sequence encodes the second framework region (FW2) of a heavy chain variable domain; and/or,

(j) the nucleotide sequence encodes $WX^1X^2QX^3X^4GX^5X^6LX^7WX^8G$ (SEQ ID NO: 2256) wherein X^1 is selected from V and I; X^2 is selected from R and K; X^3 is selected from A, M, N, R, K, T, and S; X^4 is selected from P, T, and H; X^5 is selected from Q, K, and R; X^6 is selected from G, R and S; X^7 is selected from E, D, K, Q, and A; and, X^8 is selected from M, I, and V, wherein the nucleotide sequence encodes the second framework region (FW2) of a heavy chain variable domain; and/or,

(k) the nucleotide sequence encodes $WX^1RX^2X^3X^4X^5X^6X^7LX^8WX^9X^{10}$ (SEQ ID NO: 2260) wherein X^1 is selected from I and V; X^2 is selected from Q and H; X^3 is

selected from L, P, S, and H; X^4 is selected from P and S; X^5 is selected from G and E; X^6 is selected from K and R; X^7 is selected from G and A; X^8 is selected from E and Q; X^9 is selected from I and L; and, X^{10} is selected from G and A, wherein the nucleotide sequence encodes the second framework region (FW2) of a heavy chain variable domain; and/or,

(l) the nucleotide sequence encodes $WX^1X^2X^3X^4PX^5KX^6X^7X^8X^9X^{10}IX^{11}$ (SEQ ID NO: 2240) wherein X^1 is selected from Y and F; X^2 is selected from Q and L; X^3 is selected from Q and H; X^4 is selected from K and I; X^5 is selected from G and E; X^6 is selected from A and V; X^7 is selected from P and V; X^8 is selected from K and Q; X^9 is selected from L, T, S, R, P, and V; X^{10} is selected from L and W; and, X^{11} is selected from Y and S, wherein the nucleotide sequence encodes the second framework region (FW2) of a kappa light chain variable domain; and/or,

(m) the nucleotide sequence encodes $WX^1X^2QX^3X^4GQX^5PX^6X^7LIX^8$ (SEQ ID NO: 2244) wherein X^1 is selected from Y, F, and W; X^2 is selected from L and Q; X^3 is selected from K and R; X^4 is selected from P and S; X^5 is selected from S and P; X^6 is selected from Q, K, R, and N; X^7 is selected from L and R; and, X^8 is selected from Y and W, wherein the nucleotide sequence encodes the second framework region (FW2) of a kappa light chain variable domain; and/or,

(n) the nucleotide sequence encodes $WX^1X^2QX^3PGQAPRX^4LIX^5$ (SEQ ID NO: 2248) wherein X^1 is selected from Y and F; X^2 is selected from Q and R; X^3 is selected from K and R; X^4 is selected from L and P; and X^5 is selected from Y, R, and K, wherein the nucleotide sequence encodes the second framework region (FW2) of a kappa light chain variable domain; and/or,

(o) the nucleotide sequence encodes $RFSGSX^1SX^2X^3X^4AX^5LX^6IX^7X^8X^9X^{10}X^{11}X^{12}DEAX^{13}YX^{14}C$ (SEQ ID NO: 2237) wherein X^1 is selected from K, N, S, and I; X^2 is selected from G and S; X^3 is selected from T and N; X^4 is selected from S and T; X^5 is selected from S, T, and F; X^6 is selected from A, T, and G; X^7 is selected from T, H, and S; X^8 is selected from G, N, and R; X^9 is selected from L, V, and A; X^{10} is selected from Q, E, and A; X^{11} is selected from A, T, and I; X^{12} is selected from E and G; X^{13} is selected from D and I; and, X^{14} is selected from Y and F, wherein the nucleotide sequence encodes the third framework region (FW3) of a lambda light chain variable domain; and/or,

(p) the nucleotide sequence encodes RFSGSX^{1}SGX^{2}X^{3}X^{4}X^{5}X^{6}TISSLX^{7}X^{8}X^{9}DX^{10}AX^{11}YX^{12}C (SEQ ID NO: 2241) wherein X^1 is selected from G and R; X^2 is selected from T and Q; X^3 is selected from D, E, and Y; X^4 is selected from F and Y; X^5 is selected from T and S; X^6 is selected from L and F; X^7 is selected from Q and E; X^8 is selected from P, Q, A, and S; X^9 is selected from E and D; X^{10} is selected from F, I, S, L, V, and T; X^{11} is selected from T, S, and V; and, X^{12} is selected from Y and F, wherein the nucleotide sequence encodes the third framework region (FW3) of a kappa light chain variable domain; and/or,

(q) the nucleotide sequence encodes RFSGSGSX^{1}TX^{2}FTLX^{3}ISX^{4}X^{5}X^{6}AX^{7}DVX^{8}X^{9}X^{10}X^{11}C (SEQ ID NO: 2245) wherein X^1 is selected from G and A; X^2 is selected from D and A; X^3 is selected from K, R, and T; X^4 is selected from R and S; X^5 is selected from V and L; X^6 is selected from E and Q; X^7 is selected from E and Q; X^8 is selected from G and A; X^9 is selected from V, D, and F; X^{10} is selected from Y and W; and, X^{11} is selected from Y, F, and W, wherein the nucleotide sequence encodes the third framework region (FW3) of a kappa light chain variable domain; and/or,

(r) the nucleotide sequence encodes RFSGSGSGTX^{1}X^{2}TLTISX^{3}LX^{4}X^{5}EDFAX^{6}X^{7}YC (SEQ ID NO: 2249) wherein X^1 is selected from D and E; X^2 is selected from F and S; X^3 is selected from R and S; X^4 is selected from E and Q; X^5 is selected from P and S; X^6 is selected from V and T; and, X^7 is selected from Y and F, wherein the nucleotide sequence encodes the third framework region (FW3) of a kappa light chain variable domain; and/or,

(s) the nucleotide sequence encodes X^{1}X^{2}X^{3}X^{4}SX^{5}DX^{6}X^{7}X^{8}X^{9}X^{10}X^{11}X^{12}LX^{13}X^{14}X^{15}X^{16}LX^{17}X^{18}EDTX^{19}X^{20}X^{21}X^{22}C (SEQ ID NO: 2253) wherein X^1 is selected from R and K; X^2 is selected from F and V; X^3 is selected from T, I, and A; X^4 is selected from L and I; X^5 is selected from V, R, L, and A; X^6 is selected from R, N, T, D, K, and S; X^7 is selected from S, A and V; X^8 is selected from K, R, and E; X^9 is selected from N, S, R, H, and T; X^{10} is selected from T and S; X^{11} is selected from L, A, and F; X^{12} is selected from Y and F; X^{13} is selected from Q and E; X^{14} is selected from M and V; X^{15} is selected from N, D, and S; X^{16} is selected from S, G, and I; X^{17} is selected from R

and K; X^{18} is selected from A, S, D, V, and P; X^{19} is selected from A and G; X^{20} is selected from V, M, and L; X^{21} is selected from Y and F; and, X^{22} is selected from Y and F, wherein the nucleotide sequence encodes the third framework region (FW3) of a heavy chain variable domain; and/or,

(t) the nucleotide sequence encodes $X^1X^2X^3X^4X^5X^6X^7X^8SX^9X^{10}TX^{11}X^{12}X^{13}X^{14}X^{15}X^{16}X^{17}LX^{18}X^{19}X^{20}DX^{21}X^{22}X^{23}YX^{24}C$ (SEQ ID NO: 2257) wherein X^1 is selected from R, Q, and K; X^2 is selected from V, I, F, G, and A; X^3 is selected from T, A, and K; X^4 is selected from M, I, L, and F; X^5 is selected from T and S; X^6 is selected from T, A, R, V, S, E, and L; X^7 is selected from D, E, and N; X^8 is selected from T, K, Q, S, P, R, I, N, and E; X^9 is selected from T, K, S, A, I, and V; X^{10} is selected from S, N, D, and T; X^{11} is selected from A, V, and T; X^{12} is selected from Y, and F; X^{13} is selected from M and L; X^{14} is selected from E, Q, and D; X^{15} is selected from L, M, W, and I; X^{16} is selected from R, S, D, L, K, T, and N; X^{17} is selected from S and R; X^{18} is selected from R, K, Q, and T; X^{19} is selected from S, H, F, A, and P; X^{20} is selected from D, E, and S; X^{21} is selected from T and S; X^{22} is selected from A and G; X^{23} is selected from V, F, T, and M; and, X^{24} is selected from Y, F, and L, wherein the nucleotide sequence encodes the third framework region (FW3) of a heavy chain variable domain; and/or,

(u) the nucleotide sequence encodes $RX^1X^2X^3X^4X^5DX^6SX^7X^8QX^9X^{10}LX^{11}X^{12}X^{13}X^{14}X^{15}X^{16}X^{17}X^{18}DTAX^{19}X^{20}X^{21}C$ (SEQ ID NO: 2261) wherein X^1 is selected from V and L; X^2 is selected from T and S; X^3 is selected from I and M; X^4 is selected from S and L; X^5 is selected from V, R, and K; X^6 is selected from T and K; X^7 is selected from K and R; X^8 is selected from K and N; X^9 is selected from F and V; X^{10} is selected from S and V; X^{11} is selected from R, T, K, and M; X^{12} is selected from L, I, M, and V; X^{13} is selected from S, T, and N; X^{14} is selected from S and N; X^{15} is selected from V and M; X^{16} is selected from T and D; X^{17} is selected from A and P; X^{18} is selected from A and V; X^{19} is selected from V and T; X^{20} is selected from Y and W; and, X^{21} is selected from Y, F and W, wherein the nucleotide sequence encodes the third framework region (FW3) of a heavy chain variable domain; and/or,

(v) the nucleotide sequence encodes $FGX^1GTX^2X^3TVL$ (SEQ ID NO:2238) wherein X^1 is selected from G and T; X^2 is selected from K and Q; and X^3 is selected from

L and V, wherein the nucleotide sequence encodes the fourth framework region (FW4) of a lambda light chain variable domain; and/or,

(w) the nucleotide sequence encodes $X^1GX^2GTX^3X^4X^5X^6X^7$ (SEQ ID NO: 2242) wherein X^1 is selected from F and L; X^2 is selected from Q, G, and S; X^3 is selected from K and R; X^4 is selected from V and L; X^5 is selected from E, D, and Q; X^6 is selected from I and V; and, X^7 is selected from K and T, wherein the nucleotide sequence encodes the fourth framework region (FW4) of a kappa light chain variable domain; and/or,

(x) the nucleotide sequence encodes $FGX^1GTX^2X^3X^4X^5K$ (SEQ ID NO: 2246) wherein X^1 is selected from Q, A, P, and G; X^2 is selected from K and R; X^3 is selected from V and L; X^4 is selected from E and Q; and X^5 is selected from I and L, wherein the nucleotide sequence encodes the fourth framework region (FW4) of a kappa light chain variable domain; and/or,

(y) the nucleotide sequence encodes $FX^1X^2GTX^3X^4X^5IK$ (SEQ ID NO: 2250) wherein X^1 is selected from G and C; X^2 is selected from Q, G, and P; X^3 is selected from K and R; X^4 is selected from V, L, and A; and, X^5 is selected from E and D, wherein the nucleotide sequence encodes the fourth framework region (FW4) of a kappa light chain variable domain; and/or,

(z) the nucleotide sequence encodes $WGX^1GX^2X^3VTVS$ (SEQ ID NO: 2254) wherein X^1 is selected from Q, R, and K; X^2 is selected from T, I and A; and, X^3 is selected from L, S, T, M, and P, wherein the nucleotide sequence encodes the fourth framework region (FW4) of a heavy chain variable domain; and/or,

(aa) the nucleotide sequence encodes $WGX^1GTX^2X^3TVS$ (SEQ ID NO: 2258) wherein X^1 is selected from R, Q, K, A and S; X^2 is selected from L, M, T, Q, and P; and, X^3 is selected from V and L, wherein the nucleotide sequence encodes the fourth framework region (FW4) of a heavy chain variable domain; and/or,

(ab) the nucleotide sequence encodes $WX^1X^2GX^3X^4VTVS$ (SEQ ID NO: 2262) wherein X^1 is selected from G and D; X^2 is selected from Q and R; X^3 is selected from T and S; and X^4 is selected from T, L, and M, wherein the nucleotide sequence encodes the fourth framework region (FW4) of a heavy chain variable domain.

4. The polynucleotide of any one of claims 1 to 3, further comprising a nucleotide sequence which encodes a linker sequence of formula $(Gly_xSer)_y$, wherein x and y are integers between 1 and 100.

5. The polynucleotide according to claim 4, wherein the linker sequence is interposed between a VH domain and a VL domain.

6. The polynucleotide according to 5, which encodes an scFv.

7. The polynucleotide according to any of claims 1 to 3, therein said polynucleotide encodes a therapeutic antibody or an antigen-binding fragment thereof.

8. The polynucleotide according to claim 7, wherein the therapeutic antibody is selected from the group consisting of abagovomab, abciximab, adalimumab, alemtuzumab, alirocumab, amatuximab, anrukinzumab, arcitumomab, basiliximab, bavituximab, benralizumab, bevacizumab, bezlotoxumab, bimagrumab, bococizumab, brentuximab, briakinumab, brodalumab, canakinumab, cantuzumab, carlumab, cetuximab, cixutumumab, clivatuzumab, conatumumab, crenezumab, dacetuzumab, daclizumab, dalotuzumab, denosumab, drozitumab, dupilumab, dusigitumab, eculizumab, elotuzumab, enokizumab, epratuzumab, etaracizumab, evolocumab, farletuzumab, fasinumab, fezakinumab, ficlatuzumab, figitumumab, fresolimumab, fulranumab, ganitumab, gantenerumab, gevokizumab, girentuximab, glembatumumab, ibalizumab, ibritumomab, icrucumab, inotuzumab, intetumumab, itolizumab, ixekizumab, lebrikizumab, lorvotuzumab, mavrilimumab, mepolizumab, milatuzumab, mogamulizumab, motavizumab, naptumomab, necitumumab, nivolumab, obinutuzumab, ocrelizumab, olaratumab, omalizumab, otelixizumab, oxelumab, pateclizumab, pembrolizumab, pertuzumab, ponezumab, ramucirumab, rilotumumab, rituximab, robatumumab, romosozumab, rontalizumab, samalizumab, sarilumab, secukinumab, sifalimumab, siltuximab, sirukumab, solanezumab, tabalumab, tanezumab, tenatumomab, teplizumab, tigatuzumab, tildrakizumab, tocilizumab, tositumomab, tralokinumab, trastuzumab, urelumab, ustekinumab, vedolizumab, and veltuzumab, and functional fragments thereof.

9. The polynucleotide according to any one of claims 1 to 3, wherein the nucleotide sequence is selected from SEQ ID NOS: 1979-2188, and subsequences thereof.

10. The polynucleotide according to any one of claim 1 to 9, wherein the nucleotide sequence is codon-optimized.

11. The polynucleotide according to any one of claims 1 to 10, where the nucleotide sequence is codon-optimized based on TABLE 1 or TABLE 2 (e.g., MAP1, MAP2, MAP3, MAP4, MAP5, MAP6, MAP7, MAP8, MAP9, MAP10, MAP11, MAP12, MAP13, MAP14, MAP15, MAP16 or any combination thereof).

12. The polynucleotide according to any one of claims 1 to 11, wherein the polynucleotide is an mRNA.

13. The polynucleotide according to claim 12, wherein the mRNA is synthetic.

14. The polynucleotide according to any one of claims 1 to 13, wherein

(i) at least one uridine has been replaced with 2-pseudouridine, 5-methoxyuridine, 2-thiouridine, 4-thiouridine, N1-methylpseudouridine, 5-aza-uridine, 2-thio-5-aza-uridine, 4-thio-pseudouridine, 2-thio-pseudouridine, 5-hydroxyuridine, 4-methoxy-pseudouridine, 4-methoxy-2-thio-pseudouridine, 3-methyluridine, 5-carboxymethyl-uridine, 1-carboxymethyl-pseudouridine, 5-propynyl-uridine, 1-propynyl-pseudouridine, 2-methoxy-4-thio-uridine, 5-taurinomethyluridine, 1-taurinomethyl-pseudouridine, 5-taurinomethyl-2-thio-uridine, 1-taurinomethyl-4-thio-uridine, 5-methyl-uridine, 2-methoxyuridine, 1-methyl-pseudouridine, 4-thio-1-methyl-pseudouridine, 2-thio-1-methyl-pseudouridine, 1-methyl-1-deaza-pseudouridine, 2-thio-1-methyl-1-deaza-pseudouridine, 1-ethyl-pseudouridine, or 2-thio-dihydrouridine; and/or,

(ii) at least one adenosine has been replaced with 2-aminopurine, 2,6-diaminopurine, 7-deaza-adenine, 7-deaza-8-aza-adenine, 7-deaza-2-aminopurine, 7-deaza-8-aza-2-aminopurine, 7-deaza-2,6-diaminopurine, 7-deaza-8-aza-2,6-diaminopurine, 1-methyladenosine, N6-methyladenosine, N6-isopentenyladenosine, N6-(cis-hydroxyisopentenyl)adenosine, 2-methylthio-N6-(cis-hydroxyisopentenyl)adenosine, N6-glycinylcarbamoyladenosine, N6-

threonylcarbamoyladenosine, 2-methylthio-N6-threonyl carbamoyladenosine, N6,N6-dimethyladenosine, or 7-methyladenine; and/or,

(iii) at least one guanosine has been replaced with inosine, 1-methyl-inosine, wyosine, wybutosine, 7-deaza-guanosine, 7-deaza-8-aza-guanosine, 6-thio-guanosine, 6-thio-7-deaza-guanosine, 6-thio-7-deaza-8-aza-guanosine, 7-methyl-guanosine, 6-thio-7-methyl-guanosine, 7-methylinosine, 6-methoxy-guanosine, 1-methylguanosine, N2-methylguanosine, N2,N2-dimethylguanosine, 8-oxo-guanosine, 7-methyl-8-oxo-guanosine, or 1-methyl-6-thio-guanosine; and/or,

(iv) at least one cytidine has been replaced with 5-methylcytidine, 5-aza-cytidine, pseudoisocytidine, 3-methyl-cytidine, N4-acetylcytidine, 5-formylcytidine, N4-methylcytidine, 5-hydroxymethylcytidine, 1-methyl-pseudoisocytidine, pyrrolo-cytidine, pyrrolo-pseudoisocytidine, 2-thio-cytidine, 2-thio-5-methyl-cytidine, 4-thio-pseudoisocytidine, 4-thio-1-methyl-pseudoisocytidine, 4-thio-1-methyl-1-deaza-pseudoisocytidine, 1-methyl-1-deaza-pseudoisocytidine, zebularine, 5-aza-zebularine, 5-methyl-zebularine, 5-aza-2-thio-zebularine, 2-thio-zebularine, 2-methoxy-cytidine, 2-methoxy-5-methyl-cytidine, or 4-methoxy-pseudoisocytidine.

15. A method to treat a disease or condition in a subject in need thereof comprising administering a therapeutically effective amount of (i) a polynucleotide according to any one of claims 1 to 14 or a complement thereof, (ii) a vector or set of vectors comprising said polynucleotide, (iii) a cell comprising said vector or set of vectors, (iv) a pharmaceutical composition comprising said polynucleotide, vector or set of vectors, or cell, or (v) a combination thereof.

CD00098 - IgC: Immunoglobulin constant domain (IgC)

P	C	Master	A	G	I	L	V	M	F	W	P	C	S	T	Y	N	Q	H	K	R	D	E
1	V	119-V	-2	-10	1	1	7	-2	-4	-9	0	-7	-8	-3	-8	-9	-8	-9	-9	-9	-9	-4
2	F	120-S	-3	-8	-1	-3	-1	-4	4	1	-8	-8	2	2	4	0	-1	2	1	1	-3	-3
3	L	121-I	-2	-3	4	3	4	0	-1	-9	3	-8	-6	-3	-8	-9	-8	-9	-4	-6	-9	-9
4	L	122-F	-4	-7	-1	4	-3	-1	5	-3	-2	-8	0	1	3	-3	-1	0	-2	-5	-9	-6
5	P	123-P	1	-4	-2	-2	0	-2	-2	-9	5	1	2	-2	-2	-4	1	1	-2	1	-1	-2
6	P	124-P	-2	-2	-5	-3	-3	-8	-5	-10	6	-9	2	0	-5	-3	-1	1	0	0	-5	-1
7	S	125-S	-1	-3	-8	-4	-4	-4	-1	-3	3	5	4	-1	-3	0	0	1	-1	0	-2	0
8	P	126-S	0	-1	-1	-1	-3	-3	-5	-7	3	4	2	1	-3	-1	-1	-3	-1	0	0	1
9	E	127-E	0	-1	-3	0	0	-4	-1	-2	1	-1	1	-1	-4	0	0	0	1	-2	1	3
10	E	128-Q	-2	-1	-2	-1	-1	-1	-3	-1	0	-4	1	-1	-4	-1	2	0	-2	-2	2	4
11	L	129-L	-1	-1	-1	1	-2	1	0	0	-2	-2	1	0	-1	0	-1	0	1	0	0	0
12	L	130-T	-2	0	-1	1	-3	-1	-1	-2	2	-3	1	2	-3	1	-1	-2	0	-1	0	0
13	G	131-S	-1	3	-5	-4	-5	-3	-7	-7	0	-3	2	0	-7	2	2	-1	1	0	2	1
14	G	132-G	-2	3	-5	-5	-4	-2	-8	-8	-3	-8	1	-1	-4	2	1	0	2	0	2	2
15	S	133-G	-2	-4	-2	-4	-2	-1	-1	-2	3	-8	2	3	-3	0	0	0	2	-2	-2	1
16	V	134-A	2	-4	1	-1	4	-3	-4	-9	-1	-8	0	-1	-3	2	-2	-2	-1	-2	-4	-2
17	T	135-S	-1	-5	-1	-3	0	-3	-2	-1	-1	-8	2	5	-4	0	0	-2	0	-1	-5	0
18	L	136-V	-5	-10	3	6	1	-1	0	-8	-6	-8	-9	-4	1	-10	-9	-9	-9	-9	-10	-9
19	T	137-V	-2	2	3	0	3	-2	-5	-3	-8	-8	0	4	-1	-3	-8	-8	-5	-3	-8	-5
20	C	138-C	-4	-9	-8	-6	-7	-8	-9	-9	-9	12	-7	-7	-5	-9	-10	-10	-10	-10	-10	-10
21	L	139-F	-6	-9	0	5	0	3	1	0	-9	-4	-1	-3	1	-3	0	3	-3	-2	-6	-3
22	A	140-L	4	-6	4	-1	5	-2	-3	-9	-8	-1	-2	-3	-8	-9	-8	-9	-8	-9	-9	-8
23	T	141-N	-4	-6	-5	-3	-2	-1	-2	0	-8	-8	3	5	0	0	-2	0	2	0	0	-1
24	G	142-N	-7	6	-10	-10	-10	-9	-10	-10	-8	-9	-1	-7	-9	3	-3	-1	-1	0	3	-2
25	F	143-F	-5	-4	-2	-1	-5	-1	9	-6	-4	-9	-5	-8	5	-9	-9	-1	-9	-9	-10	-9
26	Y	144-Y	-3	-5	-2	0	-5	-3	4	-2	-4	-7	1	-2	8	-3	-2	2	-5	-1	-7	-4
27	P	145-P	-5	-4	-5	-3	-5	-9	-2	-10	9	-9	-4	-4	-5	-8	-6	-3	-5	-9	-8	-5
28	P	146-K	-1	0	-9	-2	-5	-8	-9	-9	3	-2	1	-3	-3	-1	-1	-2	2	2	3	2
29	D	147-D	-1	0	-5	-6	-1	-3	-2	-9	2	-9	0	0	-4	1	-2	4	-1	-5	5	2
30	I	148-I	0	-10	6	1	5	-1	-1	-9	-3	-8	-8	-5	-8	-10	-9	-10	-9	-9	-10	-9
31	T	149-N	-2	-5	-3	-3	-2	-3	0	-3	-8	-8	2	4	0	3	1	-1	0	-1	0	2
32	V	150-V	-7	-10	4	1	6	3	3	-8	-5	-8	-8	-4	-4	-10	-4	-9	-9	-9	-10	-9
33	T	151-K	-4	-5	-4	-5	-2	-1	-3	-9	-8	-8	3	5	-5	-2	2	2	2	1	-3	1
34	W	152-W	-9	-9	-4	-2	-9	-3	-3	13	-10	-9	-9	-9	-4	-10	-9	-9	-10	-9	-11	-10
35	L	153-K	-6	-5	-1	3	-4	2	1	-1	-8	-8	0	1	0	0	0	-3	3	0	-4	0
36	K	154-I	-5	-4	-2	-2	2	-1	-8	-9	-3	-9	0	-4	-2	-2	2	0	4	4	-1	1
37	N	155-D	-3	0	-6	-7	-6	-9	-9	-1	-8	-9	0	-5	-4	6	0	0	0	-3	5	-1
38	G	156-G	-4	6	-6	-7	-9	-9	-5	-9	-3	-9	1	-1	-9	3	-2	-2	-1	-5	1	-1
39	K	157-S	-1	0	-2	-3	1	-3	-4	-9	-3	-9	0	0	-8	1	3	1	3	-1	-2	2
40	E	158-E	-1	-1	-1	-3	0	-4	-6	-9	3	-9	0	1	-4	1	0	-4	1	-2	1	3
41	L	159-R	-1	-1	3	2	3	1	-4	-2	-4	-8	0	-1	-5	-5	1	-2	-2	-1	0	0
42	T	160-Q	-2	-1	-2	-3	0	-1	-2	-9	2	-8	2	4	-3	0	1	1	-1	-2	-1	0
43	S	161-N	0	0	-3	-1	-1	-7	-5	-3	-1	-8	2	2	-8	0	1	1	0	-4	2	1
44	G	162-G	-1	3	-3	-5	0	-1	-1	-8	-2	-3	1	0	0	1	1	0	-1	-2	1	1
45	V	163-V	-2	-1	0	-1	3	-2	0	-8	-1	-7	0	1	-1	1	2	0	0	-2	-1	0
46	T	164-L	-1	-3	-2	0	1	-1	1	-5	-1	-6	1	2	2	-1	1	1	0	-2	-1	1
47	T	166-S	-1	-2	0	0	-2	-3	-1	-1	-1	-3	1	4	1	-1	-2	1	0	-1	0	0
48	T	167-W	-3	1	-5	-2	-2	-4	0	-1	1	-7	3	3	2	0	1	-3	-2	-1	0	-2
49	P	168-T	0	1	-1	-4	0	-3	-1	-8	3	-8	1	1	1	-1	-1	-1	-2	-4	1	2
50	P	169-D	1	-4	0	0	1	0	0	-1	5	-2	0	1	0	-4	-3	1	-2	-8	-2	-3
51	V	170-Q	0	-5	1	1	2	-3	-3	-9	2	-5	0	0	0	-4	2	-1	1	1	-5	-1
52	P	171-D	-2	-4	-1	0	0	1	1	-9	4	-4	-1	-3	-2	-3	3	-1	2	0	0	0
53	N	172-S	-1	-1	-4	-2	-5	-2	-9	-9	-1	-1	1	0	-8	5	3	3	-1	0	1	0
54	S	173-K	-1	2	-6	-6	-2	-4	-9	-9	0	-9	3	1	-5	1	1	2	1	-2	0	1
55	D	174-D	-3	2	-9	-6	-7	-9	-10	-10	-3	-1	1	-1	-9	4	-4	-7	-2	-4	6	0
56	G	175-S	-3	5	-4	-4	-4	-3	-9	5	-4	-9	0	-3	-4	1	-3	-1	0	-1	1	0

FIG. 1

57	T	176-T	-2	-3	-7	0	-4	-1	-3	-9	-3	-1	2	6	-4	-3	-4	-4	1	-2	-5	-5
58	Y	177-Y	-4	-7	-7	-3	-4	-2	6	-1	-10	-2	-8	-4	9	-9	-3	-6	-6	-8	-5	-6
59	S	178-S	-2	-5	-3	-6	-5	0	1	1	-8	-8	5	2	-1	0	4	0	-4	1	-4	-3
60	V	179-M	1	-3	2	1	3	5	-4	-2	-8	-3	-2	2	-8	-8	0	-8	-2	-1	-9	-5
61	S	180-S	-4	-8	-2	-3	-1	2	2	4	-8	-3	4	3	2	-2	-1	2	-4	1	-4	-5
62	S	181-S	-1	-2	-3	-8	0	-4	-9	-9	-8	-4	7	-3	-5	-2	-7	-2	-1	-4	-7	-7
63	Q	182-T	-3	-6	-1	-1	0	1	-2	4	-8	-4	0	2	4	-7	4	3	1	2	-4	-1
64	L	183-L	-1	-10	1	6	1	1	-1	-8	-9	-8	-8	-1	-2	-10	-9	-5	-9	-9	-10	-9
65	T	184-T	-1	-6	-3	-2	-3	-2	-9	-9	0	-8	1	4	-2	3	1	-2	0	2	0	1
66	V	185-L	-4	-4	5	2	5	0	2	-9	-5	-8	-5	-2	-7	-9	-9	-3	-9	-9	-10	-9
67	S	186-T	-2	-4	-2	-3	-2	-8	-9	-3	4	2	3	3	-8	1	-1	-2	1	-4	0	-2
68	P	187-K	2	-3	-2	-1	-3	-3	-4	-9	4	-8	0	1	-8	-4	0	-4	2	0	-4	0
69	S	188-D	1	0	-4	-6	-4	-1	-9	-9	-5	-3	2	2	-5	0	2	-1	1	0	1	3
70	D	189-E	-1	-4	-4	-2	-2	-1	-2	-1	-2	-4	0	-1	-4	-1	2	-2	0	-1	4	4
71	W	190-Y	-2	-2	-2	-2	-2	-3	0	11	-2	2	-2	-3	0	-2	-3	-1	-2	-3	-2	0
72	N	191-E	-2	-2	-1	-1	-2	0	-1	3	-1	-3	-2	-1	-3	4	1	1	1	0	3	0
73	S	192-R	0	-1	-4	-4	-2	-2	-9	-9	2	-1	2	-1	-3	0	-1	0	2	1	1	2
74	G	193-H	-2	5	-4	-6	-9	-4	-9	-2	-2	-9	-1	-1	-3	0	-1	2	1	1	1	-1
75	D	194-N	1	-3	-9	-9	-5	-8	-9	-9	-8	-9	1	2	-3	2	1	4	2	-1	4	1
76	T	195-S	-3	-3	1	-3	2	0	-6	-3	-2	-8	1	4	-1	1	0	0	1	0	-2	1
77	Y	196-Y	-8	-10	0	-2	0	-3	7	-1	-10	-9	-9	-8	8	-9	-9	-3	-9	-9	-10	-9
78	T	197-T	-1	-4	-2	-8	-1	-3	-4	-3	-8	-3	3	5	0	-2	1	-4	-1	0	-3	0
79	C	198-C	-7	-9	-4	-5	-7	-8	-9	-9	-9	12	-7	-7	-9	-9	-10	-10	-10	-10	-10	-10
80	V	199-E	-3	-9	1	-3	2	1	-5	1	-8	-9	-1	0	-2	-1	1	3	2	3	-7	2
81	V	200-A	1	-9	1	-6	8	-1	-4	-9	-9	-7	-8	-5	-1	-9	-9	-9	-9	-9	-10	-9
82	T	201-T	-2	-1	-6	-3	-3	-2	-9	0	-8	-9	1	3	-2	0	3	0	2	0	0	2
83	H	202-H	-8	-9	-9	-4	-6	-4	-1	-8	-9	-9	-5	-8	2	-3	-2	11	-7	-3	-8	-5
84	E	203-K	0	0	-2	-4	-2	-3	-4	1	3	-2	1	-1	1	-2	1	-4	0	0	-2	3
85	S	204-T	2	3	-6	-4	-8	-8	-9	-9	0	-1	4	1	-3	0	-5	-2	-3	-3	1	-1
86	L	205-S	-2	-2	-1	4	-2	2	-3	0	-4	-6	2	1	-6	0	-2	-3	-1	-4	-2	-3
87	P	206-T	-1	0	-3	-2	-2	-2	-2	-3	3	-3	1	1	-3	1	1	2	1	-2	0	1
88	E	207-S	-1	-1	-2	-1	-1	-2	-1	-1	-1	1	0	2	-1	-1	3	-1	1	0	-1	2
89	P	208-P	-2	-1	-1	-4	-1	-3	-1	-7	5	-2	1	2	-4	1	-2	1	1	-2	0	0
90	L	209-I	-1	-2	3	2	3	1	-1	-2	-8	-2	0	1	-4	-2	0	-4	1	-3	-5	-2
91	T	210-V	-3	-8	1	-3	0	-2	-4	-2	0	-8	2	3	-8	-1	2	-1	-2	1	0	2
92	K	211-K	-1	-2	-1	1	2	-4	-4	-9	-3	-9	-2	-3	-1	-7	1	-1	4	2	-2	0
93	S	212-S	-6	-4	-3	-5	-2	-4	-1	-2	1	1	3	3	0	1	-3	1	0	0	0	1
94	I	213-F	-2	-5	4	0	2	0	2	8	-3	-2	-3	-1	1	-4	-5	-3	-2	-2	-6	-3
95	P	214-N	-1	-4	-5	-6	-2	-3	-6	0	3	-9	2	0	-3	2	-1	-2	1	0	2	2

FIG. 1 (continuation)

CD07699 - IgC_L: Immunoglobulin (Ig) light chain constant (C) domain.

P	C	Master	A	G	I	L	V	M	F	W	P	C	S	T	Y	N	Q	H	K	R	D	E
1	A	112-A	4	-2	-4	-4	-3	-3	-4	-5	5	-3	1	-2	-4	-3	-2	-3	-2	1	-3	1
2	P	113-P	-2	-4	-4	-5	-4	-4	-5	-5	8	-4	0	-2	-4	-3	-3	-4	-2	-4	-3	-3
3	S	114-S	1	-3	-2	-3	1	-2	-4	-4	2	-3	3	4	-3	-2	-2	-3	1	-2	-3	-2
4	V	115-V	-2	-5	3	2	5	0	-2	-4	-4	-3	-4	-2	-3	-5	-4	-5	-4	-4	-5	-4
5	T	116-F	-2	-3	-2	2	-2	-1	1	-4	-3	-3	3	4	-2	-2	-2	-3	-2	1	-3	-3
6	I	117-I	-3	-5	4	4	3	0	-2	-4	-4	-3	-3	1	-3	-4	-4	-5	-4	-4	-5	-4
7	F	118-F	-3	-5	-1	2	1	-1	7	-1	-5	-4	-4	-3	3	-5	-4	-3	-4	-4	-5	-5
8	P	119-P	-2	-4	-4	-4	-4	-4	1	-4	7	4	-2	-3	-3	-4	-3	-3	-2	1	-3	-3
9	P	120-P	-2	-4	-4	-5	-4	-4	-5	-5	8	-4	-2	-3	-5	-4	-3	-4	-3	-4	-3	-3
10	S	121-S	-1	-2	-3	0	-3	-2	-2	-3	2	-3	5	-1	2	-2	-2	-2	-2	-3	-2	-2
11	S	122-D	-1	-3	-3	0	-3	-2	-4	-4	1	-3	3	3	-3	-2	-2	-3	3	-2	1	-2
12	E	123-E	1	-3	-4	-1	-3	-3	-4	-4	-3	-4	1	-2	-4	-2	0	-2	-1	2	2	5
13	E	124-Q	0	-3	-5	-4	-4	-3	-5	-4	-3	-5	-1	-2	-4	-2	4	-2	-1	-2	2	5
14	L	125-L	-3	-4	-1	4	-2	0	-2	-4	-4	-4	-3	-3	-3	1	-2	-3	-2	1	-3	2
15	E	--	-1	-2	-1	2	-1	-1	-1	-2	-1	-2	-1	-1	-2	2	-1	-1	-1	-1	2	3
16	K	126-K	-2	-3	-4	-4	-3	-3	-4	-4	-3	-4	2	3	-3	2	3	-2	3	1	-2	1
17	S	127-S	3	-2	-3	-4	-3	-3	-4	-4	-2	-3	4	2	-3	-2	-2	-3	2	-2	-2	-2
18	G	128-G	-2	4	-5	-5	-4	-4	-5	-5	-3	-4	2	1	-4	4	-2	-2	-2	-3	2	-2
19	K	129-T	-2	0	-4	-4	-3	-3	-4	-5	-3	-4	2	2	-3	-2	-1	-2	6	0	-2	-1
20	A	130-A	5	-2	-2	-3	1	-2	-4	-4	-2	-2	3	-1	-3	-3	-2	-3	-2	-3	-3	-2
21	T	131-S	-1	-3	-3	-3	-2	-2	2	-4	-3	-3	3	5	-2	-2	-2	-3	-2	-3	-2	-2
22	L	132-V	1	-4	1	4	2	0	-2	-4	-4	-3	-3	1	-3	-4	-3	-4	-4	-4	-5	-4
23	V	133-V	-2	-5	1	1	6	-1	-2	-4	-4	-2	-3	1	-3	-4	-4	-5	-4	-4	-5	-4
24	C	134-C	-2	-4	-3	-3	-2	-3	-4	-4	-4	11	-2	-2	-4	-4	-5	-5	-5	-5	-5	-5
25	L	135-L	-3	-5	0	5	1	0	1	-3	-4	-3	-4	-3	-2	-4	-3	3	-4	-4	-5	-4
26	I	136-L	0	-4	4	2	3	3	-2	-4	-4	-3	0	-2	-3	-4	-3	-4	-4	-4	-4	-4
27	N	137-N	-2	-2	-4	-5	-4	-3	-4	-5	-3	-4	3	2	-4	6	-2	-2	-2	-2	2	-2
28	D	138-N	-3	-3	-5	-5	-4	-4	-5	-5	-3	-5	-2	-2	-3	4	2	4	3	-1	5	-1
29	F	139-F	-3	-4	-1	1	-2	3	7	-1	-5	-4	0	-3	1	-4	-4	-3	-4	-4	-5	-4
30	Y	140-Y	-3	0	-3	-3	-3	-3	3	0	-4	-4	0	-3	7	-3	-2	-1	-3	-3	-3	1
31	P	141-P	-2	-4	-2	-3	2	-3	-4	-5	8	-4	-3	-2	-4	-4	-3	-4	-3	-4	-3	-3
32	G	142-R	1	3	-4	-4	-3	-3	-4	-4	-2	-3	3	-1	-3	-1	-2	-2	1	1	3	-1
33	F	143-E	0	0	-1	0	2	-2	4	-3	1	-3	0	-2	-2	-3	-3	-3	-3	-4	2	0
34	A	144-A	4	-3	3	-1	4	-1	-3	-4	-3	-2	-2	-2	-3	-4	-3	-4	-3	-4	-4	-3
35	T	145-K	-2	-3	-4	-4	-3	-3	-5	-5	-3	-4	-1	4	-4	3	2	-2	1	-2	4	2
36	V	146-V	-2	-5	2	-1	6	-1	-2	-5	-4	-2	-3	-2	-3	-5	-4	-5	-4	-4	-5	-4
37	Q	147-Q	0	-3	-4	-4	-3	-3	-4	-4	-3	-4	1	3	-3	2	3	-2	2	2	1	-1
38	W	148-W	-4	-4	-4	-3	-5	-3	-1	12	-5	-4	-4	-4	1	-5	-4	-4	-5	-4	-6	-5
39	K	149-K	-2	-4	-2	2	0	-1	-3	-4	-3	-4	-2	1	-3	-3	-1	-3	5	-1	-3	-2
40	V	150-V	1	0	4	-1	4	-1	-3	-4	-4	-3	0	-2	-3	-4	-3	-4	-3	-4	-4	-4
41	D	151-D	0	-2	-4	-4	-4	-4	-4	-5	-3	-4	-1	-2	-4	4	-2	-2	-2	-3	6	0
42	G	152-N	-1	6	-4	-4	-4	-4	-4	-4	-3	-3	0	-2	-4	3	-2	-3	-2	-3	-2	-3
43	A	153-A	2	-2	-3	-3	-2	-2	-4	-4	-2	-3	3	2	-3	2	-1	-2	3	-1	-2	-2
44	T	154-L	0	-3	-2	-1	1	-2	-3	-4	3	-3	1	3	-3	-2	-1	-3	-2	-2	-2	2

FIG. 2

45	V	155-Q	-2	-4	3	-1	4	-1	-3	-3	-3	-3	-2	-2	-3	-3	5	-2	-1	1	-3	-1
46	S	156-S	-1	-2	-3	-3	-3	-2	-4	-4	-2	-3	3	2	-3	2	4	-2	2	-1	-2	-1
47	S	157-G	0	0	-4	-5	-4	-4	-5	-5	-3	-4	3	-1	-4	1	-1	-2	-2	-2	4	3
48	G	158-N	-2	6	-5	-5	-5	-4	-5	-4	-4	-4	-2	-3	-5	3	-3	-3	-3	-4	-2	-3
49	V	159-S	-2	-4	3	-1	6	-1	-2	-4	-4	-2	1	-1	-3	-4	-3	-4	-4	-4	-4	-4
50	Q	160-Q	-2	-3	-3	1	-3	-2	2	-3	-3	-4	1	2	-2	-2	3	-2	1	-2	1	2
51	T	161-E	-2	-3	-3	-3	-2	-3	-4	-4	-3	-3	0	6	-3	2	-2	-3	-2	1	-2	0
52	S	162-S	-1	1	-4	-4	-3	-3	-4	-4	-3	-3	5	2	-4	-1	-2	-3	-2	-3	2	-2
53	T	163-V	-2	0	-3	-4	0	-3	-4	4	-3	-3	2	3	-3	2	-2	-3	2	-2	1	-2
54	T	164-T	2	-3	-2	-3	1	2	-4	-4	4	-3	-1	3	-3	-3	-2	-3	1	-2	-3	1
55	E	165-E	-2	-3	1	-2	2	-2	-4	-5	1	-3	1	3	-3	-2	-2	-3	1	-2	1	2
56	Q	166-Q	-2	-3	-4	-4	-4	-2	-5	-4	-3	-4	-2	1	-3	-2	6	-2	4	2	-2	0
57	Q	167-D	0	-3	-4	-3	-3	3	-4	-4	-3	-4	0	-2	-3	-1	5	-2	1	-1	3	2
58	S	168-S	-1	1	-3	-3	-3	-3	-3	-3	5	-3	3	-1	-3	-2	-2	-2	1	-2	-2	-2
59	D	170-D	-3	0	-4	-5	-4	-4	-5	-5	-3	-5	-1	1	-4	3	-2	-2	-2	-3	6	-1
60	N	171-S	-2	1	-5	-5	-4	-3	-4	-5	-3	-4	2	-1	-3	5	1	4	1	-2	-1	-1
61	T	172-T	-2	-3	-3	-3	-2	-2	-4	-4	-3	-3	-1	5	-3	1	2	-3	3	-1	-2	-2
62	Y	173-Y	-4	-5	-2	-2	-3	-2	5	0	-5	-4	-4	-3	8	-4	-3	0	-4	-4	-5	-4
63	S	174-S	0	-2	-3	-3	-3	4	-4	-4	-3	-3	5	-1	-3	-1	2	-2	1	-2	-2	-2
64	L	175-L	2	-3	1	2	-1	6	-2	-4	-3	-3	1	1	-3	-3	-3	-4	-3	-3	-4	-3
65	S	176-S	0	-2	1	-3	-2	-3	-4	-4	-3	-2	6	0	-3	-1	-2	-3	-2	-2	-2	-2
66	S	177-S	0	-2	-4	-4	-3	-3	-4	-4	-2	-2	6	0	-3	-1	-2	-2	-2	-2	-2	-2
67	Y	178-T	-2	-3	-3	-3	-3	-3	0	-1	-3	-3	2	4	6	-2	-2	-1	-2	1	-3	-3
68	L	179-L	-3	-5	0	6	-1	1	-1	-3	-5	-3	-4	-3	-3	-5	-4	-4	-4	-4	-5	-5
69	T	180-T	-1	-3	-2	-3	1	-2	-4	-4	-3	-2	4	5	-3	-1	-2	-3	-2	-3	-2	-2
70	L	181-L	-3	-5	2	5	3	3	-1	-4	-4	-3	-4	-2	-3	-5	-4	-4	-4	-4	-5	-4
71	T	182-S	-1	-3	-3	-3	-2	-2	-4	-4	-3	-3	4	5	-3	-1	3	-2	-2	-2	-2	-1
72	K	183-K	2	-2	-4	-4	-3	-3	-4	-4	1	-3	3	1	-3	-2	-1	-3	4	2	-2	-2
73	S	184-A	1	-3	-3	0	-3	-3	-4	-5	-3	-3	3	-1	-4	-1	-1	-3	1	-2	3	2
74	D	185-D	-3	-3	-4	-4	0	-3	-5	-5	-3	-5	-2	-2	-4	-1	4	-2	-1	-2	5	4
75	W	186-Y	-3	-3	1	-2	-2	-2	0	10	-3	-3	-3	-2	5	-3	-2	-2	-3	-3	-3	-3
76	N	187-E	-3	-3	-4	-4	-4	-3	2	5	-3	-4	0	-2	-2	5	-1	-2	2	-2	2	2
77	K	188-K	-1	-3	-4	-3	-3	3	-4	-4	-3	-4	4	-1	-3	-2	-1	-2	4	1	-2	1
78	H	189-H	-3	-3	-5	-4	-4	-3	-3	-4	-3	-4	1	-2	3	-1	-1	7	1	-2	4	-1
79	K	190-K	-2	-3	-5	-4	-4	-3	-5	-5	-3	-4	2	-2	-3	3	0	-2	4	4	-1	2
80	V	191-V	-2	-4	-1	0	2	-2	-3	-4	2	-3	1	3	-3	-2	-2	-3	-2	1	-3	1
81	Y	192-Y	-3	-5	2	1	1	3	1	0	-4	-4	-3	-3	7	-4	-3	-1	-4	-4	-5	-4
82	T	193-A	1	-2	-3	-3	-2	-3	-4	-4	-3	-2	4	6	-3	-1	-2	-3	-2	-3	-2	-2
83	C	194-C	-2	-4	-3	-3	-2	-3	-4	-4	-4	11	-2	-2	-4	-4	-5	-5	-5	-5	-5	-5
84	E	195-E	-2	-4	-4	-3	-3	3	-4	-4	-3	-4	-2	1	-3	-2	3	4	1	1	-1	4
85	V	196-V	3	-3	2	-1	5	-1	-3	-4	-3	-2	-2	-2	-3	-4	-3	-4	-3	-4	-4	-4
86	T	197-T	-2	-3	-2	-3	0	-2	-4	-4	-3	-3	-1	6	-3	-2	-2	-3	3	-1	-3	-2
87	H	198-H	-3	-4	-3	-4	0	-2	-3	-4	-4	-4	-2	-3	-1	-1	4	9	-2	-2	-3	-1
88	E	199-Q	1	2	-5	-4	-4	-3	-5	-5	2	-4	-1	-2	-4	1	2	-2	1	-2	-1	5
89	G	200-G	1	4	1	-3	-2	-3	-4	-4	-3	-3	3	2	-4	-2	-3	-3	-3	-3	-3	-3
90	L	201-L	0	-2	-1	1	1	-1	-2	-2	-1	-2	2	-1	-2	-1	-1	-2	-1	-2	2	1
91	S	202-S	-1	-1	-2	-2	-1	-1	-2	-2	-1	-2	3	2	-2	-1	0	-1	1	-1	-1	2

FIG. 2 (continuation)

92	S	203-S	1	-3	-2	0	0	-2	-3	-4	-3	-3	4	2	-3	-2	2	-3	1	-2	-3	-2
93	T	204-P	-2	-3	-3	-3	-3	-3	3	-3	5	-3	1	4	-2	-2	-2	-3	-2	-3	-3	0
94	I	205-V	-2	-4	5	-1	2	-1	-3	-4	-4	-3	2	-2	-3	-3	1	-3	1	1	-3	-2
95	T	206-T	-1	-3	-2	-3	1	-2	-4	-4	-3	-3	3	4	-3	-2	-1	-3	-1	1	-2	2
96	K	207-K	1	-3	-4	-1	-3	-2	-4	-4	-3	-4	1	-2	-3	-2	3	-2	5	0	-2	1
97	S	208-S	-1	1	-4	-4	-3	-3	-4	-5	-3	-3	5	2	-4	-1	-2	-3	-2	-3	2	1
98	I	209-F	-2	-4	4	1	2	-1	4	-3	1	4	0	-2	-1	-4	-4	-4	-4	-4	-4	-4
99	K	210-N	1	-3	-4	-4	-3	-3	-4	-5	-3	-4	-1	1	-3	3	-1	-2	5	2	-2	-1
100	R	211-R	-3	-4	-3	0	-3	-2	2	-3	1	-5	-2	-3	3	-3	-1	-2	3	5	-3	1

FIG. 2 (continuation)

CD04985 - IgC_CH1: First Ig constant domain (IgC) of Ig heavy chains.

P	C	Master	A	G	I	L	V	M	F	W	P	C	S	T	Y	N	Q	H	K	R	D	E
1	A	128-A	4	-2	-5	-5	-4	-4	0	-5	4	-4	3	-1	-4	0	-3	-4	0	-4	-4	-3
2	P	129-P	-4	-5	-6	-6	-5	-5	-7	-7	9	-6	-4	-4	-6	-5	-4	-5	-4	-5	-4	-4
3	T	130-T	-2	-4	-4	-5	-4	-4	-5	-6	-4	-4	5	5	-5	0	-3	-4	-1	1	-3	-3
4	V	131-L	-3	-6	1	4	6	-1	-3	-5	-6	-4	-5	-3	-4	-6	-5	-6	-5	-5	-6	-6
5	F	132-F	-5	-6	-3	-3	-4	-3	8	-1	-6	-5	-5	-5	6	-6	-6	-3	-6	-5	-6	-6
6	P	133-P	-1	-2	-1	-5	-5	-5	-6	-7	8	-6	-4	-4	-6	-5	-4	-5	-4	-5	-4	-4
7	L	134-L	-4	-7	-1	6	0	2	-2	-5	-6	-4	-5	-4	-4	-6	-5	-6	-5	-5	-7	-6
8	V	135-V	1	-5	3	-1	3	1	0	-5	-5	3	1	1	0	-4	-4	-5	-1	-2	-5	-4
9	P	136-S	-1	-2	-4	1	-4	-3	-5	-6	5	-5	3	0	-5	-3	2	-4	-3	-1	-4	-1
10	C	137-C	-3	-2	-4	-4	-4	-4	-1	-5	0	10	2	-3	0	-5	-5	-5	-5	-6	-5	-5
11	C	138-E	-3	3	-5	-5	-5	-5	-5	-5	0	7	3	-3	0	-1	-1	-4	-3	-1	-1	2
12	S	139-N	2	2	-5	-5	-5	-4	-6	-6	0	2	4	-1	-5	-1	1	-4	1	-3	0	1
13	S	140-S	-2	1	-5	-1	-5	-4	-5	-6	3	-5	3	1	-5	1	0	2	-3	-1	-1	1
14	S	141-N	0	0	0	0	-1	-1	-2	-2	-1	-2	2	1	-2	2	0	-1	0	-1	0	-1
15	S	142-P	-2	-3	-1	-3	-3	-3	-4	-4	4	-3	3	3	-3	2	-2	-3	-2	-3	0	-2
16	S	143-S	0	1	-4	-1	-2	0	-5	-6	-4	-5	3	0	-5	-1	-1	-4	0	-1	4	-1
17	D	144-S	-1	4	-6	-6	-5	-5	-6	-6	-4	-5	2	1	-5	0	-3	-4	-4	-4	5	-3
18	S	145-T	-3	-4	-4	-2	-1	2	0	-5	5	-5	3	1	0	2	-1	0	-1	-4	-4	-3
19	V	146-V	-2	-6	3	-1	6	1	0	-5	-5	-4	-4	-3	1	-6	-5	-5	-2	-5	-6	-5
20	T	147-A	3	-4	0	-4	1	-3	-5	-5	-4	-4	2	6	-4	-3	-4	-5	-4	-4	-4	-4
21	L	148-V	-4	-6	4	4	2	4	1	-4	-6	-4	-5	-1	2	-6	-5	-5	-5	-5	-6	-6
22	G	149-G	-1	7	-7	-6	-6	-5	-6	3	-5	-5	-3	-4	-5	-3	-5	-5	-4	-5	-4	-5
23	C	150-C	-3	-5	-4	-1	-4	-4	-5	-5	-6	11	-4	-4	-5	-6	-6	-6	-6	-6	-6	-7
24	L	151-L	-1	-6	1	6	-2	-1	-2	-5	-6	-4	-5	-4	-4	-6	-5	-6	-5	-5	-6	-6
25	A	152-A	5	-1	1	-3	4	-3	-5	-6	-4	-3	-1	-3	-4	-5	-4	-5	-4	-5	-5	-4
26	T	153-Q	-3	-5	-4	-2	-4	1	0	-5	-4	-5	-2	5	0	-3	2	2	4	1	-4	-3
27	G	154-D	-3	6	-6	-7	-6	-6	-6	-6	-5	-6	-3	-4	-6	0	-4	-4	-4	-5	5	0
28	F	155-F	-5	-6	-3	-3	-4	-3	7	-1	-6	-5	-5	-5	7	-5	-5	-2	-5	-5	-6	-6
29	L	156-L	-2	-5	0	4	-3	-2	6	-3	-6	-5	0	-1	1	-1	0	-4	-5	-5	-5	-5
30	P	157-P	-4	-2	-6	-6	-5	-5	-7	-7	9	-6	-4	-4	-6	-5	-4	-5	-4	-5	-4	-4
31	E	158-D	1	-4	-6	-6	-5	-5	-6	-6	-4	-5	2	-1	-5	0	2	-3	-1	-3	3	5
32	P	159-S	-1	-1	-1	-5	-4	-5	-6	-6	6	-5	3	2	-5	-3	-4	-5	-4	-4	0	-3
33	V	160-I	2	-5	3	2	5	2	-3	-5	-5	-4	-4	-1	-4	-5	-5	-6	-5	-5	-6	-5
34	T	161-T	0	-4	-4	-4	-4	1	-5	-6	-4	-4	3	5	-5	0	-3	-4	-3	-4	0	1
35	F	162-F	-4	-6	0	-1	5	1	7	-3	-6	-4	-5	-4	1	-6	-5	-5	-6	-6	-6	-6
36	T	163-S	-2	0	-5	-5	-4	-4	-5	-6	-4	-4	4	5	-5	0	1	-4	1	-3	-1	-3
37	W	164-W	-5	-5	-6	-5	-6	-4	-2	13	-7	-5	-6	-5	0	-7	-5	-5	-6	-6	-7	-6
38	N	165-K	-3	0	-5	-5	-5	-4	0	-5	-4	-5	3	2	0	5	-3	-3	2	0	1	-3
39	S	166-Y	-1	-4	-5	-5	-5	-5	-4	-5	-1	-5	3	-1	4	1	-2	-3	3	-3	4	1
40	G	167-K	-3	3	-6	-6	-5	-5	-6	-6	-4	-5	2	-3	-5	2	4	-3	2	-3	3	0
41	S	168-N	-1	3	-6	-6	-5	-5	-6	-6	0	-5	4	-1	-5	5	-3	-3	-1	-4	0	-3
42	N	169-N	-2	1	0	0	-4	-3	-1	-5	-5	-5	1	1	-4	4	0	-4	2	0	-3	-1
43	S	170-S	1	-2	0	-2	-2	-2	-3	-3	2	-2	2	2	-2	0	1	-2	0	-2	1	0
44	T	172-I	-1	1	1	2	0	1	-4	-5	-5	-4	1	4	-5	1	-4	-4	-4	-4	0	-2

FIG. 3

45	S	173-S	-1	1	-1	-2	-1	-4	0	-5	0	-4	3	4	-4	-3	0	-4	-4	-4	2	-3
46	G	174-S	-3	5	-5	-2	-5	-4	-5	-6	-5	-5	0	3	-5	1	0	-4	-1	-4	2	-3
47	V	175-T	0	-2	1	1	3	2	2	-4	0	-4	1	2	1	-4	-4	-5	-4	-5	-5	-5
48	K	176-R	-1	-5	-1	0	0	3	-5	-6	-4	-5	-3	-4	-4	-3	1	2	4	4	0	-1
49	T	177-G	-3	-2	-4	-5	-1	-4	-5	-6	-4	-5	0	5	-5	5	3	-3	1	-3	0	-3
50	Y	178-F	-5	-6	0	0	-3	-3	7	-1	-6	-5	-5	-5	7	-6	-5	-2	-5	-5	-6	-6
51	P	179-P	-4	-5	-6	-6	-5	-5	-7	-7	9	-6	-4	-4	-6	-5	-4	-5	-1	-5	-4	-4
52	A	180-S	4	-3	-5	-5	-4	-4	-5	-6	4	-4	4	0	-5	-3	-3	-4	0	-4	-4	-3
53	V	181-V	1	-5	2	-1	6	3	-4	-6	-5	-4	-2	-1	-4	-5	-5	-6	-5	-5	-6	-5
54	L	182-L	-4	-2	2	4	-2	2	0	-5	-5	-5	-4	0	-4	-5	3	-4	-3	1	-5	-1
55	Q	183-R	-1	-2	-2	-2	-4	-4	-5	-6	-4	-5	3	3	-5	3	4	-3	0	-1	-1	-1
56	S	--	1	-1	-2	-2	-2	-2	-2	-2	1	-2	3	1	-2	-1	2	-1	2	-1	-1	-1
57	G	184-G	-1	5	-5	-2	-2	-4	-5	-5	-4	-4	3	-3	-5	3	-3	-4	1	-3	-3	-1
58	G	185-G	-3	6	-6	-6	-6	-5	-6	-6	0	-5	-1	-1	-6	2	-4	-4	-4	-5	2	-4
59	L	186-K	0	-5	-3	3	-3	-2	-1	-5	0	-4	-1	4	0	-4	-3	1	2	0	-5	-4
60	Y	187-Y	-5	-6	-4	-2	-4	-4	1	-1	-6	-5	-5	-4	9	-5	-4	-1	-1	-4	-6	-5
61	T	188-A	1	-4	0	-1	-1	-3	-5	-5	-4	-4	2	6	-4	-3	-1	-4	-3	0	-1	-3
62	A	189-A	3	2	-4	1	-3	3	-1	-5	-4	-4	0	2	-4	-4	2	-4	-3	1	-4	-4
63	S	190-T	-1	-4	-1	-4	2	-3	-5	-6	-4	-4	5	4	-4	-3	-3	-4	-3	-4	-4	-4
64	S	191-S	-2	-1	-5	-5	-4	-4	-5	-6	-4	-4	7	1	-5	0	-3	-4	-3	-4	-3	-3
65	Q	192-Q	-4	-5	-4	0	0	-3	-5	-5	-4	-5	-2	-3	-4	-3	7	2	0	1	-3	0
66	L	193-V	-1	-6	3	5	3	2	-3	-5	-6	-4	-5	-4	-4	-6	-5	-6	-5	-5	-6	-6
67	T	194-L	-1	-5	-3	1	-1	-3	-5	-5	-4	-4	-1	5	-4	-1	3	-4	-1	2	0	-3
68	V	195-L	-4	-6	4	3	5	2	-3	-5	-5	-4	-2	-3	-4	-6	-5	-6	-5	-5	-6	-6
69	P	196-P	1	-4	-5	-5	-4	-5	-6	3	7	-5	2	1	-5	0	-4	-4	-1	-4	-4	-4
70	A	197-S	3	0	-5	-3	-4	-4	0	-5	0	-5	2	-3	-4	-3	1	-4	0	0	0	3
71	S	198-K	2	-1	-5	-5	-1	-4	-5	-5	-4	-4	4	1	0	0	0	-4	1	-3	2	0
72	E	199-D	-3	-4	-5	-5	-2	-4	-5	3	-1	-5	2	-3	-4	0	4	-3	-1	-3	4	4
73	W	200-V	-2	0	0	-2	-1	-2	1	8	-2	6	-2	-2	-1	-2	-2	-2	-2	-2	-2	-2
74	K	203-G	-1	0	-3	0	-1	1	-3	6	1	-4	-1	-3	-3	-1	2	-3	3	-1	2	-2
75	G	204-T	2	2	-5	-5	-5	-5	-6	-6	-1	3	2	-1	-5	2	-3	0	2	-3	1	2
76	K	205-N	0	1	-2	-5	-4	-4	-6	-6	1	-5	2	1	-5	1	0	-4	3	2	1	-3
77	E	206-E	-1	-1	-6	-6	-5	-4	-6	-6	-4	-6	0	-1	-5	2	4	-3	4	-2	1	4
78	S	207-H	0	-4	0	-2	-4	2	-4	-5	1	-5	3	1	2	3	-1	1	0	-3	0	-1
79	F	208-V	-4	-6	3	-1	2	-2	7	-2	-6	-5	-5	-4	5	-6	-5	-3	-6	-5	-6	-6
80	Y	209-V	-4	-5	-2	-4	1	-4	1	-3	-5	-5	-1	2	5	-4	1	-3	2	1	-4	2
81	C	210-C	-3	-5	-4	-4	-4	-4	-5	-5	-6	11	-4	-4	-5	-6	-6	-6	-6	-6	-6	-7
82	K	211-K	-2	-4	-5	-3	-1	-4	-5	-6	-4	-5	2	2	-4	3	0	3	3	2	-3	1
83	V	212-V	2	-5	1	-2	7	-2	-4	-6	-5	-4	-4	-1	-4	-5	-5	-6	-5	-5	-6	-5
84	E	213-Q	1	-4	-5	-5	-5	-4	-6	-6	-4	-5	-2	2	-4	1	3	3	3	0	-3	4
85	H	214-H	-4	-5	-6	-2	-6	-4	-4	-5	-5	-6	-1	-4	1	1	-2	10	-3	0	-4	-3
86	P	215-P	1	1	-5	-3	-5	-4	-4	-4	4	-5	2	-1	3	-1	-3	-4	1	-4	1	-1
87	S	216-N	2	2	-3	-3	-3	-3	-4	-4	3	-3	2	0	-3	1	-2	-3	-2	-3	1	0
88	T	217-G	0	0	-3	0	-2	-2	0	-3	3	-3	1	3	-3	-2	0	-3	2	-2	-3	-2
89	S	218-N	-1	-2	-4	-2	1	-4	-5	-6	-1	-5	3	-3	-5	3	2	-3	2	1	0	0
90	V	219-K	0	2	1	-4	3	-3	-5	3	-5	2	0	1	-4	0	2	-4	0	1	-4	-4
91	D	220-E	-4	-4	-1	-5	-5	-5	-6	-6	1	-5	-1	3	-5	3	0	3	-1	-3	4	3

FIG. 3 (continuation)

92	K	221-K	-1	-5	0	2	2	-2	0	-5	-5	-5	-4	-4	1	-4	-1	-4	5	-2	-1	-3
93	T	222-D	-4	-5	-5	-5	-1	-4	-5	3	1	-5	-3	3	-4	1	0	2	3	-1	1	3
94	V	223-V	-3	-6	5	0	5	-2	-3	-5	-1	-4	-4	-1	1	-5	-5	-5	1	-4	-6	-5
95	P	224-P	-1	-2	-5	-6	-2	-5	-6	-6	7	-5	-1	-1	-5	0	-3	-4	-1	-1	2	2

FIG. 3 (continuation)

CD04986 - IgC_CH2: Second Ig constant domain (IgC) of Ig heavy chains.

P	C	Master	A	G	I	L	V	M	F	W	P	C	S	T	Y	N	Q	H	K	R	D	E
1	C	244-C	-1	-3	-3	-3	-2	-2	-4	-4	5	8	2	-1	-3	-2	-2	-3	-2	-3	-3	-2
2	C	245-C	-1	2	0	-1	3	-1	-3	-3	-3	7	-2	-1	-3	-3	-3	-4	-3	-4	-4	-4
3	K	246-H	-2	-3	-4	-3	-3	-2	-3	-3	-2	-4	-1	-2	-1	-1	3	5	3	0	-1	3
4	P	247-P	-1	-3	-4	-4	-3	-3	-4	-4	8	-4	-1	-2	-4	-3	-2	-3	-2	-3	-2	-2
5	S	248-R	0	-1	-3	-3	-3	-2	-3	-4	-2	-2	5	0	-2	0	0	-1	0	3	-1	-1
6	L	249-L	-2	-4	1	4	3	1	-1	-3	-3	-2	-3	-1	-2	-4	-3	-4	-3	-3	-4	-3
7	S	250-S	1	-1	-3	-3	-2	-2	-3	-4	-1	-2	5	1	-2	0	-1	-2	-1	-1	-1	-1
8	L	251-L	-2	-4	1	5	0	2	0	-2	-4	-2	-3	-2	-2	-4	-3	-4	-3	-3	-4	-4
9	Q	252-H	-2	-3	-1	2	-2	0	-2	-3	-3	-3	-2	-2	-1	-1	5	5	-1	-1	-2	0
10	R	253-R	-1	3	-4	-3	-3	-2	-4	-3	-3	-4	-1	-2	-3	-1	0	-1	1	5	-2	-1
11	P	254-P	-1	-2	-3	-3	-3	-3	-4	-4	7	-3	3	-1	-3	-1	-1	-2	-1	-2	-2	-1
12	A	255-A	3	-1	-3	-3	-2	-2	-3	-4	-2	-2	3	0	-3	-1	-1	-2	-1	-2	3	-1
13	L	256-L	-1	-3	0	4	0	1	-1	-3	-3	-2	2	-1	-2	-2	-2	-3	-2	-2	-3	-2
14	E	257-E	-1	2	-1	2	-2	-1	-2	-3	-3	-3	-1	-2	-3	-2	0	-2	-1	-2	-1	4
15	D	258-D	-2	-2	-4	-4	-4	-3	-4	-4	-2	-4	-1	-2	-3	0	4	-1	0	-1	6	1
16	L	259-L	-1	-3	0	4	0	1	-1	-3	-3	-2	2	-1	-2	-2	-2	-3	-2	-2	-3	-2
17	L	260-L	-1	-2	1	4	0	1	0	-1	-2	-1	-2	-1	-1	-2	-1	-2	-2	-2	-2	-2
18	L	261-L	-1	-3	0	4	0	1	-1	-3	-3	-2	2	-1	-2	-2	-2	-3	-2	-2	-3	-2
19	G	262-G	-1	6	-4	-4	-4	-3	-4	-3	-3	-3	-1	-2	-4	3	-2	-2	-2	-3	-1	-2
20	S	263-S	0	-1	-3	-4	-3	-3	-3	-4	-2	-2	4	0	-3	0	-1	-2	-1	-2	4	0
21	S	264-E	-1	-2	-4	-4	-3	-3	-4	-4	-2	-3	3	0	-3	3	0	-1	-1	-1	3	3
22	A	265-A	5	0	-2	-2	-1	-2	-3	-3	-1	-1	1	-1	-2	-2	-1	-2	-1	-2	-2	-2
23	S	266-N	3	-1	-3	-3	-2	-2	-3	-4	-2	-2	3	0	-3	3	-1	-1	-1	-2	-1	-1
24	L	267-L	-2	-4	3	5	1	1	0	-3	-4	-2	-3	-2	-2	-4	-3	-4	-3	-3	-4	-4
25	T	268-T	-1	-3	0	2	0	0	-2	-3	-2	-2	0	5	-2	-1	-2	-3	-2	-2	-2	-2
26	C	269-C	-1	-3	-2	-2	-1	-2	-3	-3	-4	10	-2	-2	-3	-3	-4	-4	-4	-4	-4	-4
27	T	270-T	-1	-3	0	2	0	0	-2	-3	-2	-2	0	5	-2	-1	-2	-3	-2	-2	-2	-2
28	L	271-L	-2	-4	1	5	0	2	0	-2	-4	-2	-3	-2	-2	-4	-3	-4	-3	-3	-4	-4
29	S	272-T	-1	-2	-3	-3	-2	-2	-3	-4	-2	-2	3	3	-2	3	-1	-1	0	3	-1	-1
30	G	273-G	0	5	-4	-4	-3	-3	-4	-3	-2	-3	3	-1	-3	-1	-2	-2	-2	-2	-2	-2
31	L	274-L	-2	-4	0	4	0	1	1	0	-4	-2	-3	-2	5	-4	-3	-1	-3	-3	-4	-3
32	R	275-R	-1	-2	-2	1	-2	-1	-2	-3	-2	-3	3	-1	-2	-1	0	-2	0	4	-2	-1
33	D	276-D	-1	-1	-2	-2	-2	-2	-2	-2	-1	-2	0	-1	-2	1	-1	-1	-1	-1	6	1
34	P	277-A	2	-2	-3	-3	-2	-3	-4	-4	7	-3	-1	-1	-3	-3	-2	-3	-2	-3	-2	-2
35	E	278-S	-1	-2	-4	-4	-3	-3	-4	-4	5	-3	2	-1	-3	-1	0	-2	-1	-1	0	4
36	G	279-G	0	6	-5	-4	-4	-3	-4	-3	-3	-3	-1	-2	-4	-1	-3	-3	-2	-3	-2	-3
37	A	280-V	5	-1	0	-1	2	-1	-3	-3	-2	-1	0	-1	-2	-3	-2	-3	-2	-2	-3	-2
38	V	281-T	-1	-3	0	-1	3	-1	-2	-4	-2	-2	-1	2	-2	-2	-1	-2	3	0	-2	-1
39	F	282-F	-2	-4	1	0	3	0	6	-1	-4	-2	-3	-2	2	-4	-4	-2	-4	-3	-4	-4
40	T	283-T	0	-2	-2	-2	-1	-2	-3	-3	-2	-2	3	5	-2	0	-1	-2	-1	-2	-1	-1
41	W	284-W	-3	-3	-3	-2	-4	-2	0	12	-4	-3	-4	-3	2	-5	-3	-3	-4	-3	-5	-4
42	T	285-T	-1	-3	-2	-3	-2	-2	-3	-3	-2	-2	0	4	-3	-1	0	-2	-1	-1	0	4
43	P	286-P	-2	-3	-4	-3	-3	-3	-4	-4	7	-4	-1	-2	-3	-2	-1	-2	0	4	-2	-1
44	S	287-S	0	-2	-3	-3	-2	-2	-3	-4	-2	-2	3	3	-3	0	-1	-2	-1	-2	4	0
45	G	288-S	-1	3	-2	-2	-2	-2	3	-2	-2	-2	2	2	-1	-1	-2	-2	-2	-2	-2	-2

FIG. 4

46	G	289-G	0	5	-4	-4	-3	-3	-4	-3	-2	-3	3	-1	-3	-1	-2	-2	-2	-2	-2	-2
47	K	290-K	-1	-2	-4	-3	-3	-2	-4	-4	-2	-4	-1	-1	-3	-1	1	-1	5	1	0	4
48	E	291-S	-1	-2	-4	-4	-3	-3	-4	-4	-2	-3	2	-1	-3	0	1	-1	0	-1	3	5
49	A	292-A	3	4	-3	-3	-2	-2	-4	-4	4	-2	0	-1	-3	-2	-2	-3	-2	-3	-2	-2
50	V	293-V	-1	-4	1	0	4	0	4	-2	-4	-2	-3	-1	0	-4	-3	-3	-3	-3	-4	-3
51	Q	294-Q	-2	-3	-1	2	-1	0	-2	-3	-2	-3	-1	-2	-2	-1	6	-1	0	0	-2	0
52	T	295-G	-1	2	-1	1	-1	-1	-2	-3	-2	-2	0	3	-3	-1	-1	-2	2	-1	-2	-2
53	S	296-P	0	-2	-3	-3	-3	-2	-4	-4	4	-2	4	0	-3	-1	-1	-2	2	-1	-1	-1
54	A	297-P	4	-1	-3	-3	-2	-2	-4	-4	6	-2	0	-1	-3	-2	-2	-3	-2	-2	-2	-2
55	E	298-E	-1	-3	-1	-2	1	-2	-3	-4	-2	-3	-1	-1	-3	-1	1	-1	2	0	0	5
56	S	299-R	-1	-2	-2	1	-2	-1	-2	-3	-2	-2	3	-1	-2	-1	3	-1	0	3	-2	-1
57	D	300-D	-2	-2	-4	-4	-4	-2	-4	-4	-2	-4	-1	-1	-3	3	4	-1	0	-1	5	1
58	Q	301-L	-1	-2	-2	1	-2	-1	-2	-3	-2	-3	2	-1	-1	-1	4	5	-1	-1	-2	0
59	C	302-C	-1	-3	-2	-3	-2	-3	-3	-4	-3	9	-1	-2	-3	-1	-2	-3	-3	-3	4	-1
60	G	303-G	0	6	-5	-4	-4	-3	-4	-3	-3	-3	-1	-2	-4	-1	-3	-3	-2	-3	-2	-3
61	C	304-C	-1	-3	-2	-2	-2	-2	-3	-3	-3	9	-2	-2	-3	-2	-2	-2	-1	4	-3	-3
62	Y	305-Y	-2	-4	-2	-2	-2	-2	3	2	-4	-3	-2	-2	8	-3	-2	1	-3	-2	-4	-3
63	S	306-S	1	-1	-3	-3	-2	-2	-3	-4	-1	-2	5	1	-2	0	-1	-2	-1	-1	-1	-1
64	V	307-V	-1	-4	1	0	4	0	-2	-3	-3	7	-2	-1	-2	-4	-3	-4	-3	-4	-4	-4
65	S	308-S	0	-1	-2	-3	-2	-2	-3	-3	-2	-2	4	4	-2	0	-1	-2	-1	-2	-1	-1
66	S	309-S	1	-1	-3	-3	-2	-2	-3	-4	-1	-2	5	1	-2	0	-1	-2	-1	-1	-1	-1
67	V	310-V	-1	-4	2	0	5	0	-2	-4	-3	-1	-2	-1	-2	-4	-3	-4	-3	-3	-4	-3
68	L	311-L	-2	-4	1	5	0	2	0	-2	-4	-2	-3	-2	-2	-4	-3	-4	-3	-3	-4	-4
69	P	312-P	-1	-3	-4	-3	-3	-3	-4	-4	7	-4	-1	-2	-3	-2	-1	-2	3	-1	-2	-1
70	G	313-G	-1	5	-1	2	-2	-1	-2	-3	-3	-3	-2	-2	-3	-2	-3	-3	-3	-3	-3	-3
71	C	314-C	-1	-3	-2	-2	-2	-2	-3	-3	-3	9	-1	-1	-3	-2	-1	-3	3	-1	-3	-2
72	A	315-A	4	-1	-3	-2	-1	-2	-3	-3	-2	-2	0	-1	-2	-2	-1	-2	0	4	-2	-1
73	E	316-E	2	-2	-3	-3	-3	-2	-4	-4	-2	-3	0	-1	-3	-1	1	-1	0	-1	1	5
74	E	317-P	2	-2	-3	-3	-2	-2	-4	-4	4	-3	-1	-1	-3	-2	0	-1	0	3	-1	3
75	W	318-W	-3	-3	-3	-2	-4	-2	0	12	-4	-3	-4	-3	2	-5	-3	-3	-4	-3	-5	-4
76	N	319-N	-2	-2	-4	-4	-4	-3	-4	-4	-2	-4	0	-1	-3	6	0	0	0	-1	1	3
77	K	320-H	2	-2	-3	-3	-2	-2	-3	-3	-2	-2	2	-1	-1	-1	0	5	3	0	-2	-1
78	G	321-G	-1	5	-4	-4	-4	-3	-4	-3	-3	-4	-1	-2	-3	-1	-1	-2	0	4	-2	-2
79	E	322-K	2	-2	-3	-3	-2	-2	-3	-3	-2	-3	0	2	-3	-1	0	-1	2	0	-1	3
80	P	323-T	-1	-2	-3	-3	-2	-2	-4	-4	5	-3	2	2	-3	-1	-1	-2	2	-1	-2	-1
81	F	324-F	-3	-4	-1	-1	-2	-1	6	1	-4	-3	-3	-3	6	-3	-3	0	-3	-3	-4	-4
82	A	325-T	3	-1	-2	-3	-2	-2	-3	-4	-2	-2	0	2	-3	3	-1	-1	2	-1	-1	-1
83	C	326-C	-1	-3	-2	-2	-1	-2	-3	-3	-4	10	-2	-2	-3	-3	-4	-4	-4	-4	-4	-4
84	T	327-T	-1	-3	-2	-2	-1	-2	-3	-3	-2	-2	0	5	-2	-1	-1	-2	0	4	-2	-1
85	V	328-A	2	-3	2	0	5	0	-2	-4	-3	-1	-1	-1	-2	-3	-3	-3	-3	-3	-4	-3
86	T	329-A	2	-2	-2	-2	-1	-1	-3	-3	-2	-1	1	5	-2	-1	-1	-2	-1	-2	-2	-2
87	H	330-Y	-2	-3	-3	-3	-3	-2	0	-1	-3	-4	-2	-2	4	0	0	8	-2	-1	-2	-1
88	P	331-P	3	-2	-3	-3	-2	-3	-4	-4	7	-2	-1	-1	-3	-3	-2	-3	-2	-3	-2	-2
89	E	332-E	-1	3	-4	-4	-3	-3	-4	-3	-2	-4	-1	-2	-3	-1	0	-1	0	-1	0	5
90	G	333-S	0	4	2	-1	-1	-1	-2	-2	-2	-2	2	-1	-2	-1	-1	-2	-1	-2	-1	-2
91	D	--	-1	-1	-2	-2	-2	-1	-2	-2	-1	-2	0	-1	-1	0	0	-1	2	0	4	0
92	Q	334-K	-1	2	-4	-3	-3	-2	-4	-3	-2	-3	2	-1	-3	-1	4	-1	3	0	-1	0

FIG. 4 (continuation)

93	E	335-T	-1	-3	-1	-2	2	-2	-3	-4	-2	-3	-1	2	-3	-1	0	-2	-1	-2	3	3
94	I	336-P	-1	-3	3	-1	0	-1	-2	-4	4	-2	2	2	-2	-2	-2	-3	-2	-3	-2	-2
95	L	337-L	-2	-4	0	4	-1	1	-1	-3	5	-2	-2	-2	-2	-4	-3	-3	-3	-3	-3	-3
96	T	338-T	-1	-3	-1	-1	-1	-1	4	-1	-2	-2	0	5	0	-1	-2	-2	-2	-2	-2	-2
97	P	339-A	2	2	-3	-4	-3	-3	-4	-4	5	-3	-1	-2	-3	-1	-2	-2	-2	-3	3	-1
98	T	340-T	-1	-3	2	-1	0	-1	-3	-3	-2	-2	0	4	-2	-1	-1	-2	3	0	-2	-1
99	I	341-L	-2	-4	4	2	1	0	-1	-3	-4	7	-2	-2	-2	-4	-3	-4	-4	-4	-4	-4

FIG. 4 (continuation)

CD07696 - IgC_CH3: Third Ig constant domain (IgC) of Ig heavy chains.

P	C	Master	A	G	I	L	V	M	F	W	P	C	S	T	Y	N	Q	H	K	R	D	E
1	V	118-A	1	-4	2	-1	4	-1	2	-4	-4	-3	-2	1	-2	-4	-3	-4	-3	-4	-3	1
2	T	119-Y	-2	-3	-3	-3	-3	-3	2	-2	-3	-3	2	4	4	2	-2	-2	-3	-3	1	-2
3	L	120-L	0	-5	3	3	3	0	4	-3	-4	-3	-3	-2	-1	-5	-4	-4	-4	-4	-5	-4
4	L	121-S	1	-4	3	4	-1	0	-2	-4	2	-3	1	-2	-3	-4	-3	-4	-3	-4	-4	-4
5	P	122-R	-2	-4	-5	-5	-4	-4	-5	-5	8	-5	-2	-3	-5	-3	-3	-4	-2	1	-3	-3
6	P	123-P	-2	-3	-4	-4	-4	-4	-5	-5	8	-4	1	1	-4	-3	-3	-4	-2	-3	-3	-3
7	S	124-S	-1	-3	-3	-3	-3	-3	4	-3	2	-3	5	-1	-1	-2	2	-3	-2	-3	-2	-2
8	P	125-P	1	-4	1	-3	1	-3	2	-4	6	-4	-2	-2	-3	-3	1	-3	-2	-3	-3	1
9	K	126-F	0	-3	-3	0	-3	-3	2	-4	-3	-4	-2	-3	-2	-2	-1	-3	3	-1	4	3
10	Q	127-D	-2	-3	1	-3	-2	3	-4	-4	-3	-4	1	-2	-3	-1	5	-2	-2	-2	4	-1
11	L	128-L	0	-4	2	3	2	-1	4	-3	-4	-3	-3	-3	-1	-4	-3	-4	-4	-4	-4	0
12	L	129-F	-3	-5	-1	4	-1	0	4	-2	-4	-3	-4	-3	3	-4	-3	-3	-4	-4	-4	1
13	L	130-I	-2	-3	2	3	1	5	-1	-2	-2	-2	-2	-2	-2	-3	2	-2	-2	-2	-3	-2
14	T	131-R	-2	1	-4	-4	-3	-3	-4	-4	-3	-4	1	4	-4	2	-1	-2	-1	2	-1	3
15	K	132-K	0	1	-4	-3	-3	3	-4	-5	-3	-4	1	-2	-4	-2	-1	-3	5	-1	1	-1
16	S	133-S	1	-2	-3	-4	-3	-3	-4	-4	-3	-3	5	4	-3	-1	-2	-3	-2	-2	-2	1
17	V	134-P	3	-3	-1	-2	4	-2	-4	-5	4	-3	-2	1	-3	-4	-3	-4	-3	-4	-4	-3
18	K	135-T	-2	-3	0	-3	-2	-2	-4	-5	-3	-4	-1	4	-3	2	-1	-3	5	-1	-2	-2
19	L	136-I	-3	-5	2	5	3	0	-1	-4	-4	-3	-4	-2	-3	-5	-4	-5	-4	-4	-5	-5
20	T	137-T	-2	-3	-2	-3	-2	-2	-3	-3	-3	-3	0	7	2	-2	-2	-3	-2	-3	-3	-3
21	C	138-C	-2	-4	-3	-3	-2	-3	-4	-4	-5	11	-2	-2	-4	-4	-5	-5	-5	-5	-5	-5
22	L	139-L	-3	-5	4	4	3	0	-2	-4	-4	-3	-4	-2	-3	-4	-3	-4	-3	1	-5	-4
23	V	140-V	0	-4	4	-1	6	-1	-2	-5	-4	-2	-3	-2	-3	-5	-4	-5	-4	-4	-5	-4
24	T	141-V	-1	-3	-1	-2	2	-2	2	-4	-3	-3	3	4	-2	-2	-3	-3	-3	-3	-3	-3
25	N	142-D	-3	-3	-5	-5	-5	-4	-5	-6	-3	-5	-1	-2	-4	5	-1	-2	-1	1	5	1
26	L	143-L	2	-4	2	3	1	5	-2	-4	-4	-3	0	-2	-3	-4	-3	-4	-3	-3	-4	-4
27	T	144-A	1	2	-2	-2	-2	-2	-2	-2	3	-2	-1	3	-2	-1	-1	-2	-1	-2	-1	1
28	S	146-S	-1	-1	-2	-2	-2	-2	-2	-2	-1	-2	4	3	-2	-1	-1	-2	-1	-1	2	-1
29	P	147-K	-3	-4	1	0	-2	-2	-3	-3	5	-4	-2	1	2	-3	-2	-3	1	-3	-3	1
30	E	148-G	-2	1	-4	-1	-4	-3	-3	-3	-3	-4	0	-2	2	-2	-1	3	1	-2	3	3
31	G	149-T	-2	4	-5	-5	-4	-4	-5	-5	-3	-4	1	1	-4	2	-2	-3	-2	-3	4	2
32	V	150-V	-2	-5	2	2	5	0	2	-3	-4	-3	-4	-2	-2	-5	-4	-5	-4	-4	-5	-4
33	N	151-Q	-3	-2	-4	-4	-3	-3	-4	-5	-3	-4	-1	4	-4	7	2	-1	-2	-2	-1	-2
34	V	152-L	-2	-5	4	1	5	0	-2	-4	-4	-3	-4	-2	-3	-5	-4	-5	-4	-4	-5	-4
35	T	153-T	-1	-3	-3	-3	-2	-3	-4	-4	-3	-3	3	5	-3	-1	2	-3	-2	-2	1	-2
36	W	154-W	-4	-4	-4	-3	-5	-3	-1	12	-5	-4	-5	-4	1	-6	-4	-4	-5	-4	-6	-5
37	T	155-S	-1	-3	1	-3	-2	-2	-4	-4	-3	-3	3	4	-3	2	-2	-3	1	-2	-2	-2
38	R	156-R	-3	-4	-5	-4	-4	-3	-5	-4	-4	-5	-2	-3	-3	-2	4	-2	1	7	-3	-1
39	G	157-A	2	4	-4	-4	-4	-3	-5	-4	-3	-4	1	-2	-4	-2	2	-3	-2	-3	-2	3
40	S	158-S	-2	-2	-4	-4	-3	-3	-4	-4	-2	-3	3	-1	-3	4	2	-2	1	-1	4	1
41	G	159-G	-1	4	-3	-3	-2	-2	-3	-3	-2	-3	-1	2	-2	-1	-1	-2	2	2	2	-1
42	E	160-K	-2	-3	-2	0	1	-2	-3	-3	-2	-3	-1	1	-2	-2	0	-2	3	-1	-1	4
43	P	161-P	1	-3	-2	0	1	-2	-3	-4	4	-3	1	1	-3	-2	-2	-3	3	-2	-3	-2
44	V	162-V	1	-4	-1	2	3	-1	-3	-4	-3	-3	0	1	-3	-3	-2	3	-3	-3	-3	1
45	K	163-N	-2	-3	-2	1	1	3	-3	-4	-3	-3	1	2	-3	3	-2	-3	3	-2	-2	-2

FIG. 5

46	A	164-H	3	-3	-2	-3	2	-2	-4	-4	-3	-3	1	3	-3	-2	1	4	-2	-3	-3	-2
47	S	165-S	-2	-3	1	-2	1	-2	2	-3	-3	-3	3	-2	2	-2	-2	3	-2	1	-3	1
48	I	166-T	-2	2	3	-2	1	-2	-4	-4	-3	-3	-2	3	-4	-2	3	-3	-2	-3	1	-2
49	Q	167-R	-2	-3	-4	-4	-4	-3	-4	5	-3	-4	2	1	-2	2	5	3	-1	2	-2	-1
50	P	168-K	2	0	1	-3	-2	-3	-3	-3	4	-3	-2	1	2	2	-2	-2	1	-3	-3	-3
51	P	169-E	-2	-3	-4	-4	-3	-3	-5	-5	4	-4	1	3	-4	-2	-1	-3	3	3	-2	1
52	E	170-E	-2	0	1	-3	-3	-3	-4	-4	-3	-5	-2	-2	-4	-2	4	-2	2	-1	-1	5
53	S	171-K	-2	-3	-4	-4	-4	-3	2	-4	2	-4	3	-2	-2	-2	2	4	3	-1	1	-1
54	H	172-Q	3	-3	-4	-4	-3	-3	-4	-4	2	-4	-2	-2	-2	-2	2	7	-1	1	-2	2
55	N	173-R	-2	-2	-3	-3	-3	-2	-2	-2	2	-3	1	1	2	4	-1	-1	-1	2	2	-1
56	N	174-N	-1	-1	-2	-2	-2	-2	-2	-2	-2	-2	2	-1	-2	6	-1	-1	-1	-1	2	-1
57	A	175-G	3	3	-2	-2	-2	-2	-3	-2	-2	-2	-1	-1	-2	-2	-1	-2	-1	-2	-1	3
58	T	176-T	1	-2	-2	1	-1	-1	-2	-3	2	-2	0	5	-2	-2	-2	-3	-2	-2	-2	-2
59	L	177-L	-3	-4	-2	3	-2	-1	4	-3	-4	-4	-2	1	-2	-2	3	-3	-3	-3	3	-2
60	S	178-T	-1	-3	1	-3	-2	-2	2	-4	-3	-3	4	3	-2	-2	-2	-3	-2	1	-2	1
61	V	179-V	1	-4	3	-1	6	-1	-2	-5	-4	-2	-3	-2	-3	-5	-4	-5	-4	-4	-5	-4
62	T	180-T	-2	-4	1	-2	2	3	-3	-4	-3	-3	-2	4	-3	-2	2	-3	1	-2	-3	1
63	S	181-S	1	3	-4	-4	-3	-3	-4	-4	-3	-3	5	-1	-4	-1	-2	-3	-2	-3	-2	-2
64	T	182-T	-2	-4	3	-2	-1	-2	-4	-4	-3	-4	-1	4	-3	-2	-1	-3	2	-2	-2	3
65	L	183-L	3	-3	0	4	1	0	-2	-4	-4	-3	-2	-2	-3	-4	-3	-4	-3	-4	-5	-4
66	N	184-P	-2	-3	-4	-4	-3	-3	-5	-5	4	-4	1	3	-4	5	-2	-2	1	-2	-2	-2
67	V	185-V	-2	-5	5	0	5	-1	-2	-4	-4	-3	-4	-2	-3	-5	-4	-5	-4	-4	-5	-4
68	S	186-G	-1	1	-3	-3	1	-3	-4	-4	-3	7	4	-1	-4	-2	-3	-3	-3	-3	-3	-3
69	T	187-T	2	-3	-3	-3	-2	-3	-4	-4	2	-3	-1	5	-4	-2	-2	-3	1	-2	-2	1
70	D	188-R	-3	1	-5	-4	-4	-3	-5	-5	-3	-5	-2	-3	-3	-1	3	3	2	3	3	2
71	D	189-D	0	-3	-5	-5	-4	-4	-5	-5	-3	-5	-2	-2	-4	-1	2	-2	-2	-2	6	3
72	W	190-W	-4	-4	-4	-3	-5	-3	-1	12	-5	-4	-5	-4	1	-6	-4	-4	-5	-4	-6	-5
73	I	191-I	-3	-4	4	2	-1	-1	-3	-5	-4	-4	-2	-2	-3	3	-2	-3	-3	-3	4	1
74	S	192-E	1	1	-4	-4	-3	-3	-4	-5	-3	-3	4	-1	-4	2	-1	-2	-2	-2	-1	3
75	G	193-G	-2	6	-5	-5	-5	-4	-5	-4	-4	-4	-2	-3	-3	-2	-3	4	1	-3	-3	-3
76	D	194-E	-3	-3	-5	-4	-4	-4	-3	-4	-3	-5	-2	-3	2	-1	-1	4	3	-1	4	3
77	T	195-T	-2	1	-4	-4	-3	-3	-3	-3	-3	-4	-1	3	2	-2	2	-2	2	2	1	2
78	F	196-Y	-4	-5	-2	-1	-3	-2	8	0	-5	-4	-4	-4	5	-5	-4	-2	-5	-4	-5	-5
79	Q	197-Q	-2	-3	-3	-4	1	-3	-4	-4	-3	-4	1	1	-3	3	4	-2	2	-1	-1	3
80	C	198-C	-2	-4	-3	-3	-2	-3	-4	-4	-5	11	-2	-2	-4	-4	-5	-5	-5	-5	-5	-5
81	K	199-R	0	-4	4	-2	1	-2	-3	-4	-3	-3	-2	1	-3	-3	-2	-3	4	2	-3	-2
82	V	200-V	1	-4	1	-1	6	-1	-3	-5	-4	-2	-3	-2	-3	-4	-4	-5	-4	-4	-5	-4
83	T	201-T	-2	-3	-2	0	1	-2	-3	-5	-3	-3	2	4	-3	3	-2	-3	-2	-3	2	-2
84	H	202-H	-3	-4	-4	0	-4	-2	-3	-4	-4	-4	-3	-3	0	-1	-1	10	-2	-2	-3	-2
85	P	203-P	0	-3	-3	-4	0	-3	-5	-5	6	-4	-2	1	-4	-3	3	-3	-2	-2	-3	1
86	D	204-H	-3	-3	-4	0	-4	-3	-4	-5	-3	-5	-2	-3	-3	-1	-2	4	-2	-3	7	0
87	L	205-L	-3	-5	0	5	-1	6	2	-3	-5	-3	-4	-3	-2	-5	-3	-4	-4	-4	-5	-4
88	P	206-P	1	-3	-3	-1	-3	-3	-4	-5	7	-4	0	-2	-4	-3	-3	-4	-3	-4	-3	-3
89	T	207-R	-2	-3	-3	-4	-3	-3	-4	-4	-3	-3	2	4	-3	-2	-1	-3	3	2	-2	1
90	P	208-A	2	-3	-4	-4	-3	-3	-5	-5	7	-4	-2	-2	-4	-3	2	-3	-2	-3	-3	-2
91	L	209-L	-3	-5	4	3	2	-1	-2	-4	-4	-3	-3	-2	-3	-4	-2	-4	2	2	-4	-3
92	R	210-M	-2	-4	2	-2	3	3	-3	-4	-4	-4	-2	1	-3	-3	-1	-3	1	4	-3	1

FIG. 5 (continuation)

93	K	211-R	-2	0	-4	0	-3	-2	-4	-4	-3	-4	1	-2	-3	-2	2	-2	3	3	-2	2
94	S	212-S	-1	-3	-4	-4	-3	-3	-4	-4	-3	-3	5	4	-3	-1	-2	-3	1	1	-2	-2
95	I	213-T	-3	-5	6	3	1	0	-2	-4	-4	-3	-3	1	-3	-4	-4	-5	-4	-4	-5	-5
96	Q	214-T	-2	-3	-3	-3	0	-2	-4	-4	-3	-3	2	3	-3	2	4	3	2	-1	-2	-1

FIG. 5 (continuation)

CD00099 - IgV: Immunoglobulin variable domain (IgV).

P	C	Master	A	G	I	L	V	M	F	W	P	C	S	T	Y	N	Q	H	K	R	D	E
1	L	10-V	-1	0	2	3	4	3	-6	-8	-4	-7	-3	-4	-3	-7	-2	-7	-1	-3	-5	-1
2	S	11-S	-2	-7	1	2	1	-1	-4	1	-4	-2	4	4	-7	-6	-3	-3	-4	-4	-4	-2
3	V	12-V	1	-4	2	-2	5	0	-3	-8	-7	0	-2	1	-3	-7	-6	-7	3	1	-7	-5
4	S	13-S	0	-2	-7	-2	-1	-2	-3	-2	0	-7	2	0	-7	-1	4	-1	3	0	-6	2
5	E	14-E	0	-3	-1	-3	1	-6	-7	-8	4	-7	-1	2	-1	-6	0	-1	-3	1	-6	4
6	G	15-G	-5	7	-9	-9	-8	-8	-4	-8	-7	-8	1	-4	-8	1	-3	-1	-2	-1	0	-2
7	E	16-A	1	0	-8	-6	-3	-2	-8	-8	-7	-8	-1	-1	-7	0	4	0	2	-1	-1	4
8	S	17-L	0	-3	-1	-5	-2	-2	-4	-8	0	-7	4	4	-7	2	-1	-6	-1	-1	-2	-3
9	V	18-V	3	-7	0	2	5	-5	-4	-8	-7	-2	-1	-1	-7	-7	-7	-3	-4	-7	-8	-7
10	T	19-L	-3	-7	-1	-3	-2	-2	-4	-8	-7	-7	2	5	-2	2	0	-1	2	0	-2	0
11	L	20-L	-7	-9	5	5	-2	3	2	-7	-8	-7	-8	-4	-6	-9	-8	-8	-8	-8	-9	-8
12	S	21-R	-5	-7	-5	-4	-1	-6	-7	-8	0	-7	4	4	-2	3	-1	2	-3	2	-4	0
13	C	22-C	-6	-8	-7	-7	-6	-7	-8	-8	-8	12	-6	-6	-8	-8	-8	-8	-9	-9	-9	-9
14	T	23-N	-3	-5	0	-6	0	-6	-7	-8	-3	-7	2	3	-2	3	-1	-1	3	0	-1	2
15	Y	24-Y	1	-2	-2	0	2	0	1	-6	-1	-7	-2	-2	6	-7	1	4	-7	-5	-8	-7
16	S	25-S	-3	-5	-2	-5	-7	-7	-2	-8	-3	-7	5	3	-3	-1	-2	-2	2	0	0	1
17	G	26-S	-4	3	0	-3	-2	1	-7	-8	-3	-7	3	3	1	2	-1	-6	-1	-6	2	-6
18	S	27-S	-1	-1	1	-1	-1	-4	3	1	0	-4	3	1	0	1	-2	-4	-2	-2	-1	-2
19	F	--	-1	-1	1	0	0	-1	4	2	-1	0	2	0	1	-2	-1	0	-2	-3	-1	-2
20	S	28-V	-1	2	-1	-5	-3	-3	-8	-2	-1	-7	3	3	-7	2	0	-2	1	-2	1	0
21	S	29-P	-2	-1	-3	-2	-2	-1	1	-7	-4	-7	3	1	3	3	-3	5	-2	-6	0	1
22	Y	30-P	-1	-2	-4	-4	-1	-2	1	-6	1	-7	0	2	5	3	-4	-2	-3	-3	2	-1
23	Y	31-Y	0	0	-2	-7	-3	0	0	4	-2	-1	1	0	6	2	-2	0	-4	2	-1	-4
24	I	32-L	-1	-8	5	2	3	6	1	-1	-8	-6	-7	-1	-6	-8	-7	-8	-7	-8	-9	-8
25	F	33-F	-1	-1	-7	-3	-5	-6	5	-6	-5	0	1	-2	5	2	1	6	-6	-4	-4	-4
26	W	34-W	-8	-8	-7	-2	-2	-6	-4	13	-9	-8	-8	-8	-3	-9	-7	-8	-8	-8	-10	-8
27	Y	35-Y	-7	-8	-1	-1	1	-6	1	-4	-8	-8	-7	-7	9	-3	-7	2	-7	-7	-8	-7
28	R	36-V	-3	-7	-7	-3	-2	-6	-8	-1	-7	-8	-6	-6	-7	-6	5	-2	3	7	-7	-5
29	Q	37-Q	-6	-1	-4	-3	-7	-6	-8	3	-7	-8	-5	-2	-7	-3	8	-1	-1	0	-6	-1
30	K	38-Y	1	-7	-3	0	-5	-1	-2	-7	0	-7	-1	0	2	0	0	-1	3	3	1	-3
31	P	39-P	-1	-4	-7	-1	-7	-3	-4	-8	7	-8	-1	-2	-8	-3	2	-3	-3	-3	1	-2
32	G	40-N	-1	6	-6	-4	-5	-5	-6	-1	-4	-5	1	-3	-5	2	-2	-2	-3	-1	0	-1
33	K	41-Q	-4	0	-7	-5	-6	-2	-7	-7	-3	-7	0	-2	-3	1	5	2	4	1	-2	1
34	G	42-G	3	4	-8	-6	-5	-2	-8	-8	-1	-7	2	0	-3	0	3	2	-4	-3	-2	0
35	P	43-L	-2	-8	-1	4	-3	3	1	1	6	-7	-7	-3	-3	-4	-7	-2	-4	-1	-7	-2
36	E	44-Q	-5	-7	-2	-4	1	-2	-8	-8	-7	-8	-4	2	-7	-2	4	-2	2	0	-5	5
37	L	45-L	-5	-4	-3	3	-3	1	5	7	-8	-7	-1	-3	2	-4	-4	0	-3	0	-8	0
38	L	46-L	-6	-9	4	5	3	2	-3	-7	-8	-7	-8	-6	-6	-9	-3	-8	-8	-8	-9	-8
39	I	47-L	1	1	4	2	0	0	2	-7	-8	4	-1	-1	1	-7	-7	-1	-7	-3	-8	-8
40	Y	48-K	-1	-2	-1	-1	-1	-6	2	1	-7	-7	0	0	7	0	-2	1	-1	1	-7	-2
41	I	49-Y	-1	0	3	0	2	1	1	2	-7	-7	1	-1	3	-4	1	-2	-4	-1	-5	-1
42	S	50-T	0	-2	-1	-1	0	-5	-2	0	-4	-6	2	1	3	3	0	-5	0	-1	0	-3
43	S	51-S	-2	1	-1	-2	-3	-2	-1	0	1	-6	4	1	1	2	-2	-6	0	-3	0	-1
44	N	52-A	-1	2	-2	-3	0	-1	-2	-1	-1	-6	1	1	-3	3	-1	-1	1	-1	2	-2
45	G	53-A	-1	3	-3	-1	-2	1	-2	-5	-1	-5	2	0	1	0	-1	-1	0	1	-2	0

FIG. 6

46	S	54-T	-1	2	-3	-1	-1	-1	-2	-3	-1	-3	2	1	-1	0	0	-1	2	-3	0	0
47	Q	--	0	-1	-1	-1	-1	-1	-1	2	0	-3	0	1	0	2	2	2	1	0	-2	0
48	Y	--	-2	-2	-2	-1	1	-2	0	-2	-1	-3	0	0	4	-2	1	1	-2	0	1	2
49	A	55-L	1	1	-3	-2	-5	-2	-5	-5	0	-5	1	1	0	2	1	0	2	-1	0	0
50	G	56-V	-3	2	-3	-4	-1	-6	0	-6	2	-6	2	0	-6	1	0	-1	-1	-2	2	1
51	G	57-K	0	3	-4	-1	-1	-1	1	-7	-4	-6	1	0	-3	0	-3	-6	3	-5	1	1
52	V	58-G	-1	-2	1	-1	2	-5	2	-6	0	-6	-1	2	0	-4	0	1	1	-4	0	1
53	K	59-I	-2	-1	-3	-4	-6	-6	-3	-7	4	-7	1	-2	-2	1	0	0	4	-3	1	0
54	G	60-N	-1	4	-3	-8	-5	-7	0	-1	-1	-8	2	-4	-7	1	-2	-6	0	-2	3	1
55	R	61-G	-6	0	-3	-7	-8	0	-1	-1	-7	-9	-4	-6	-7	1	-4	-2	3	7	-2	-2
56	F	62-F	0	-8	1	0	-1	-5	7	-5	-8	4	-3	-7	5	-8	-4	-6	-4	-8	-8	-5
57	S	63-E	-2	-7	0	-1	-2	-6	-4	-8	-3	-7	4	4	-1	-4	1	-2	-3	-2	0	1
58	G	64-A	3	3	3	-1	1	-1	-1	-1	-2	-7	0	0	-7	-7	-7	-7	-7	-3	-7	-2
59	T	65-E	-3	-3	0	-7	-1	-6	-2	-8	-7	-7	4	4	-7	1	-1	-2	0	0	0	1
60	R	66-F	-2	0	-1	1	-1	-5	2	-7	-4	0	-1	-1	-2	-3	0	2	2	4	-7	-2
61	D	67-K	-3	-2	0	-4	-1	-3	-3	-8	1	-8	2	-2	-3	3	-1	0	-1	-2	4	0
62	S	68-K	0	1	-4	-5	-4	-3	-4	-8	-1	-7	2	1	-7	2	-2	-6	3	1	2	0
63	S	69-S	1	1	-4	-2	-4	-4	-4	-4	-2	-4	3	1	0	2	-1	-3	0	-1	1	0
64	K	70-E	0	-1	-1	-2	-2	0	-2	-2	-2	-2	2	1	-2	1	0	-1	3	1	-1	0
65	S	71-T	-4	-3	-4	-4	-6	-1	-4	-1	-6	-7	3	3	-2	3	-3	1	3	0	0	0
66	S	72-S	-4	-5	-2	-3	-3	-1	-3	-7	-4	-7	3	3	1	-2	3	0	0	1	2	-2
67	F	73-F	1	-2	-1	-1	3	-5	5	1	-4	3	3	-3	0	-4	-7	-7	-7	-4	-7	-7
68	T	74-H	-2	-7	1	-2	-2	-6	0	-6	-2	-7	3	4	4	0	-4	3	-4	-1	-4	-3
69	L	75-L	-7	-9	-2	7	-3	3	1	-7	-8	-7	-8	-7	-6	-9	-7	-8	-8	-8	-9	-8
70	T	76-T	-5	-7	0	-2	-4	-3	-7	-1	-7	-7	0	5	1	1	3	4	1	0	-3	0
71	I	77-K	-4	-6	7	3	2	2	-5	-8	-8	-7	-7	-6	-7	-8	-8	-8	-2	-8	-9	-8
72	S	78-P	-1	-4	-7	-3	-5	-2	-4	2	-2	-7	3	2	-3	3	3	0	1	1	-3	0
73	S	79-S	1	1	-8	-6	-7	-7	-2	-8	0	-7	3	-4	-3	4	-1	0	-3	1	3	-2
74	L	80-A	2	-7	0	5	3	3	-6	-8	-2	-6	0	-2	-7	-8	-7	-8	-7	-7	-8	-7
75	Q	81-H	-4	-2	-4	-3	-7	-6	-8	-8	-7	-8	0	2	-1	-5	5	1	2	3	-2	2
76	P	82-M	1	-2	-5	2	2	-1	-3	-7	5	-7	-1	0	-7	-4	-4	-7	1	-1	-3	-1
77	E	83-S	-2	0	-7	-6	-3	-1	-8	-8	-6	-7	1	-2	-7	2	1	-5	-1	0	-1	6
78	D	84-D	-4	-7	-9	-9	-8	-8	-9	-9	-4	-9	-5	-6	-8	-4	2	-1	-6	-7	8	-4
79	S	85-A	0	-2	-4	-5	-3	-3	-2	-8	-6	-6	5	4	-7	-2	-3	-6	-4	-6	-4	1
80	A	86-A	6	4	-7	-8	-6	-7	-8	-8	-7	-6	-1	-6	-8	-6	-6	-7	-6	-3	0	-6
81	V	87-E	-5	-8	3	0	5	1	-1	-8	-7	-7	-2	4	-6	-4	-2	-8	-3	-7	0	-4
82	Y	88-Y	-7	-9	-7	-4	-6	-6	-2	-3	-8	-8	-7	-7	10	-8	-7	-3	-7	-7	-9	-7
83	Y	89-F	-7	-9	0	2	-6	-5	6	-4	-8	-8	-7	-7	8	-8	-7	-1	-3	-4	-9	-8
84	C	90-C	-6	-8	-7	-7	-6	-7	-8	-8	-8	12	-6	-6	-8	-8	-8	-8	-9	-9	-9	-9
85	A	91-A	6	-2	-3	-3	-2	1	-7	-8	-6	-1	-1	-1	-7	-7	1	-7	1	-6	-7	-6
86	V	92-V	2	-2	1	1	3	-5	-2	-7	-7	1	2	0	1	-6	-1	-7	-2	2	-7	-3
87	S	93-S	-4	1	-1	-3	0	-1	-1	7	-4	-7	4	0	2	-6	-4	-3	-6	0	0	1
88	L	94-E	-1	1	-4	1	-1	-1	-2	0	0	-7	0	-1	1	-1	-2	1	-4	0	2	1
89	S	95-S	1	2	-7	-2	-3	3	-2	0	0	-2	3	0	0	-1	0	-3	-1	2	-2	-2
90	G	96-P	-3	3	-2	-2	-3	-6	-1	2	2	-2	2	-4	3	0	-1	1	-6	1	0	-4
91	G	97-F	-1	2	-3	-3	-3	-1	-2	0	3	-4	2	0	1	2	-1	-4	-1	0	1	-3
92	T	99-N	-1	2	-3	-1	-2	-3	-2	0	0	-1	1	2	4	1	0	-1	0	-2	-1	-3

FIG. 6 (continuation)

93	Y	100-E	0	0	-1	-1	-1	-4	1	3	-5	-5	-2	-1	5	1	1	-2	-3	-2	3	-1
94	K	101-K	-4	-3	0	-6	-1	-2	2	1	1	-8	-3	1	2	-3	-1	0	5	-1	-1	2
95	L	102-L	1	-4	0	3	-1	0	-2	0	-1	-7	-6	-6	3	1	1	-6	2	-4	1	-2
96	Y	103-T	-6	-2	3	-1	3	-5	1	-6	-8	-7	-6	3	6	-3	-4	-1	-2	-5	-7	1
97	F	104-F	-7	-4	-3	-3	-6	-1	9	7	-9	-2	-4	-7	-3	-8	-8	-7	-1	-7	-8	-3
98	G	105-G	-2	8	-9	-9	-8	-8	-9	-8	-7	-8	0	-7	-8	-6	-7	-7	-7	-8	-7	-7
99	Q	106-T	1	1	-8	-5	-7	-7	-8	-8	2	-7	2	0	-7	2	4	-3	2	-1	-1	-1
100	G	107-G	-3	8	-9	-9	-9	-8	-9	-8	-8	-8	-5	-7	-9	-6	-7	-7	-7	-8	-7	-7
101	T	108-T	-1	-7	-6	-4	-5	-6	-7	-8	-6	-6	-1	8	-7	-5	-6	-7	-6	-6	-6	-6
102	R	109-R	-3	-7	0	-2	-1	1	-8	-8	-5	-8	-2	-2	-7	-3	3	-6	4	6	-6	0
103	L	110-L	-6	-9	1	6	3	-3	-5	-7	-8	-6	-8	-3	-6	-9	-8	-8	-8	-8	-9	-8
104	T	111-T	-4	-7	3	-3	1	-6	-4	-8	-7	-7	1	6	0	-4	-2	2	-3	-7	-3	-1
105	V	112-I	-6	-9	2	-3	8	-4	-3	-8	-8	-6	-3	-5	-6	-8	-8	-9	-8	-8	-9	-8

FIG. 6 (continuation)

CD04981 - IgV_H: Immunoglobulin (Ig) heavy chain (H), variable (V) domain.

P	C	Master	A	G	I	L	V	M	F	W	P	C	S	T	Y	N	Q	H	K	R	D	E
1	L	5-L	0	-5	0	5	-1	0	-1	-3	-4	-3	-3	0	-3	-5	-4	-4	-4	-4	-5	-4
2	V	6-E	-2	-4	-1	2	3	-1	-3	-4	-4	-3	0	-2	-3	-3	2	-3	2	-2	0	0
3	E	7-Q	-3	-4	-5	-4	-4	-3	-5	-4	-3	-5	-2	-2	-3	-2	6	-1	0	-1	-1	5
4	S	8-S	-1	-2	-4	-4	-4	-3	-4	-5	3	-3	6	-1	-4	-1	-2	-3	-2	-3	-2	1
5	G	9-G	0	6	0	-4	-3	-4	-4	-4	-4	-4	0	-3	-4	-2	-3	-4	-3	-4	-3	0
6	P	10-A	2	2	-4	-4	-3	-4	-5	-5	6	-3	2	-2	-4	-3	-3	-3	-3	-3	-3	-3
7	G	11-E	-2	5	-3	-3	1	2	-4	-4	-4	-4	-2	-3	-4	-2	-2	-3	-2	-3	-2	3
8	L	12-V	-2	-5	0	4	3	0	-2	-4	-4	-3	0	2	-3	-4	-3	-4	-4	-4	-4	-4
9	V	13-K	-2	-4	2	-1	5	2	-3	-5	-4	-3	-3	-2	-3	-4	-3	-4	3	-2	-4	-3
10	K	14-K	-2	-3	-4	-1	-4	-2	-4	-4	-3	-4	0	-2	-3	-2	3	-2	6	2	-2	-1
11	P	15-P	0	-4	-4	-5	-4	-4	-5	-5	8	-4	-2	-3	-5	-4	-3	-4	-3	-4	-3	-3
12	G	16-G	-1	6	-5	-5	-4	-4	-5	-4	-3	-4	3	-2	-4	-2	-3	-3	-3	-3	1	-3
13	Q	17-S	1	0	-4	-4	-4	-3	-5	-4	-3	-4	0	-2	-3	-2	5	-2	-1	1	-2	4
14	S	18-S	1	-2	-3	-4	-3	-3	-4	-5	2	-3	5	4	-3	-1	-2	-3	-2	-3	-2	-2
15	L	19-V	-3	-5	0	5	3	0	-2	-4	-4	-3	-4	-3	-2	-5	-4	2	-4	-4	-5	-4
16	K	20-Q	-2	-3	-4	-4	-3	-3	-4	-4	-3	-4	3	2	-3	-2	1	-2	4	3	-2	-1
17	L	21-V	-3	-5	3	5	2	0	-1	-4	-5	-3	-4	-3	-3	-5	-4	-5	-4	-4	-5	-5
18	S	22-S	-1	-3	-3	-3	0	-3	-4	-4	-3	-3	5	5	-3	-1	-2	-3	-2	-3	-2	-2
19	C	23-C	-2	-4	-3	-3	-2	-3	-4	-4	-5	11	-3	-3	-4	-4	-5	-5	-5	-5	-5	-6
20	K	24-K	1	-3	-3	-4	-3	-3	-4	-5	-3	-3	0	4	-3	-2	-1	-3	5	1	-3	-2
21	A	25-A	4	2	1	-2	4	-2	-3	-4	-3	-2	0	-2	-3	-3	-3	-4	-3	-4	-4	-3
22	S	26-S	-1	-2	-4	-4	-3	-3	-4	-5	-3	-3	6	2	-3	-1	-2	-3	-2	-2	-2	-2
23	G	27-G	-2	7	-5	-5	-5	-4	-4	-3	-4	-4	-2	-3	2	-2	-4	-3	-3	-4	-3	-4
24	F	28-G	-3	-1	-2	0	0	-2	6	-1	-5	-4	0	-3	5	-3	-4	-2	-4	-4	0	-4
25	T	29-T	-2	-3	0	-3	0	-2	-4	-4	-3	-3	3	5	-3	1	-2	-3	-2	-3	-2	1
26	F	30-F	-3	-4	-1	2	-2	-1	7	-1	-4	-3	0	-3	3	-4	-4	-3	-4	-3	-4	-4
27	T	31-S	-1	0	-3	-4	-3	-3	-4	-4	-3	-3	4	5	-4	1	-2	-3	0	-2	-2	-2
28	S	32-M	-1	0	-4	-4	-3	2	-4	-5	-3	-3	5	-1	-4	4	1	-2	-2	-2	-2	-2
29	Y	33-Y	0	-4	-3	-3	-3	-3	0	0	-4	-4	0	0	8	1	-3	3	-3	-3	-4	-3
30	G	34-G	1	3	-4	-4	-3	-3	-2	8	-4	3	0	1	2	-3	-3	-3	-3	-4	0	-3
31	V	35-F	-2	-5	3	0	4	6	1	4	-4	-3	-3	1	-2	-4	-3	-4	-4	-4	-5	-4
32	N	36-N	2	3	-4	-4	-4	-3	-2	-3	-4	-4	0	-2	3	4	-2	4	-3	-3	-2	-3
33	W	37-W	-4	-4	-4	-3	-5	-3	-1	12	-6	-4	-5	-4	1	-6	-4	-4	-5	-4	-6	-5
34	V	38-V	-2	-5	5	0	5	-1	1	-4	-4	-3	-4	-2	-2	-5	-4	-5	-4	-5	-5	-5
35	R	39-R	0	-4	-5	-4	-4	-3	-5	-5	-4	-5	-2	-3	-3	-2	-1	-2	3	7	-3	-2
36	Q	40-Q	-3	-4	-5	-4	-4	-2	-5	-4	-3	-5	-2	-2	-3	-2	8	-1	2	2	-2	0
37	A	41-A	4	-2	-3	0	-2	-2	-4	-4	4	-3	0	-2	-4	-3	-2	-3	-2	1	-3	-3
38	P	42-P	3	-3	-4	-4	-3	-4	-5	-5	7	-4	-2	-2	-4	-3	1	-3	-2	-3	-3	-2
39	G	43-G	-2	6	-5	-5	-5	-4	-5	-4	-4	-4	0	-3	-5	-2	-2	-3	-2	1	-3	1
40	K	44-H	-3	-3	-5	-4	-4	-3	-5	-5	-3	-5	-2	-2	-3	-2	2	4	6	2	-2	-1
41	G	45-G	0	6	-5	-5	-4	-4	-5	-5	-4	-4	-2	0	-5	-2	-3	-4	-3	-4	2	-3
42	L	46-L	-3	-5	-1	5	-1	0	-1	5	1	-3	-4	-3	-2	-5	-3	-4	-3	0	-5	-4
43	E	47-E	-3	-4	-5	-5	-4	-4	-5	-5	-3	-5	-2	-2	-4	2	0	-2	1	-2	0	7
44	W	48-W	-4	0	-4	-3	-4	3	-1	12	-5	-4	-4	-4	0	-5	-4	-4	-4	-4	-6	-5
45	I	49-M	-3	-5	5	2	4	5	-2	-4	-4	-3	-4	-2	-3	-5	-4	-5	-4	-4	-5	-4

FIG. 7

46	G	50-G	2	6	-4	-4	-1	-4	-5	-4	-4	-4	-2	-3	-4	-2	-3	-4	-2	1	-3	-3
47	F	51-G	0	1	0	0	0	-2	3	6	-4	-4	-2	0	-2	2	-3	-3	-3	0	-3	-1
48	I	52-I	-3	-5	5	1	3	2	-2	-3	-4	-3	-3	0	2	-4	-3	-4	-3	0	-5	-4
49	N	53-I	-3	-3	0	-1	0	-2	-2	6	-4	-4	1	0	4	3	-2	-2	-2	1	0	-3
50	P	54-P	-2	0	1	-3	-3	-3	-3	-3	4	-4	1	2	3	1	-3	-2	-3	-3	2	-2
51	G	55-I	-2	4	0	-4	0	-3	-4	-5	-4	-4	0	-2	-4	2	-2	-3	-2	0	4	-2
52	G	56-F	0	4	-4	-4	-4	-4	0	-4	-3	-4	2	0	-3	1	-2	-3	-2	-3	1	0
53	G	57-G	-1	4	-3	-4	0	-3	-4	-4	-3	-3	3	1	-4	-2	-2	-3	0	-3	-2	0
54	E	58-T	-2	0	0	-3	-3	-3	-3	4	-3	-4	1	2	1	1	2	-2	1	1	1	2
55	T	59-S	0	-3	0	-3	-2	-2	-2	-2	-3	-3	2	4	5	-2	1	-2	-2	-3	-3	-2
56	Y	60-N	-1	0	-2	-2	0	-2	-1	-1	-2	-2	1	1	4	2	-1	3	-2	-2	1	-1
57	Y	61-Y	-3	-4	1	-2	-2	-2	2	0	-4	-3	0	-2	8	-3	-3	0	-3	-3	-4	-3
58	A	62-A	5	-2	-3	-4	-2	-3	-4	-5	-3	-3	1	1	-4	3	-2	-3	1	-2	-2	-2
59	D	63-Q	-3	-3	-4	-4	-1	-3	-5	-5	3	-5	0	-2	-4	-1	4	-2	-2	-2	5	3
60	S	64-K	2	-2	-4	-4	-3	-3	-4	-5	-3	-3	4	1	-3	-2	-1	-3	4	-1	-2	0
61	V	65-F	-3	-5	0	3	4	0	5	-3	-5	-3	-4	-3	-1	-5	-4	-4	-4	-4	-5	-5
62	K	66-R	-2	-3	-4	-4	-4	-3	-5	-5	-3	-5	-2	0	-3	-2	3	-2	6	2	-2	-1
63	G	67-G	-1	5	-5	-5	-4	-4	-5	-5	-3	-4	3	-2	-4	4	-3	-3	-3	-3	-2	-3
64	R	68-R	-3	-4	-5	-4	-4	-3	-5	-5	-4	-5	-2	-3	-3	-2	0	-2	3	7	-3	-2
65	F	69-V	3	-4	2	-1	2	-1	6	-2	-4	-3	-3	-2	0	-4	-4	-4	-4	-4	-5	-4
66	T	70-T	0	-3	-3	-3	-2	-3	-4	-4	-3	-3	2	6	-3	-2	-2	-3	-2	-3	-2	0
67	I	71-F	-3	-5	6	2	0	2	4	-3	-5	-3	-4	-3	-2	-5	-4	-4	-4	-5	-5	-5
68	T	72-T	-1	-3	-3	-3	-2	-3	-4	-4	-3	-3	4	5	-3	1	-2	-3	-2	-3	-2	-2
69	R	73-A	1	-4	-2	-1	1	-2	-4	-4	1	-4	-2	-2	-3	-3	1	-3	1	5	-3	1
70	D	74-D	-3	-3	-1	-5	-4	-4	-5	-6	-3	-5	-2	-3	-5	3	-2	-3	-2	-3	7	1
71	T	75-Q	2	-2	-4	-4	-3	-3	-4	-5	-3	-3	1	4	-4	4	1	-2	-2	-2	2	-2
72	S	76-A	2	2	-4	-4	-3	-3	-4	-5	-3	-3	5	-1	-4	1	-2	-3	-2	-3	-2	-2
73	K	77-T	0	-3	-1	-4	-3	-3	-4	-5	-3	-4	2	0	-3	1	1	-2	5	0	-2	-1
74	S	78-S	-1	-2	-5	-5	-4	-3	-4	-5	-3	-3	5	-1	-4	5	-2	-2	1	-2	-1	-2
75	T	79-T	-2	-3	-3	-3	-2	2	-4	-4	-3	-3	-1	5	-3	-2	5	-2	-1	-2	-3	-1
76	V	80-A	3	-3	0	1	4	-1	1	-4	-4	-2	-2	-2	-2	-4	-3	-4	-3	-4	-4	-4
77	Y	81-Y	-3	-4	-3	-3	-3	-3	4	0	-4	-4	0	0	8	-3	-3	0	-3	-3	-4	-4
78	L	82-M	-3	-5	0	5	-1	5	-1	-3	-5	-3	-4	-3	-3	-5	-3	-4	-4	-4	-5	-4
79	Q	83-E	-2	-4	-4	-4	-4	-3	-5	-4	-3	-5	-2	1	-3	-2	6	-2	1	1	1	2
80	L	84-L	0	0	1	4	-1	6	-2	-3	-4	-3	-3	-3	-3	-4	-3	-4	-4	-4	-5	-4
81	N	85-T	-2	-2	-5	-5	-4	-3	-5	-5	-3	-4	3	1	-4	6	2	-2	-2	-2	1	1
82	S	86-N	-1	1	-5	-5	-4	-3	-5	-5	1	-3	5	-1	-4	4	1	-2	-2	-2	-1	-2
83	L	87-L	-3	-5	0	5	2	3	-1	-4	-5	-3	-4	-3	-3	-5	-4	-5	-4	-4	-5	-5
84	T	88-R	-2	-3	-4	-4	-3	-3	-4	-4	-3	-4	-1	4	-3	-2	2	-2	4	3	-2	1
85	P	89-S	0	0	-2	-3	2	-3	-4	-5	5	-3	2	2	-4	1	-3	-3	-3	-3	-3	-3
86	E	90-D	-3	-4	-5	-5	-4	-4	-5	-5	-3	-5	-2	-3	-4	-2	0	-2	1	-2	2	7
87	D	91-D	0	-3	-5	-5	-5	-5	-5	-6	-3	-5	-2	-3	-5	0	-2	-3	-2	-3	7	0
88	T	92-T	-1	-3	-3	-3	-2	-3	-4	-4	-3	-3	3	6	-3	-1	-2	-3	-2	-3	-2	-2
89	A	93-A	6	0	-3	-3	-2	-3	-4	-4	-3	-2	0	-2	-4	-3	-3	-3	-2	-3	-3	-3
90	V	94-V	-2	-4	1	-1	6	3	-3	-4	-4	-3	0	2	-3	-4	-3	-4	-4	-4	-4	-4
91	Y	95-Y	-4	-5	-3	-3	-3	-3	2	1	-5	-4	-3	-3	9	-4	-3	0	-4	-3	-5	-4
92	Y	96-Y	-4	-5	-3	-3	-3	-3	4	1	-5	-4	-4	-3	9	-4	-3	0	-4	-4	-5	-4

FIG. 7 (continuation)

93	C	97-C	-2	-4	-3	-3	-2	-3	-4	-4	-5	11	-3	-3	-4	-4	-5	-5	-5	-5	-5	-6
94	A	98-A	6	-2	-3	-3	-2	-3	-4	-4	-3	-2	0	2	-3	-3	-2	-3	-2	-3	-3	-3
95	R	99-R	0	-4	-4	-4	-4	-3	-3	-4	-3	-4	0	0	2	-2	-1	-2	1	6	-3	1
96	G	100-D	-2	4	-3	-4	1	-3	-4	-5	-3	-4	1	1	-4	-2	-2	2	-3	-3	3	0
97	L	101-F	0	1	-2	2	-2	2	1	-2	3	-3	-2	-3	3	1	-3	-2	-3	-3	-3	-3
98	G	102-G	0	4	-4	-4	-4	-3	-3	-3	3	-4	-2	-3	3	1	-2	-2	-2	2	-2	0
99	G	103-P	-2	3	-3	-4	0	-3	-3	4	2	-4	0	0	1	3	-2	2	-3	-3	1	-2
100	Y	104-D	-1	2	-2	-2	-2	-2	-1	-1	-1	-2	2	1	4	-1	-1	-1	-1	-2	1	-1
101	G	115-G	-1	3	-2	-2	-2	-1	-1	-1	-1	-2	0	-1	4	-1	-1	-1	-1	1	-1	-1
102	Y	116-R	1	-2	-2	-2	-2	-2	-1	-1	-2	3	1	-1	5	-2	1	2	-1	1	-2	-1
103	G	117-G	-2	3	-4	-4	-4	-3	0	-4	-3	3	2	-2	-3	-1	-2	2	-2	0	2	0
104	Y	118-F	2	0	1	0	0	-2	1	-2	-3	-3	-2	-2	5	0	-3	3	-3	-3	0	-3
105	F	119-F	-3	-5	1	-1	1	2	7	4	0	-4	-3	0	0	-4	-4	-3	0	-3	-5	-4
106	D	120-D	-1	-3	-5	-5	-5	-4	-5	-6	-3	-5	-2	-3	-4	2	-1	-2	-2	0	6	3
107	Y	121-F	-3	-5	0	1	-2	-2	2	0	-5	-4	-4	-3	8	-4	-3	3	-4	-3	-5	-4
108	W	122-W	-4	0	-4	-3	-5	-3	-1	12	-5	-4	-4	-4	1	-5	-4	-4	-5	-4	-6	-5
109	G	123-G	0	7	-5	-5	-5	-4	-5	-4	-4	-4	-2	-3	-5	-2	-3	-4	-3	-4	-3	-4
110	Q	124-Q	-3	-4	-5	-4	-4	-2	-5	-4	-3	-5	-2	-2	-3	-2	7	3	3	2	-2	0
111	G	125-G	1	7	-5	-5	-5	-4	-5	-4	-4	-4	-2	-3	-5	-2	-3	-4	-3	-4	-3	-4
112	T	126-T	0	-3	-2	0	-2	-2	-4	-4	-3	-3	1	6	-3	-2	-2	-3	-2	-3	-3	-3
113	L	127-L	-2	-4	-1	2	1	5	-3	-4	1	-3	0	0	-3	-3	2	-3	0	-2	-3	0
114	V	128-V	-2	-5	1	0	6	-1	-3	-5	-4	-3	-3	1	-3	-4	-4	-5	-4	-4	-5	-4
115	T	129-T	0	-3	-2	-3	1	-2	-4	-4	-3	-2	0	7	-3	-2	-2	-4	-2	-3	-3	-3
116	V	130-V	-2	-5	1	-1	6	-1	1	-4	-4	-3	0	-2	-2	-4	-4	-5	-4	-4	-5	-4
117	S	131-S	-1	-2	-4	-4	-3	-3	-4	-5	-3	-3	6	3	-3	-1	-2	-3	-2	-2	-2	-2

FIG. 7 (continuation)

CD04980 - IgV_L_kappa: Ig light chain, kappa type, variable (V) domain.

P	C	Master	A	G	I	L	V	M	F	W	P	C	S	T	Y	N	Q	H	K	R	D	E
1	I	2-T	-2	-4	3	2	3	0	-2	-4	-3	-2	-2	1	-2	-3	-3	4	-3	-3	-4	-3
2	V	3-T	0	-3	0	-2	3	-1	-3	-4	-3	-2	-1	3	-3	-2	2	-3	-2	-2	-2	2
3	L	4-V	-2	-4	1	3	3	5	-1	-3	-3	-3	-3	-2	-2	-3	-2	-3	-3	-3	-3	1
4	T	5-T	-1	-3	-1	-2	2	-1	-3	-4	-2	-2	0	6	-3	-1	-2	-3	-2	-2	-2	-2
5	Q	6-Q	-2	-3	-4	-3	-3	-1	-4	-3	-2	-4	-1	-2	-3	-1	7	-1	0	0	-1	1
6	S	7-S	0	-1	-4	-4	-3	-3	-4	-4	-2	-2	6	1	-3	0	-1	-2	-1	-2	-1	-1
7	P	8-P	-2	-3	-4	-4	-4	-4	-5	-5	8	-4	-2	-2	-4	-3	-2	-3	-2	-3	-3	-2
8	A	9-A	3	2	-3	0	-2	-2	-3	-4	-2	-3	1	-1	-3	-1	-2	-3	-2	-3	3	-1
9	T	10-S	-1	1	-3	-3	-2	-2	-1	-2	-3	-2	3	4	4	-1	-2	-1	-2	-2	-2	-2
10	L	11-L	-2	-4	0	5	1	1	-1	-3	-4	-2	-3	-2	-2	-4	-3	-4	1	-2	-4	-3
11	S	12-S	1	-2	-2	1	-2	3	-2	-4	-2	-2	4	0	-3	-1	-1	-2	-2	-2	-2	-2
12	V	13-V	1	-4	1	2	5	0	-2	-4	-3	-2	-2	-1	-2	-4	-3	-4	-3	-3	-4	-3
13	S	14-A	1	-2	-3	-3	-2	-2	-4	-4	-2	-2	4	2	-3	-1	2	-2	2	-1	-2	-1
14	P	15-I	2	-3	3	-1	2	-1	-3	-4	5	-3	-2	-2	-3	-3	-3	-4	-3	-3	-3	-3
15	G	16-G	-1	7	-5	-5	-4	-4	-4	-4	-3	-4	-1	-3	-4	-1	-3	-3	-3	-3	-2	-3
16	E	17-E	-2	-3	-4	-4	-4	-3	-4	-4	-2	-5	-1	-2	-3	-1	3	-1	0	-1	3	6
17	S	18-K	-1	-2	-4	-3	-3	-3	-4	-4	2	-3	3	1	-3	-1	-1	-2	2	3	-2	-1
18	A	19-V	4	-2	0	-1	4	-1	-3	-4	-3	-2	-1	-1	-3	-3	-3	-3	-2	-3	-4	-3
19	T	20-T	0	-2	-2	-3	-2	-2	-3	-4	-2	-2	4	5	-3	-1	-1	-2	-2	-2	-2	-2
20	I	21-I	-2	-5	6	3	1	0	-1	-3	-4	-2	-4	-2	-2	-4	-4	-4	-4	-4	-4	-4
21	S	22-R	0	-2	-3	-3	-2	-2	-4	-4	-2	-2	5	3	-3	-1	-1	-2	-1	2	-2	-1
22	C	23-C	-1	-4	-2	-2	-2	-3	-4	-3	-4	10	-2	-2	-4	-4	-4	-4	-4	-5	-5	-5
23	K	24-I	-2	0	1	-2	-2	-2	-3	-4	-3	-3	-1	2	-3	-2	0	-2	5	3	-2	-1
24	A	25-T	4	-1	-2	-3	-1	-2	-3	-4	-2	-2	2	3	-3	-2	-2	-3	-2	-2	-2	-2
25	S	26-S	0	-1	-4	-4	-3	-3	-4	-4	-2	-2	6	0	-3	2	-1	-2	-1	-2	-1	-1
26	Q	27-T	-1	-2	-3	-2	-2	3	-3	-3	-2	-3	3	2	-3	2	5	-1	-1	-1	-1	0
27	S	28-D	-1	2	1	-3	-2	-2	-3	-4	-2	-3	4	-1	-3	-1	-1	-2	1	-2	2	-1
28	V	29-I	-2	-4	4	1	4	0	1	-3	-4	-2	0	-1	-2	-3	-3	-4	-3	-3	-4	-3
29	S	--	0	3	1	-2	-1	-1	-2	-2	-1	-2	3	0	-2	-1	-1	-2	-1	-2	-1	-1
30	S	30-D	-1	2	-4	-4	-3	-3	-2	-3	-2	-3	3	-1	3	-1	-1	-2	1	-2	3	-1
31	N	31-D	-2	0	-3	-3	-3	-3	-2	-3	-3	-3	1	3	3	4	-2	-1	-2	-2	2	-1
32	Y	32-D	-2	-3	-3	-3	-3	-3	-1	-2	-3	4	0	-2	5	4	-2	-1	-2	1	2	-2
33	L	33-M	-2	-5	2	4	0	5	-1	-3	-4	-2	-3	-2	-2	-4	-3	-4	-3	-3	-5	-4
34	A	34-N	4	-1	-3	-3	-2	-2	-4	-4	-2	-2	3	-1	-3	4	-1	-2	-1	-2	-1	-2
35	W	35-W	-4	-4	-4	-3	-4	-3	0	12	-5	-3	-4	-4	1	-5	-3	-4	-4	-4	-6	-4
36	Y	36-Y	-3	-4	-2	0	-2	-1	2	1	-4	-3	-3	-3	8	-3	-3	1	-3	-3	-4	-3
37	Q	37-Q	-2	-3	-2	2	-2	-1	-3	-3	-3	-4	-2	-2	-2	-2	6	-1	0	-1	-2	1
38	Q	38-Q	-2	-3	-3	0	-3	-1	-3	-3	-3	-4	-1	-2	-3	-1	7	-1	0	0	-1	1
39	K	39-K	-2	-3	-4	-4	-3	-2	-4	-4	-2	-4	-1	-2	-3	-1	0	-2	6	3	-2	0
40	P	40-P	-2	-3	-4	-4	-3	-4	-5	-5	8	-4	1	-2	-4	-3	-2	-3	-2	-3	-2	-2
41	G	41-G	-1	7	-5	-5	-4	-4	-4	-4	-3	-4	-1	-3	-4	-1	-3	-3	-3	-3	-2	-3
42	Q	42-E	-2	-3	-4	-4	-3	-2	-4	-4	-2	-4	-1	-2	-3	-1	5	-1	3	0	-1	4
43	A	43-P	4	-1	-3	-3	-2	-2	-4	-4	3	-2	2	2	-3	-2	-2	-3	-2	-2	-2	-2
44	P	44-P	-2	-3	-4	-4	-4	-4	-5	-5	8	-4	-2	-2	-4	2	-2	-3	-2	-3	-2	-2
45	K	45-K	-2	-3	-4	-4	-3	-2	-4	-4	-2	-4	-1	-2	-3	-1	0	-2	6	4	-2	0
46	L	46-F	-3	-5	0	5	-1	1	3	-2	-4	-3	-3	-2	-1	-4	-3	-3	-2	1	-4	-4
47	L	47-F	-3	-5	0	5	0	1	3	-2	-4	-3	-4	-2	-1	-5	-3	-4	-4	-3	-5	-4
48	I	48-I	-2	-5	7	1	2	0	-1	-4	-4	-2	-4	-2	-2	-4	-4	-4	-4	-4	-4	-4
49	Y	49-S	-2	-3	-3	-2	-2	-2	1	1	-3	-3	2	-2	8	-2	-2	0	-2	-3	-3	-3
50	G	50-E	1	3	-3	0	-3	-2	-2	-3	-3	-3	1	-2	2	-2	-2	-2	-2	-3	1	1
51	A	51-G	3	3	-2	-3	1	-2	-3	-4	-3	-2	0	2	-3	2	-2	-3	-2	-3	-2	-2
52	S	52-N	-1	-2	-4	-4	-3	-3	-4	-4	-2	-3	5	0	-3	4	2	-1	-1	-1	-1	-1
53	T	53-T	-1	-2	-3	-3	-2	-2	-4	-4	-2	-3	2	4	-3	3	-1	-2	3	-1	-1	-1
54	L	54-L	-2	-3	0	4	-1	0	-1	-3	-3	-3	-2	-2	-2	-2	-1	-2	-1	4	-3	-2
55	A	55-R	3	-2	-3	-3	-3	-2	-4	-4	-2	-3	-1	-2	-2	-1	4	3	-1	2	2	0
56	S	56-P	0	-2	-3	-3	-2	-2	-4	-4	3	-2	4	4	-3	-1	-1	-2	-1	-2	-2	-2
57	G	57-G	-1	6	-5	-5	-4	-4	-4	-4	-3	-4	-1	-2	-4	2	-3	-3	-2	-3	-2	-3
58	T	58-V	-2	-4	3	-1	3	-1	-2	-4	-3	-2	-1	4	-3	-2	-3	-3	-3	-3	1	-2
59	P	59-P	-1	-3	-4	-4	-3	-3	-5	-5	8	-4	1	-2	-4	-3	-2	-3	-2	-3	-2	-2
60	S	60-S	-1	2	-4	-4	-4	-3	-4	-4	-2	-3	3	-1	-4	0	-1	-2	-1	-2	4	2

FIG. 8

61	R	61-R	-3	-3	-4	-3	-4	-2	-4	-4	-3	-5	-2	-2	-3	-1	0	-1	1	7	-3	-1
62	F	62-F	-3	-4	1	0	-1	-1	8	0	-5	-3	-4	-3	2	-4	-4	-2	-4	-4	-5	-4
63	S	63-S	0	-1	-3	-3	-2	-2	-4	-4	-2	-2	6	2	-3	0	-1	-2	-1	-2	-1	-1
64	G	64-S	-1	6	-5	-5	-4	-4	-4	-4	-3	-3	1	-2	-4	-1	-3	-3	-2	-3	-2	-3
65	S	65-S	0	-1	-3	-3	-2	-2	-3	-4	-2	-2	5	2	-3	0	-1	-2	-1	-2	-1	-1
66	G	66-G	-1	6	-4	-4	-4	-4	-2	-2	-3	-4	-1	-3	3	-2	-3	-2	-3	-3	-3	-3
67	S	67-Y	0	-2	-3	-3	-3	-3	-2	-2	-2	-2	5	0	3	-1	-1	-1	-1	-2	-2	-1
68	G	68-G	-1	6	-5	-5	-4	-4	-4	-4	-3	-4	-1	-3	-4	-1	-2	-3	-2	1	2	-2
69	T	69-T	-1	-3	-2	-2	-1	-2	-2	-2	-2	-2	0	6	3	-1	-2	-2	-2	-2	-2	-2
70	D	70-D	-3	-3	-2	-4	1	-3	-4	-5	-3	-4	-1	-2	-4	0	-2	-2	-2	-3	7	0
71	F	71-F	-3	-4	-1	-1	-2	-1	8	0	-5	-4	-4	-3	2	-4	-4	-2	-4	-4	-5	-4
72	T	72-V	-1	-3	-1	-2	1	-1	-3	-4	-2	-2	0	6	-3	-1	-2	-3	-2	-2	-2	-2
73	L	73-F	-2	-5	1	5	2	1	3	-2	-4	-2	-3	-2	-1	-4	-3	-4	-4	-3	-5	-4
74	T	74-T	-1	-2	-2	-3	-2	-2	-4	-4	-2	-2	2	5	-3	3	-1	-2	1	-1	-1	-1
75	I	75-I	-2	-5	7	1	2	0	-1	-4	-4	-2	-4	-2	-2	-4	-4	-4	-4	-4	-4	-4
76	S	76-E	-1	0	-4	-4	-3	-3	-4	-4	-2	-3	4	-1	-3	-1	0	-2	0	4	-1	2
77	R	77-N	-2	3	-4	-4	-4	-3	-4	-4	-3	-4	1	-2	-3	2	-1	-2	0	5	-2	-2
78	V	78-M	-2	-4	1	3	4	5	-1	-3	-4	-2	-3	-2	-2	-4	-3	-4	-3	-3	-5	-4
79	E	79-L	-2	-3	-3	0	-3	-2	-4	-4	-2	-4	-1	-2	-3	-1	4	-1	0	-1	0	5
80	P	80-S	2	-2	-3	-3	-2	-2	2	-3	6	-3	1	1	-2	-2	-2	-3	-2	-3	-3	-2
81	E	81-E	-2	-3	-4	-4	-4	-3	-4	-4	-2	-5	-1	-2	-3	-1	3	-1	0	-1	0	6
82	D	82-D	-3	-2	-4	-5	-4	-4	-5	-6	-3	-5	-1	-2	-4	0	-1	-2	-2	-3	7	1
83	A	83-V	3	-3	0	2	1	-1	4	-2	-3	-2	-1	-2	0	-4	-3	-3	-3	-3	-4	-3
84	A	84-A	4	4	-3	-3	-2	-3	-4	-4	-2	-2	0	-2	-3	-2	-2	-3	-2	-3	-3	-2
85	V	85-D	-2	-4	2	-1	4	-1	-3	-4	-3	-3	-2	-1	-3	-2	-2	-3	-3	-3	4	-1
86	Y	86-Y	-3	-4	-2	-2	-2	-2	2	1	-4	-4	-3	-3	9	-3	-3	1	-3	-3	-4	-3
87	Y	87-Y	-3	-4	-2	-2	-2	-2	2	1	-4	-4	-3	-3	9	-3	-3	1	-3	-3	-4	-3
88	C	88-C	-1	-4	-2	-2	-2	-3	-4	-3	-4	10	-2	-2	-4	-4	-4	-4	-4	-5	-5	-5
89	Q	89-L	-2	-4	-1	2	1	4	-2	-3	-3	-3	-2	-2	-2	-2	5	-2	-1	-1	-2	0
90	Q	90-Q	-1	2	-4	-4	-3	-2	-4	-3	-2	-4	1	-1	-3	-1	6	-1	1	0	-1	1
91	Y	91-S	-2	3	-2	0	-2	-2	1	0	-4	-3	1	-2	6	-2	-2	-1	-3	-3	-3	-3
92	G	92-D	-2	3	-4	-4	-4	-3	-1	5	-3	-4	1	-2	3	2	-2	4	-2	-2	2	-2
93	T	93-T	-1	-3	-3	-3	-3	-2	-3	6	-2	-3	2	3	-1	-1	-1	4	-1	-2	-1	3
94	F	94-L	-2	-3	2	2	-1	-1	4	-2	2	-3	2	-1	0	-3	-3	-3	-3	-3	-3	-3
95	P	95-P	0	-3	-1	-2	2	-2	-2	-3	6	-3	-2	-2	2	-3	-3	-3	-2	-3	-3	-3
96	Y	96-L	-3	-4	1	2	-1	-1	0	6	2	-3	-3	-3	5	-4	-3	-1	-3	-3	-4	-3
97	T	97-T	-1	-3	-2	-2	-1	-2	-3	-4	-2	-2	0	6	-3	2	-2	-2	-2	-2	-2	-2
98	F	98-F	-3	-4	-1	-1	-2	-1	8	0	-5	-4	-4	-3	2	-4	-4	-2	-4	-4	-5	-4
99	G	99-G	-1	7	-5	-5	-4	-4	-4	-4	-3	-4	-1	-3	-4	-1	-3	-3	-3	-3	-2	-3
100	G	100-S	2	4	-4	-4	-3	-3	-4	-4	-2	-3	1	-1	-3	-1	2	-2	1	-1	-2	-1
101	G	101-G	-1	7	-5	-5	-4	-4	-4	-4	-3	-4	-1	-3	-4	-1	-3	-3	-3	-3	-2	-3
102	T	102-T	-1	-3	-2	-2	-1	-2	-3	-4	-2	-2	1	7	-3	-1	-2	-3	-2	-2	-2	-2
103	K	103-K	-2	-3	-3	-2	-3	3	-4	-4	-3	-4	-2	-2	-3	-1	0	-2	5	5	-2	-1
104	L	104-L	-2	-5	1	5	2	1	-1	-3	-4	-2	-3	-2	-2	-5	-3	-4	-4	-3	-5	-4
105	E	105-E	-2	-3	-3	0	-3	-2	-3	-4	-2	-4	1	-1	-3	-1	0	-1	-1	-1	0	6
106	I	106-I	-2	-5	5	3	3	0	-1	-4	-4	-2	-3	-2	-2	-4	-4	-4	-4	-4	-4	-4

FIG. 8 (continuation)

CD04984 - IgV_L_lambda: Ig light chain, lambda type, variable (V) domain.

P	C	Master	A	G	I	L	V	M	F	W	P	C	S	T	Y	N	Q	H	K	R	D	E
1	L	11-A	2	-3	0	3	3	0	-2	-4	-3	-2	-2	-2	-3	-4	-2	-3	-3	-3	-3	1
2	S	12-S	0	-2	-3	-4	-3	-3	-4	-4	-2	-2	6	2	-3	-1	-1	-2	-1	-2	-2	-1
3	V	13-F	1	1	0	-1	4	-1	2	-3	-3	-3	-2	2	-2	-3	-2	-3	-3	-3	-3	0
4	S	14-S	1	-2	-3	-3	-3	-3	-4	-4	-2	-2	5	2	-3	-1	-1	-2	1	-2	-2	-1
5	P	15-L	-2	-4	-1	2	2	-1	-3	-4	7	-3	-3	-2	-3	-4	-3	-4	-3	-4	-4	-3
6	G	16-G	-1	7	-5	-5	-5	-4	-5	-4	-3	-4	-1	-3	-4	-2	-3	-3	-3	-4	-3	-3
7	E	17-A	2	1	-4	-4	-3	-3	-4	-4	-2	-4	-1	-2	-3	-2	4	-2	1	-1	-1	4
8	T	18-S	-1	-3	-2	-3	-2	-2	-4	-4	-2	-2	3	6	-3	-1	-2	-3	-1	1	-2	-2
9	V	19-A	3	-3	1	-1	5	-1	-3	-4	-3	-2	-2	-1	-3	-4	-3	-4	-3	-3	-4	-3
10	T	20-K	-1	-3	-3	-3	-2	-2	-4	-4	-2	-3	1	4	-3	-1	-1	-2	3	3	-2	-1
11	I	21-L	-3	-5	5	4	1	1	-1	-3	-4	-3	-4	-2	-2	-5	-4	-4	-4	-4	-5	-4
12	T	22-T	-1	-2	-2	-3	-2	-2	-4	-4	-2	-2	3	6	-3	-1	-2	-3	-2	-2	-2	-2
13	C	23-C	-2	-4	-3	-3	-2	-3	-4	-4	-4	11	-2	-2	-4	-4	-4	-4	-4	-5	-5	-5
14	T	24-T	-1	-3	-3	-3	-2	-2	-4	-4	-2	-3	3	5	-3	-1	2	-2	-1	2	-2	-1
15	G	25-L	-2	4	-2	2	-2	-2	2	-3	-4	-3	1	-2	-2	-2	-3	-3	-2	1	-3	-3
16	S	26-S	-1	-2	-4	-4	-3	-3	-4	-4	-2	-3	5	0	-3	2	-1	-2	-1	-2	2	1
17	S	27-S	-1	3	-4	-4	-3	-3	-4	-4	-3	-3	2	2	-3	3	2	-2	1	-1	-1	-1
18	G	28-Q	1	4	1	1	-2	-1	-3	-4	-3	-3	1	-2	-3	-2	2	-3	-2	-3	-3	-2
19	N	--	1	-1	-1	1	-1	-1	-1	-2	-1	-1	1	0	-1	3	-1	-1	-1	-1	-1	-1
20	I	29-H	-1	1	3	-1	1	-1	-2	-2	-2	-2	-1	2	-1	-1	-1	4	-2	-2	-2	-2
21	S	30-S	2	2	-3	-4	-3	-3	-4	-4	2	-3	3	2	-4	-1	-2	-3	-2	-3	1	-2
22	G	31-T	-1	4	-4	-4	-3	-3	-4	-4	-3	-3	2	2	-4	-1	-2	-3	1	-2	1	-2
23	N	32-Y	0	-3	-3	-3	-3	-3	3	-1	-3	-4	-1	-2	6	5	-2	0	1	-2	-2	-2
24	Y	33-T	-3	-4	-3	-2	-2	-2	3	0	-4	-3	-2	3	6	-2	-2	5	-3	-3	-3	-3
25	V	34-I	2	-3	3	-1	5	-1	-2	-4	-3	-2	-2	-1	-3	-4	-3	-4	-3	-4	-4	-3
26	N	35-E	-3	-3	-4	-3	-3	-3	2	-2	-3	4	-2	-2	3	4	2	5	-2	-2	-1	2
27	W	36-W	-4	-4	-4	-3	-4	-3	0	12	-5	-4	-4	-4	1	-5	-3	-4	-4	-4	-6	-4
28	Y	37-Y	-3	-4	-1	-2	1	-2	4	1	-4	-3	-3	-3	8	-4	-3	0	-3	-3	-5	-4
29	Q	38-Q	-2	-3	-4	-3	-4	-2	-5	-3	-3	-4	-1	-2	-3	-1	8	-1	0	0	-2	1
30	Q	39-Q	-2	-3	-4	-4	-4	-2	-4	-3	-3	-4	-1	-2	-2	-1	7	4	0	0	-1	3
31	K	40-Q	-2	-3	1	0	-2	-2	-3	-4	-3	-4	-2	-2	-3	-2	3	-2	5	3	-2	-1
32	P	41-P	1	-3	-3	0	-3	-3	-4	-5	7	-3	1	-2	-4	-3	-2	-3	-2	-3	-3	-2
33	G	42-L	-2	5	-3	0	-4	-3	-4	-4	2	-4	-2	-3	-4	-1	-3	-3	-3	-3	3	-2
34	S	43-K	-1	-2	-4	-4	-3	-3	-4	-4	-2	-3	3	1	-2	2	2	4	3	-1	-2	-1
35	A	44-P	3	-2	-3	0	-2	-2	-3	-4	4	-3	1	-2	-3	2	-2	-3	-2	1	-2	-2
36	P	45-P	-2	-4	1	-3	-2	-3	2	-4	8	-4	-2	-2	-2	-4	-3	-3	-3	-4	-3	-3
37	R	46-K	-2	-4	1	0	1	-1	-3	-4	-3	-3	-2	1	-3	-2	-1	-2	3	4	-3	-2
38	Y	47-Y	-2	1	-1	3	1	-1	0	-1	-4	-3	-2	1	6	-3	-3	-1	-3	-3	-4	-4
39	L	48-V	-2	-5	2	4	4	0	-1	-3	-4	-2	-3	-2	-2	-5	-4	-4	-4	-4	-5	-4
40	I	49-M	-3	-5	6	3	1	4	-1	-4	-4	-3	-4	-2	-3	-5	-4	-4	-4	-4	-5	-4
41	Y	50-E	-3	1	-3	-3	-3	-3	3	-1	-4	-4	-2	-3	6	-2	-1	-1	-1	3	-3	1
42	E	51-L	0	1	-2	1	-2	3	2	-2	-3	-3	-2	-2	3	-3	2	-2	-2	-2	-2	2
43	D	52-K	-2	-3	-4	-4	-3	-3	-3	-3	-3	-4	-1	2	2	3	-1	-1	3	-1	4	-1
44	S	53-K	-1	-2	-4	-4	-3	-3	-4	-4	-2	-3	3	1	-3	4	-1	-2	3	-1	-1	-1
45	D	54-D	-3	-3	-4	-4	-4	-3	-4	-4	-3	-5	-1	-2	-3	2	3	4	2	2	4	0
46	R	59-T	-2	-3	-4	-3	-3	-2	-4	-4	-3	-4	-1	1	-3	-1	4	-1	1	6	-2	-1
47	P	60-G	-2	4	-4	-4	-3	-4	-5	-4	6	-4	-1	1	-4	-2	-3	-3	-2	-3	-3	-3
48	S	61-D	1	-2	-4	-4	-3	-3	-4	-5	3	-3	4	-1	-3	-1	-2	-2	-2	-2	3	-1
49	G	62-G	-1	7	-5	-5	-5	-4	-5	-4	-3	-4	-1	-3	-4	-2	-3	-3	-3	-4	-3	-3
50	I	63-I	1	-4	5	0	4	0	-2	-4	-4	-2	-3	-2	-3	-4	-4	-4	-4	-4	-4	-4
51	P	64-P	-2	-3	-4	-4	-4	-4	-5	-5	8	-4	-2	-2	-4	-3	-3	-4	-2	-3	-3	-2
52	D	65-D	1	-2	-4	-4	-4	-4	-4	-5	-2	-4	1	-2	-4	-1	-1	-2	-1	-2	5	4
53	R	66-R	-3	-4	-4	-4	-4	-3	-4	-4	-3	-5	-2	-2	-3	-2	0	-1	1	7	-3	-1
54	F	67-F	-4	-5	-1	-1	-2	-1	8	0	-5	-4	-4	-3	2	-4	-5	-3	-4	-4	-5	-5
55	S	68-S	1	-1	-4	-4	-3	-3	-4	-4	-2	-2	6	0	-3	-1	-1	-2	-1	-2	-2	-1
56	G	69-G	-1	7	-5	-5	-5	-4	-5	-4	-3	-4	-1	-3	-4	-2	-3	-3	-3	-4	-3	-3
57	S	70-S	0	-2	-3	-4	-3	-3	-4	-4	-2	-2	6	2	-3	-1	-1	-2	-1	-2	-2	-1
58	K	71-S	-2	-3	1	0	-2	-2	-3	-4	-3	-3	2	2	-3	2	-1	-2	4	-1	-2	-2
59	S	72-S	-1	-2	1	-3	-2	-3	-4	-5	-2	-3	4	-1	-3	-1	-2	-2	-2	-3	4	-1
60	G	73-G	1	6	-5	-4	-4	-3	-4	-4	-3	-3	1	-2	-4	-2	-2	-3	-2	1	-3	-3

FIG. 9

61	N	74-A	1	-2	-4	-4	-3	-3	-4	-5	-3	-3	0	3	-3	6	-1	-1	-2	-2	2	-1
62	T	75-D	-1	-3	1	-3	-1	-2	-3	-4	-2	-3	2	4	-3	-1	-2	-3	2	-2	2	-1
63	A	76-R	5	3	-3	-3	-2	-3	-4	-4	-2	-2	0	-2	-3	-2	-2	-3	-2	2	-3	-2
64	S	77-Y	1	-3	1	-2	-1	-2	-1	-2	-3	-2	3	3	5	-2	-2	-2	-2	-3	-3	-2
65	L	78-L	-3	-5	0	6	0	1	-1	-3	-4	-3	-4	-2	-2	-5	-3	-4	-4	-4	-5	-4
66	T	79-S	1	-2	-2	-3	-2	-2	-4	-4	-2	-2	2	6	-3	-1	-2	-3	-1	1	-2	-2
67	I	80-I	-3	-5	7	0	2	0	-1	-4	-4	-3	-4	-2	-3	-5	-4	-5	-4	-4	-5	-5
68	S	81-S	0	-2	-3	-3	-3	-3	-4	-4	-2	-2	5	3	-3	-1	-1	-2	1	-2	-2	-1
69	G	82-N	-1	6	-5	-5	-5	-4	-5	-4	-3	-4	-1	-3	-4	2	-3	-3	-3	-3	-2	-3
70	A	83-I	4	-2	2	3	0	-1	-2	-4	-3	-2	-1	1	-3	-3	-3	-3	-3	-3	-4	-3
71	Q	84-Q	-2	-3	-3	0	-3	-1	-4	-3	-3	-4	-2	-2	-3	-1	7	-1	2	0	-2	0
72	T	85-P	2	-3	-2	0	-2	-2	-3	-4	4	-2	1	4	-3	-2	-2	-3	-2	-3	-3	-2
73	E	86-E	-2	-4	-4	-1	-3	-3	-4	-4	-3	-5	-2	-2	-3	-2	1	-1	-1	-1	0	6
74	D	87-D	-3	-3	-5	-5	-5	-4	-5	-6	-3	-5	-1	-2	-4	0	-2	-2	-2	-3	7	0
75	E	88-E	-2	-3	-5	-4	-4	-3	-5	-4	-2	-5	-1	-2	-3	-1	1	-1	0	-1	3	6
76	A	89-A	6	-1	-3	-3	-1	-2	-4	-4	-2	-2	0	-1	-3	-3	-2	-3	-2	-3	-3	-2
77	D	90-I	0	-3	3	-2	-1	-2	-3	-5	-3	-4	-2	-2	-4	-1	-2	-3	-2	-3	6	-1
78	Y	91-Y	-3	-4	-3	-2	-2	-2	2	1	-4	-4	-3	-3	9	-3	-3	1	-3	-3	-4	-3
79	Y	92-I	-3	-5	1	-2	-2	-2	4	1	-4	-4	-3	-3	8	-4	-3	0	-3	-3	-5	-4
80	C	93-C	-2	-4	-3	-3	-2	-3	-4	-4	-4	11	-2	-2	-4	-4	-4	-4	-4	-5	-5	-5
81	Q	94-G	3	1	-3	0	-2	-2	-3	-4	-3	-3	-1	-2	-3	-2	5	-2	-1	-2	-2	-1
82	V	95-V	-1	-3	2	1	3	-1	-2	-4	-3	-2	3	2	-3	-2	-2	-3	-3	-3	-3	-3
83	W	96-G	-2	3	-4	-3	-4	-3	0	10	-4	-4	0	-3	4	-3	-3	-2	-3	-3	-4	-3
84	D	97-D	-3	-3	-4	-4	-4	-3	-1	-2	-3	-4	-2	-3	5	-1	-2	5	-2	-3	6	-1
85	S	98-T	2	1	-3	-3	-2	-3	-4	-4	-2	-2	5	2	-3	-1	-2	-3	-2	-2	-2	-2
86	N	103-F	0	1	-3	0	-3	-2	2	-3	-3	-3	2	-1	-2	4	-2	-2	-2	1	-2	-2
87	S	104-V	1	1	-2	-3	1	-3	-3	-4	-3	-3	3	-1	-2	2	-2	4	-2	-2	2	-2
88	Y	105-Y	0	1	-1	0	2	-1	0	6	-4	-3	-3	-2	6	-4	-3	-2	-3	-3	-4	-3
89	V	106-V	-2	-5	3	1	5	0	-2	-4	-4	-2	-3	-2	-3	-4	-3	-4	-3	1	-4	-4
90	F	107-F	-4	-5	-1	-1	-2	-1	8	0	-5	-4	-4	-3	2	-4	-5	-3	-4	-4	-5	-5
91	G	108-G	-1	7	-5	-5	-5	-4	-5	-4	-3	-4	-1	-3	-4	-2	-3	-3	-3	-4	-3	-3
92	G	109-G	-1	6	-5	-5	-4	-4	-4	-4	-3	-4	1	-2	-4	-1	-3	-3	-3	-3	-2	-3
93	G	110-G	-1	7	-5	-5	-5	-4	-5	-4	-3	-4	-1	-3	-4	-2	-3	-3	-3	-4	-3	-3
94	T	111-T	-1	-3	-2	-2	-1	-2	-3	-4	-2	-2	0	7	-3	-1	-2	-3	-2	-2	-2	-2
95	K	112-K	-2	-3	-4	-4	-4	-2	-4	-4	-3	-5	-2	-2	-3	-1	5	-1	4	4	-2	0
96	L	113-V	-3	-5	1	5	2	1	-1	-3	-4	-2	-4	-2	-2	-5	-3	-4	-4	-4	-5	-4
97	T	114-T	-1	-3	-2	-2	-1	-2	-3	-4	-2	-2	0	7	-3	-1	-2	-3	-2	-2	-2	-2
98	V	115-V	-1	-5	2	0	6	0	-2	-4	-4	-2	-3	-1	-2	-4	-4	-5	-4	-4	-5	-4

FIG. 9 (continuation)

VL
FW1
CDR1
FW2
CDR2
LC_Abagovomab DIELTQSPASLSASVGETVTITCQAS------ENIYS------YLAWHQQKQGKSPQLLVYNAKTLAGGVSS
LC_Abciximab DILMTQSPSSMSVSLGDTVSITCHAS------QGISS------NIGWLQQKPGKSFMGLIYYGTNLVDGVPS
LC_Arcitumomab QTVLSQSPAILSASPGEKVTMTCRAS------SSVT------YIHWYQQKPGSSPKSWIYATSNLASGVPA
LC_Tenatumomab DIVMTQAAPSVPVTPGESVSISCRSSKSLLH-SNGNT------YLYWFLQRPGQSPQLLIYRMSNLASGVPD
LC_Naptumomab SIVMTQTPTSLLVSAGDRVTITCKAS------QSVSN------DVAWYQQKPGQSPKLLISYTSSRYAGVPD
LC_Bococizumab DIQMTQSPSSLSASVGDRVTITCRAS------QGIS------SALAWYQQKPGKAPKLLIYSASYRYTGVPS
LC_Pateclizumab DIQMTQSPSSLSASVGDRVTITCRAS------QAVS------SAVAWYQQKPGKAPKLLIYSASHRYTGVPS
LC_Elotuzumab DIQMTQSPSSLSASVGDRVTITCKAS------QDVG------IAVAWYQQKPGKVPKLLIYWASTRHTGVPD
LC_Tigatuzumab DIQMTQSPSSLSASVGDRVTITCKAS------QDVG------TAVAWYQQKPGKAPKLLIYWASTRHTGVPS
LC_Pertuzumab DIQMTQSPSSLSASVGDRVTITCKAS------QDVS------IGVAWYQQKPGKAPKLLIYSASYRYTGVPS
LC_Trastuzumab DIQMTQSPSSLSASVGDRVTITCRAS------QDVN------TAVAWYQQKPGKAPKLLIYSASFLYSGVPS
LC_Enokizumab DIQMTQSPSSLSASVGDRVTITCKAS------QHVI------THVTWYQQKPGKAPKLLIYGTSYSYSGVPS
LC_Dacetuzumab DIQMTQSPSSLSASVGDRVTITCRSS------QSLVHSNGN-TFLHWYQQKPGKAPKLLIYTVSNRFSGVPS
LC_Inotuzumab DVQVTQSPSSLSASVGDRVTITCRSS------QSLANSYGN-TFLSWYLHKPGKAPQLLIYGISNRFSGVPD
LC_Anrukinzumab DIQMTQSPSSLSASVGDRVTITCKAS------ESVDNY-GK-SLMHWYQQKPGKAPKLLIYRASNLESGVPS
LC_Omalizumab DIQLTQSPSSLSASVGDRVTITCRAS------QSVDYD-GD-SYMNWYQQKPGKAPKLLIYAASYLESGVPS
LC_Ocrelizumab DIQMTQSPSSLSASVGDRVTITCRAS------SSVS------YMHWYQQKPGKAPKPLIYAPSNLASGVPS
LC_Rontalizumab DIQMTQSPSSLSASVGDRVTITCRAS------QSVSTS-SY-SYMHWYQQKPGKAPKVLISYASNLESGVPS
LC_Farletuzumab DIQLTQSPSSLSASVGDRVTITCSVS------SSISS------NNLHWYQQKPGKAPKPWIYGTSNLASGVPS
LC_Veltuzumab DIQLTQSPSSLSASVGDRVTMTCRAS------SSVS------YIHWFQQKPGKAPKPWIYATSNLASGVPV
LC_Teplizumab DIQMTQSPSSLSASVGDRVTITCSAS------SSVS------YMNWYQQTPGKAPKRWIYDTSKLASGVPS
LC_Clivatuzumab DIQLTQSPSSLSASVGDRVTMTCSAS------SSVSS------SYLYWYQQKPGKAPKLWIYSTSNLASGVPA
LC_Daclizumab DIQMTQSPSTLSASVGDRVTITCSAS------SSIS------YMHWYQQKPGKAPKLLIYTTSNLASGVPA
LC_Motavizumab DIQMTQSPSTLSASVGDRVTITCSAS------SRVG------YMHWYQQKPGKAPKLLIYDTSKLASGVPS
LC_Benralizumab DIQMTQSPSSLSASVGDRVTITCGTS------EDIIN------YLNWYQQKPGKAPKLLIYHTSRLQSGVPS
LC_Romosozumab DIQMTQSPSSLSASVGDRVTITCRAS------QDISN------YLNWYQQKPGKAPKLLIYYTSRLLSGVPS
LC_Tanezumab DIQMTQSPSSLSASVGDRVTITCRAS------QSISN------NLNWYQQKPGKAPKLLIYYTSRFHSGVPS
LC_Tocilizumab DIQMTQSPSSLSASVGDRVTITCRAS------QDISS------YLNWYQQKPGKAPKLLIYYTSRLHSGVPS
LC_Gevokizumab DIQMTQSTSSLSASVGDRVTITCRAS------QDISN------YLSWYQQKPGKAVKLLIYYTSKLHSGVPS
LC_Bevacizumab DIQMTQSPSSLSASVGDRVTITCSAS------QDISN------YLNWYQQKPGKAPKVLIYFTSSLHSGVPS

FIG. 10

FIG. 10 (continuation)

FIG. 10 (continuation)

FIG. 10 (continuation)

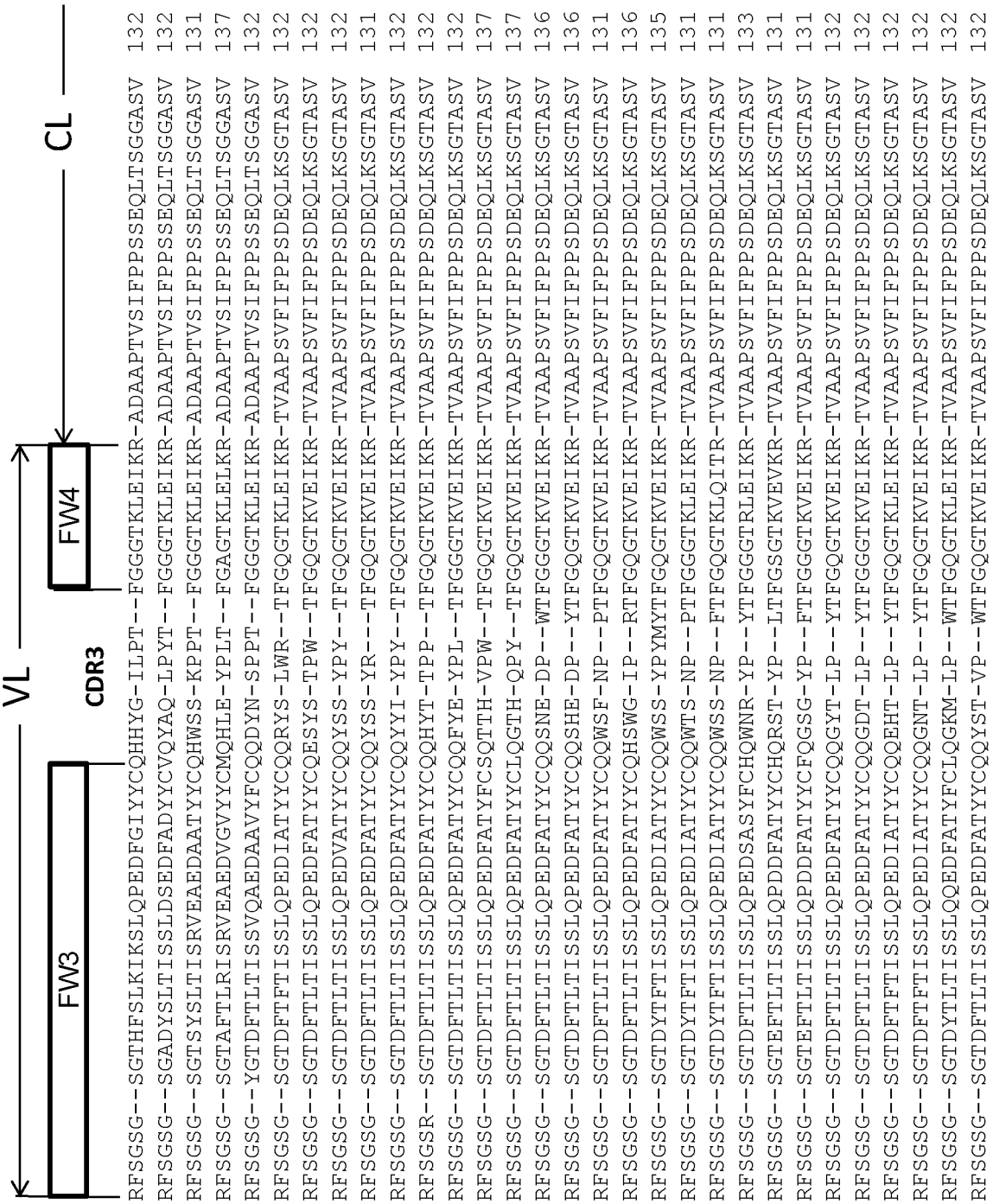


FIG. 10 (continuation)

FIG. 10 (continuation)

FIG. 10 (continuation)

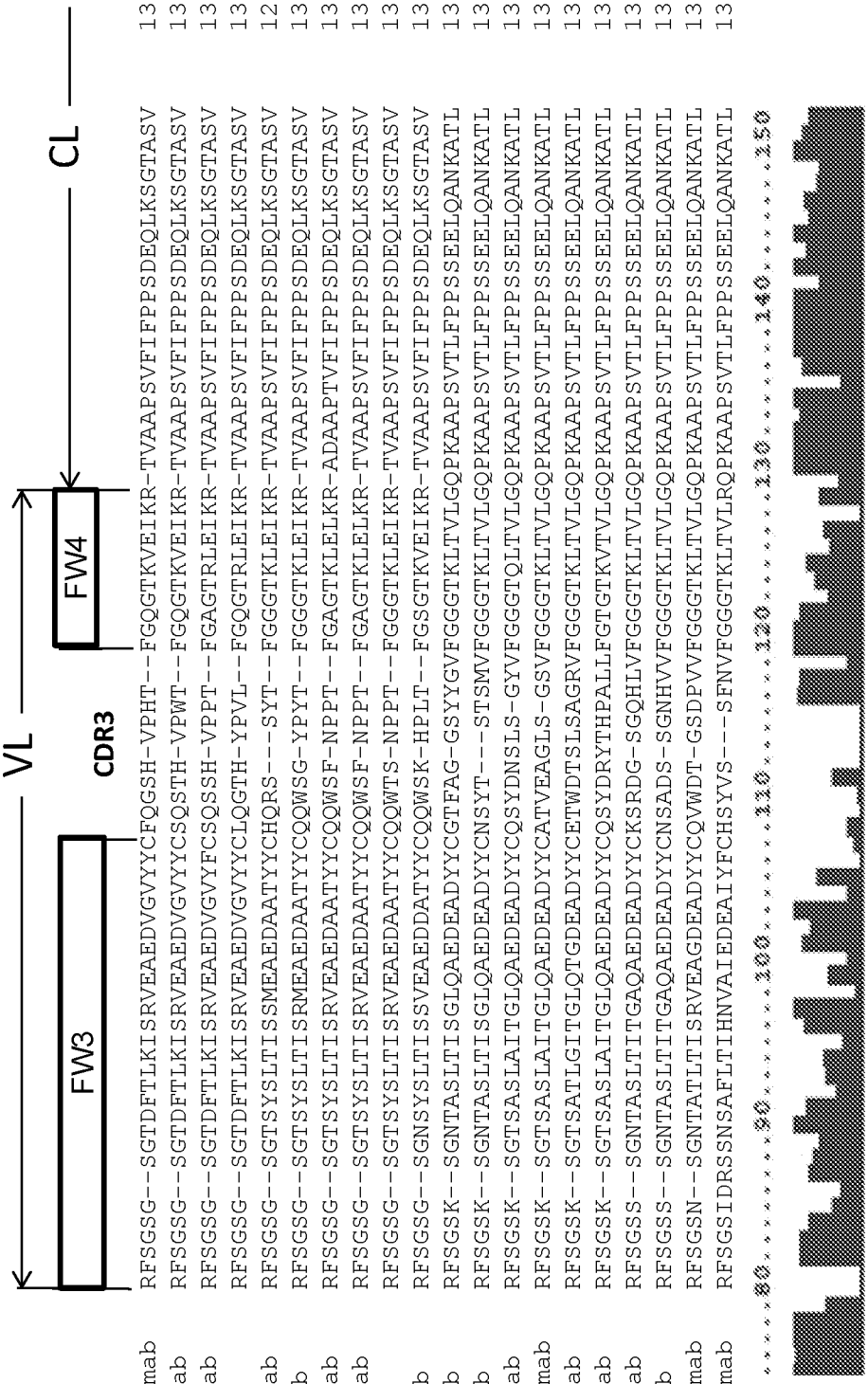


FIG. 10 (continuation)

FIG. 10 (continuation)

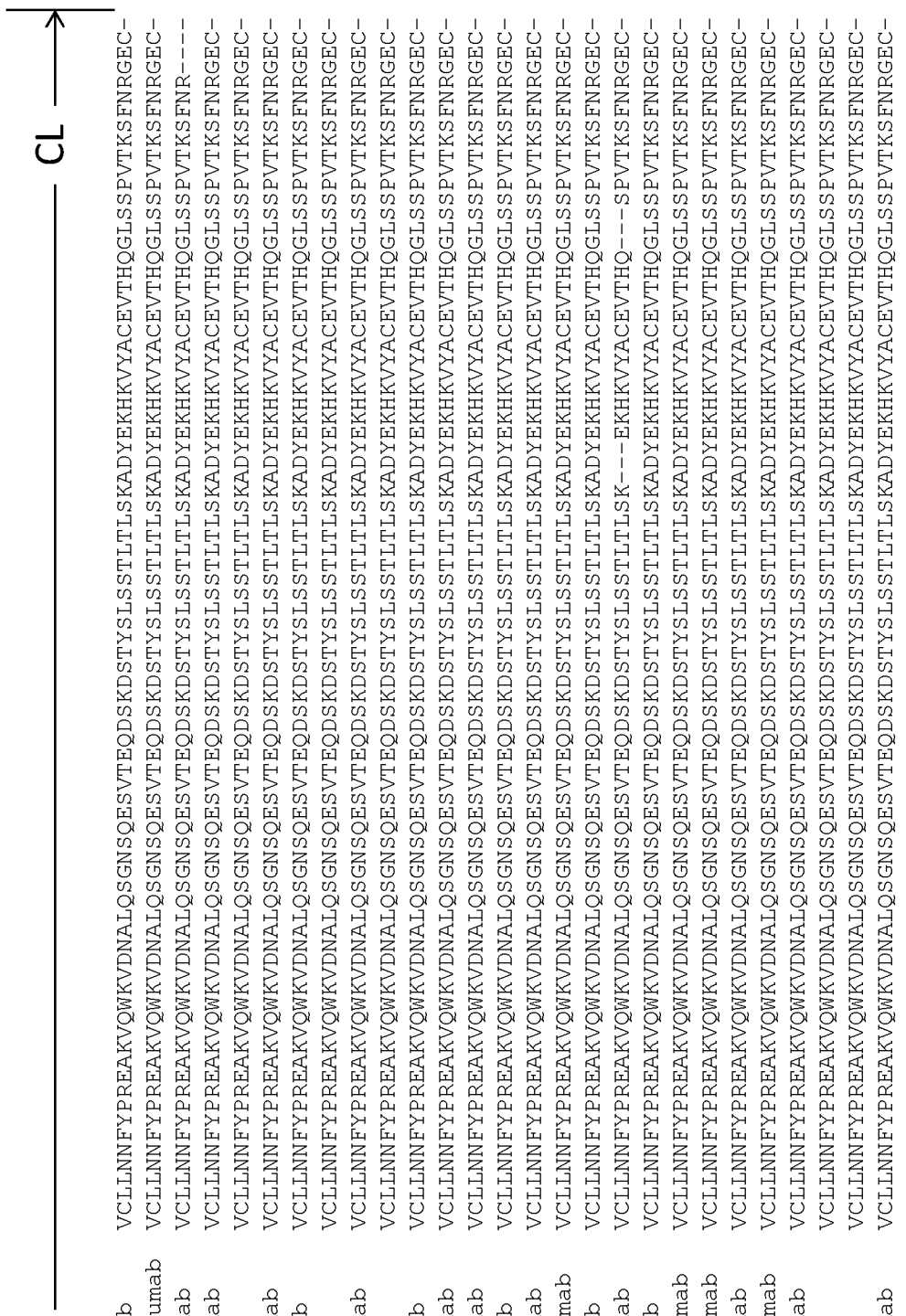


FIG. 10 (continuation)

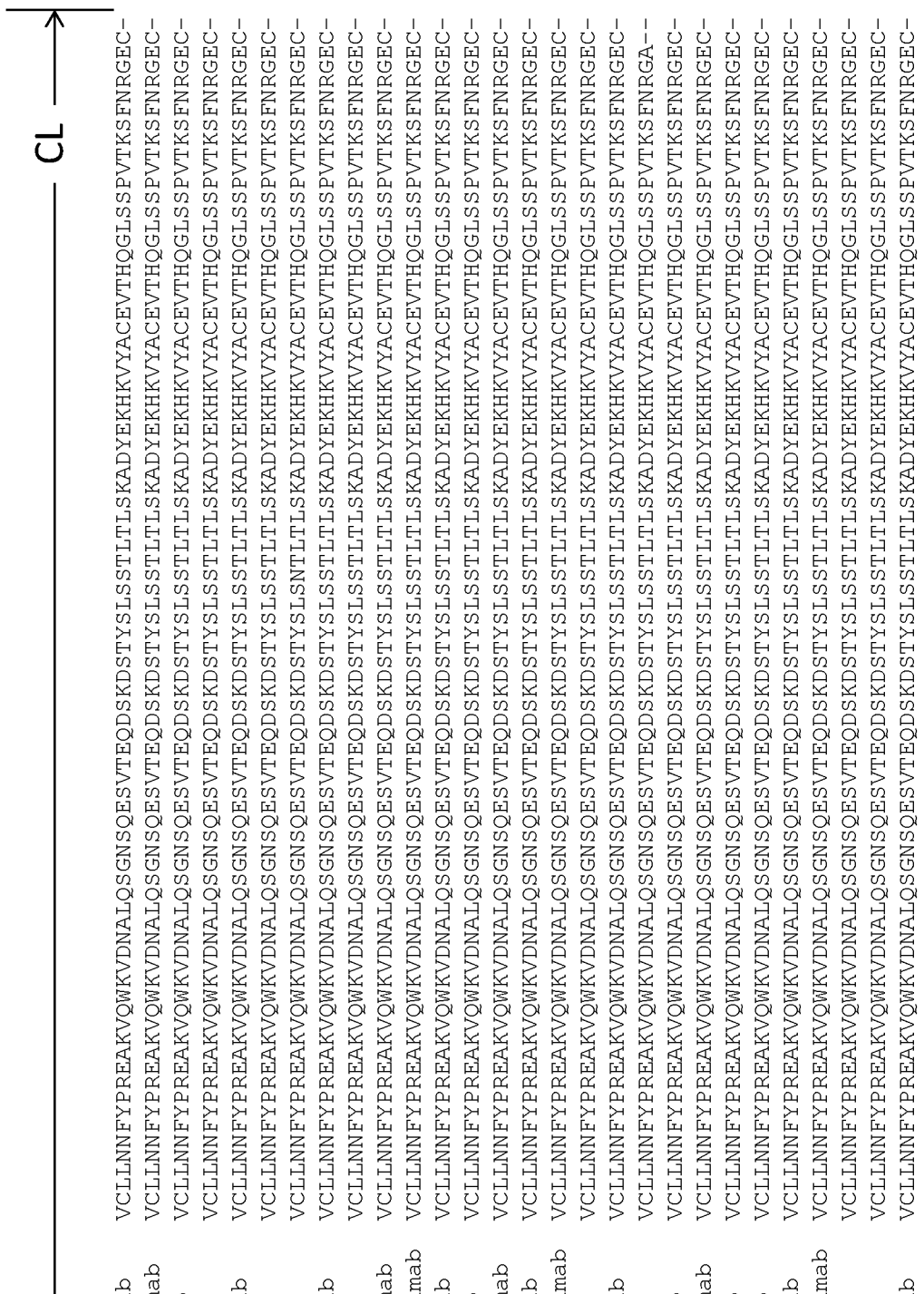


FIG. 10 (continuation)

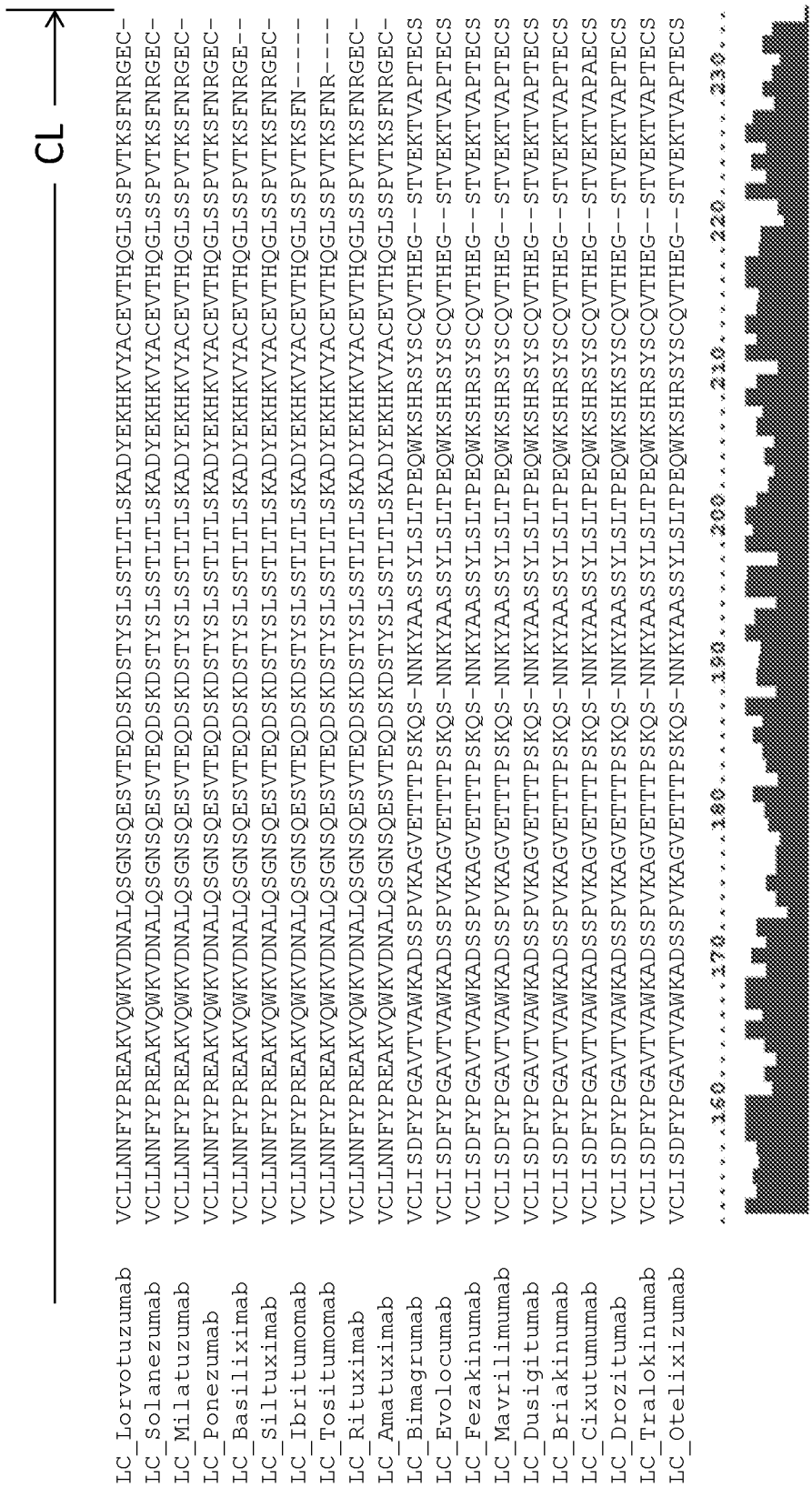


FIG. 10 (continuation)

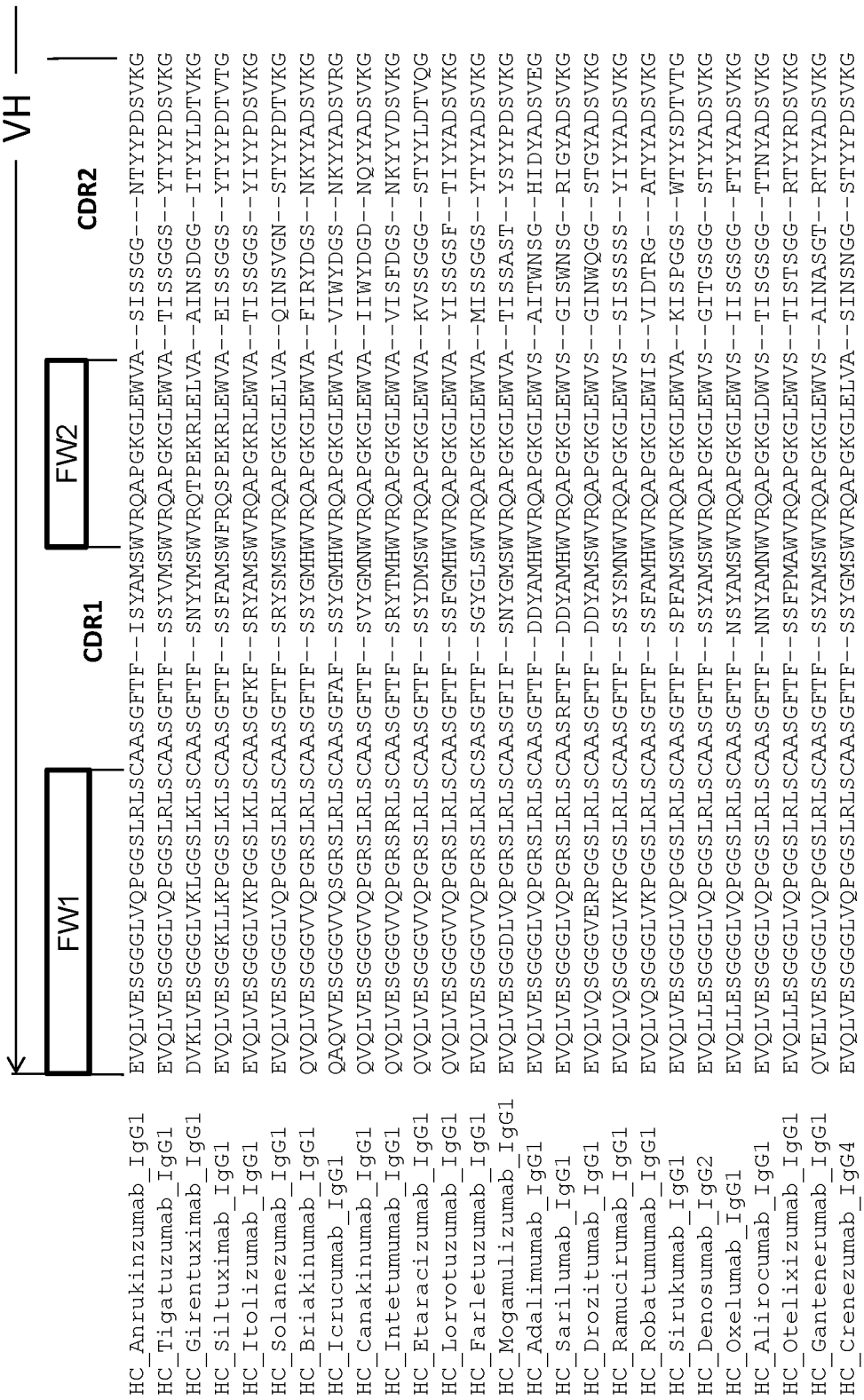

VH
FW1
CDR1
FW2
CDR2
HC_Anrukinzumab_IgG1 EVQLVESGGGLVQPGGSLRLSCAASGFTF--ISYAMSWVRQAPGKGLEWVA--SISSGG--NTYYPDSVKG
HC_Tigatuzumab_IgG1 EVQLVESGGGLVQPGGSLRLSCAASGFTF--SSYVMSWVRQAPGKGLEWVA--TISSGGS--YTYYPDSVKG
HC_Girentuximab_IgG1 DVKLVESGGGLVKLGGSLKLSCAASGFTF--SNYYMSWVRQTPEKRLELVA--AINSDGG--ITYYLDTVKG
HC_Siltuximab_IgG1 EVQLVESGGKLLKPGGSLKLSCAASGFTF--SSFAMSWFRQSPEKRLEWVA--EISSGGS--YTYYPDTVTG
HC_Itolizumab_IgG1 EVQLVESGGGLVKPGGSLKLSCAASGFKF--SRYAMSWVRQAPGKRLEWVA--TISSGGS--YIYYPDSVKG
HC_Solanezumab_IgG1 EVQLVESGGGLVQPGGSLRLSCAASGFTF--SRYSMSWVRQAPGKGLELVA--QINSVGN--STYYPDTVKG
HC_Briakinumab_IgG1 QVQLVESGGGVVQPGRSLRLSCAASGFTF--SSYGMHWVRQAPGKGLEWVA--FIRYDGS--NKYYADSVKG
HC_Icrucumab_IgG1 QAQVVESGGGVVQSGRSLRLSCAASGFAF--SSYGMHWVRQAPGKGLEWVA--VIWYDGS--NKYYADSVRG
HC_Canakinumab_IgG1 QVQLVESGGGVVQPGRSLRLSCAASGFTF--SVYGMNWVRQAPGKGLEWVA--IIWYDGD--NQYYADSVKG
HC_Intetumumab_IgG1 QVQLVESGGGVVQPGRSRRLSCAASGFTF--SRYTMHWVRQAPGKGLEWVA--VISFDGS--NKYYVDSVKG
HC_Etaracizumab_IgG1 QVQLVESGGGVVQPGRSLRLSCAASGFTF--SSYDMSWVRQAPGKGLEWVA--KVSSGGG--STYYLDTVQG
HC_Lorvotuzumab_IgG1 QVQLVESGGGVVQPGRSLRLSCAASGFTF--SSFGMHWVRQAPGKGLEWVA--YISSGSF--TIYYADSVKG
HC_Farletuzumab_IgG1 EVQLVESGGGVVQPGRSLRLSCSASGFTF--SGYGLSWVRQAPGKGLEWVA--MISSGGS--YTYYADSVKG
HC_Mogamulizumab_IgG1 EVQLVESGGDLVQPGRSLRLSCAASGFIF--SNYGMSWVRQAPGKGLEWVA--TISSAST--YSYYPDSVKG
HC_Adalimumab_IgG1 EVQLVESGGGLVQPGRSLRLSCAASGFTF--DDYAMHWVRQAPGKGLEWVS--AITWNSG--HIDYADSVEG
HC_Sarilumab_IgG1 EVQLVESGGGLVQPGRSLRLSCAASRFTF--DDYAMHWVRQAPGKGLEWVS--GISWNSG--RIGYADSVKG
HC_Drozitumab_IgG1 EVQLVQSGGGVERPGGSLRLSCAASGFTF--DDYAMSWVRQAPGKGLEWVS--GINWQGG--STGYADSVKG
HC_Ramucirumab_IgG1 EVQLVQSGGGLVKPGGSLRLSCAASGFTF--SSYSMNWVRQAPGKGLEWVS--SISSSSS--YIYYADSVKG
HC_Robatumumab_IgG1 EVQLVQSGGGLVKPGGSLRLSCAASGFTF--SSFAMHWVRQAPGKGLEWIS--VIDTRG---ATYYADSVKG
HC_Sirukumab_IgG1 EVQLVESGGGLVQPGGSLRLSCAASGFTF--SPFAMSWVRQAPGKGLEWVA--KISPGGS--WTYYSDTVTG
HC_Denosumab_IgG2 EVQLLESGGGLVQPGGSLRLSCAASGFTF--SSYAMSWVRQAPGKGLEWVS--GITGSGG--STYYADSVKG
HC_Oxelumab_IgG1 EVQLLESGGGLVQPGGSLRLSCAASGFTF--NSYAMSWVRQAPGKGLEWVS--IISGSGG--FTYYADSVKG
HC_Alirocumab_IgG1 EVQLVESGGGLVQPGGSLRLSCAASGFTF--NNYAMNWVRQAPGKGLDWVS--TISGSGG--TTNYADSVKG
HC_Otelixizumab_IgG1 EVQLLESGGGLVQPGGSLRLSCAASGFTF--SSFPMAWVRQAPGKGLEWVS--TISTSGG--RTYYRDSVKG
HC_Gantenerumab_IgG1 QVELVESGGGLVQPGGSLRLSCAASGFTF--SSYAMSWVRQAPGKGLEWVS--AINASGT--RTYYADSVKG
HC_Crenezumab_IgG4 EVQLVESGGGLVQPGGSLRLSCAASGFTF--SSYGMSWVRQAPGKGLELVA--SINSNGG--STYYPDSVKG

FIG. 11

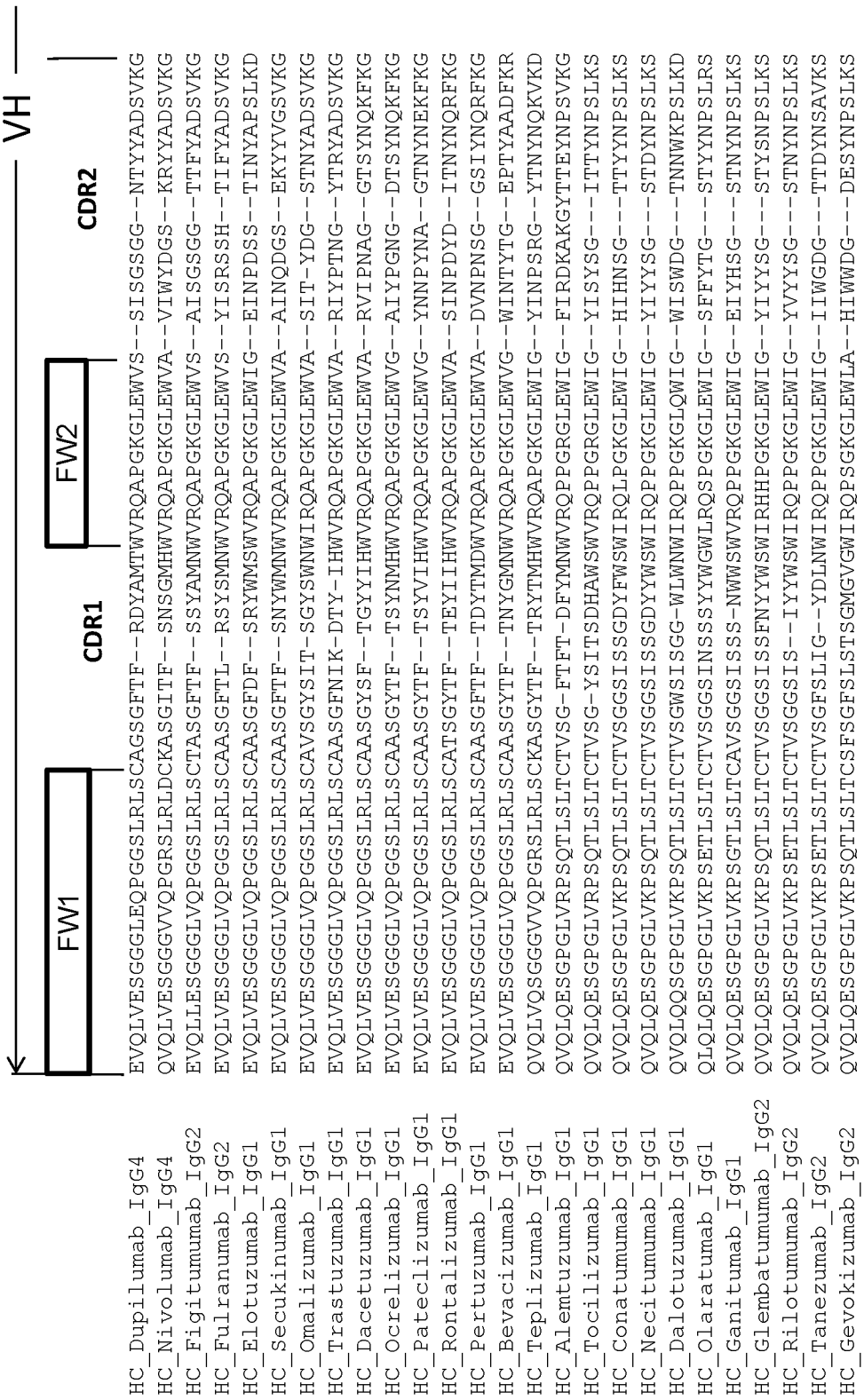
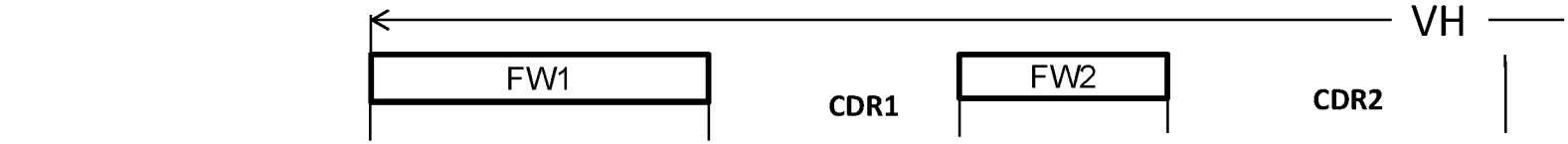


FIG. 11 continuation




```
HC_Tabalumab_IgG4       QVQLQQWGAGLLKPSETLSLTCAVYGGSFSG--YYWSWIRQPPGKGLEWIG--EINHSG---STNYNPSLKS
HC_Urelumab_IgG4        QVQLQQWGAGLLKPSETLSLTCAVYGGSFSG--YYWSWIRQSPEKGLEWIG--EINHGG---YVTYNPSLES
HC_Lebrikizumab_IgG4    QVTLRESGPALVKPTQTLTLTCTVSGFSLSA--YSVNWIRQPPGKALEWLA--MIWGDG---KIVYNSALKS
HC_Mepolizumab_IgG1     QVTLRESGPALVKPTQTLTLTCTVSGFSLTSY--SVHWVRQPPGKGLEWLGVSRIWASG---GTDYNSALM-
HC_Motavizumab_IgG1     QVTLRESGPALVKPTQTLTLTCTFSGFSLSTAGMSVGWIRQPPGKALEWL--ADIWWDD---KKHYNPSLKD
HC_Cetuximab_IgG1       QVQLKQSGPGLVQPSQSLSITCTVSGFSLTN--YGVHWVRQSPGKGLEWLG--VIWSGG---NTDYNTPFTS
HC_Bezlotoxumab_IgG1    EVQLVQSGAEVKKSGESLKISCKGSGYSF--TSYWIGWVRQMPGKGLEWMG--IFYPGDS--STRYSPSFQG
HC_Ustekinumab_IgG1     EVQLVQSGAEVKKPGESLKISCKGSGYSF--TTYWLGWVRQMPGKGLDWIG--IMSPVDS--DIRYSPSFQG
HC_Cantuzumab_IgG1      QVQLVQSGAEVKKPGETVKISCKASDYTF--TYYGMNWVKQAPGQGLKWMG--WIDTTTG--EPTYAQKFQG
HC_Milatuzumab_IgG1     QVQLQQSGSELKKPGASVKVSCKASGYTF--TNYGVNWIKQAPGQGLQWMG--WINPNTG--EPTFDDDFKG
HC_Epratuzumab_IgG1     QVQLVQSGAEVKKPGSSVKVSCKASGYTF--TSYWLHWVRQAPGQGLEWIG--YINPRND--YTEYNQNFKD
HC_Veltuzumab_IgG1      QVQLQQSGAEVKKPGSSVKVSCKASGYTF--TSYNMHWVKQAPGQGLEWIG--AIYPGMG--DTSYNQKFKG
HC_Daclizumab_IgG1      QVQLVQSGAEVKKPGSSVKVSCKASGYTF--TSYRMHWVRQAPGQGLEWIG--YINPSTG--YTEYNQKFKD
HC_Clivatuzumab_IgG1    QVQLQQSGAEVKKPGASVKVSCEASGYTF--PSYVLHWVKQAPGQGLEWIG--YINPYND--GTQYNEKFKG
HC_Benralizumab_IgG1    EVQLVQSGAEVKKPGASVKVSCKASGYTF--TSYVIHWVRQRPGQGLAWMG--YINPYND--GTKYNERFKG
HC_Ficlatuzumab_IgG1    _VQLVQPGAEVKKPGTSVKLSCKASGYTF--TTYWMHWVRQAPGQGLEWIG--EINPTNG--HTNYNQKFQG
HC_Vedolizumab_IgG1     QVQLVQSGAEVKKPGASVKVSCKGSGYTF--TSYWMHWVRQAPGQRLEWIG--EIDPSES--NTNYNQKFKG
HC_Bimagrumab_IgG1      QVQLVQSGAEVKKPGASVKVSCKASGYTF--TSSYINWVRQAPGQGLEWMG--TINPVSG--STSYAQKFQG
HC_Fezakinumab_IgG1     QVQLVQSGAEVKKPGASVKVSCKASGYTF--TNYYMHWVRQAPGQGLEWVG--WINPYTG--SAFYAQKFRG
HC_Sifalimumab_IgG1     QVQLVQSGAEVKKPGASVKVSCKASGYTF--TSYSISWVRQAPGQGLEWMG--WISVYNG--NTNYAQKFQG
HC_Carlumab_IgG1        QVQLVQSGAEVKKPGSSVKVSCKASGGTF--SSYGISWVRQAPGQGLEWMG--GIIPIFG--TANYAQKFQG
HC_Cixutumumab_IgG1     EVQLVQSGAEVKKPGSSVKVSCKASGGTF--SSYAISWVRQAPGQGLEWMG--GIIPIFG--TANYAQKFQG
HC_Enokizumab_IgG1      QVQLVQSGAEVKKPGSSVKVSCKASGGTF--SYYWIEWVRQAPGQGLEWMG--EILPGSG--TTNPNEKFKG
HC_Obinutuzumab_IgG1    QVQLVQSGAEVKKPGSSVKVSCKASGYAF--SYSWINWVRQAPGQGLEWMG--RIFPGDG--DTDYNGKFKG
HC_Tildrakizumab_IgG1   QVQLVQSGAEVKKPGASVKVSCKASGYIF--ITYWMTWVRQAPGQGLEWMG--QIFPASG--SADYNEKFEG
HC_Amatuximab_IgG1      QVQLQQSGPELEKPGASVKISCKASGYSF--TGYTMNWVKQSHGKSLEWIG--LITPYNG--ASSYNQKFRG
```

FIG. 11 continuation

SUBSTITUTE SHEET (RULE 26)

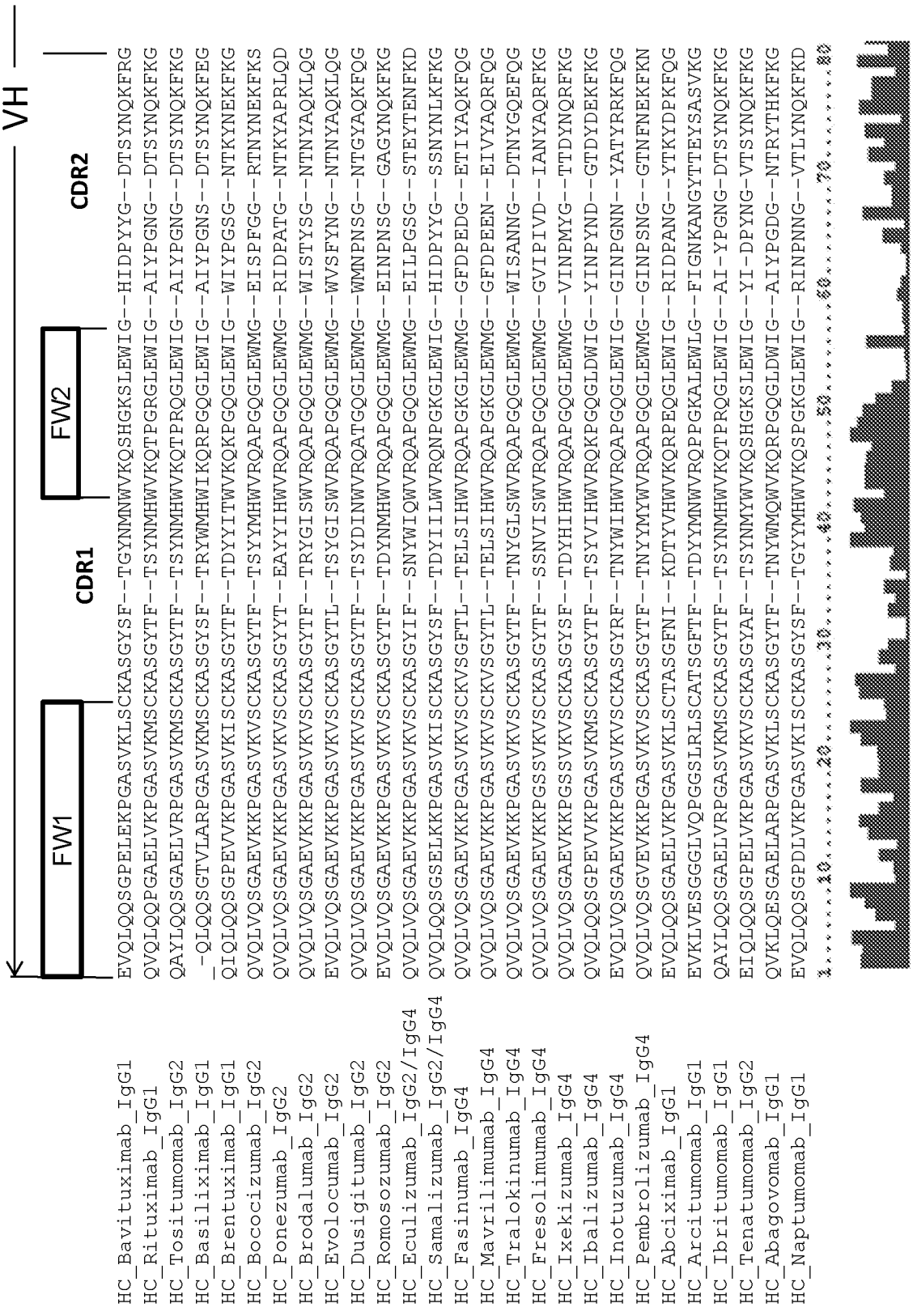


FIG. 11 continuation

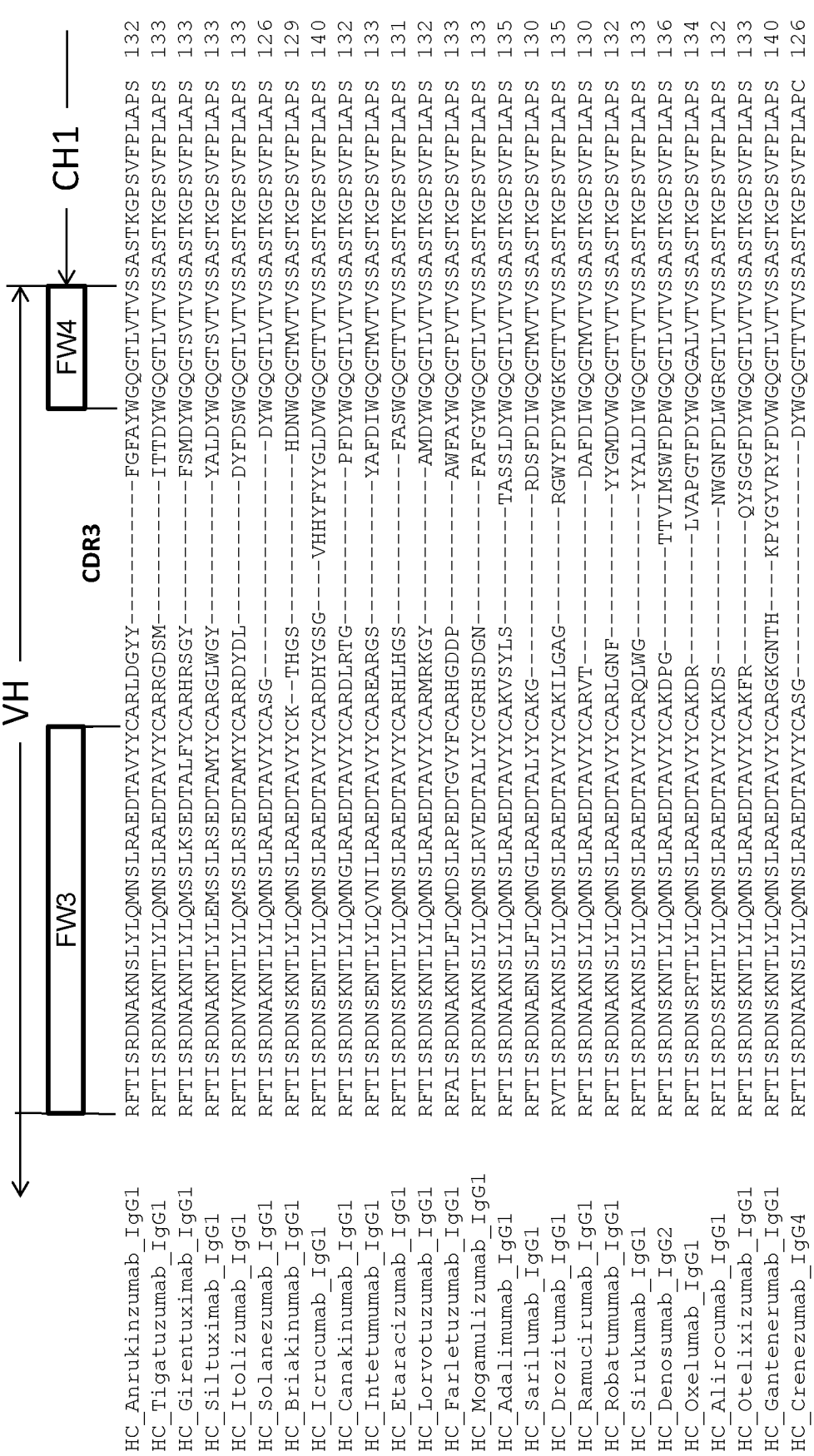


FIG. 11 continuation

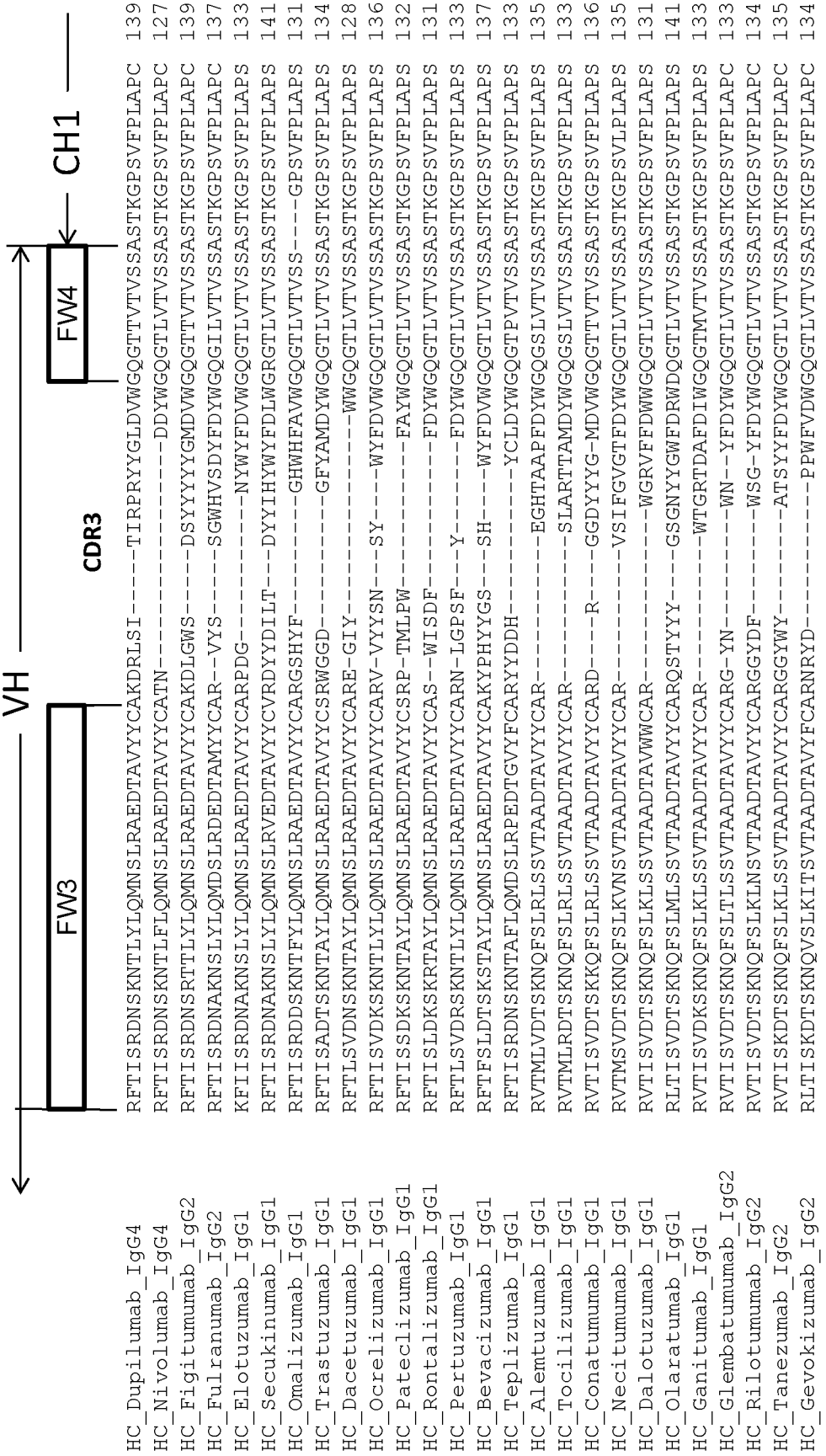


FIG. 11 continuation

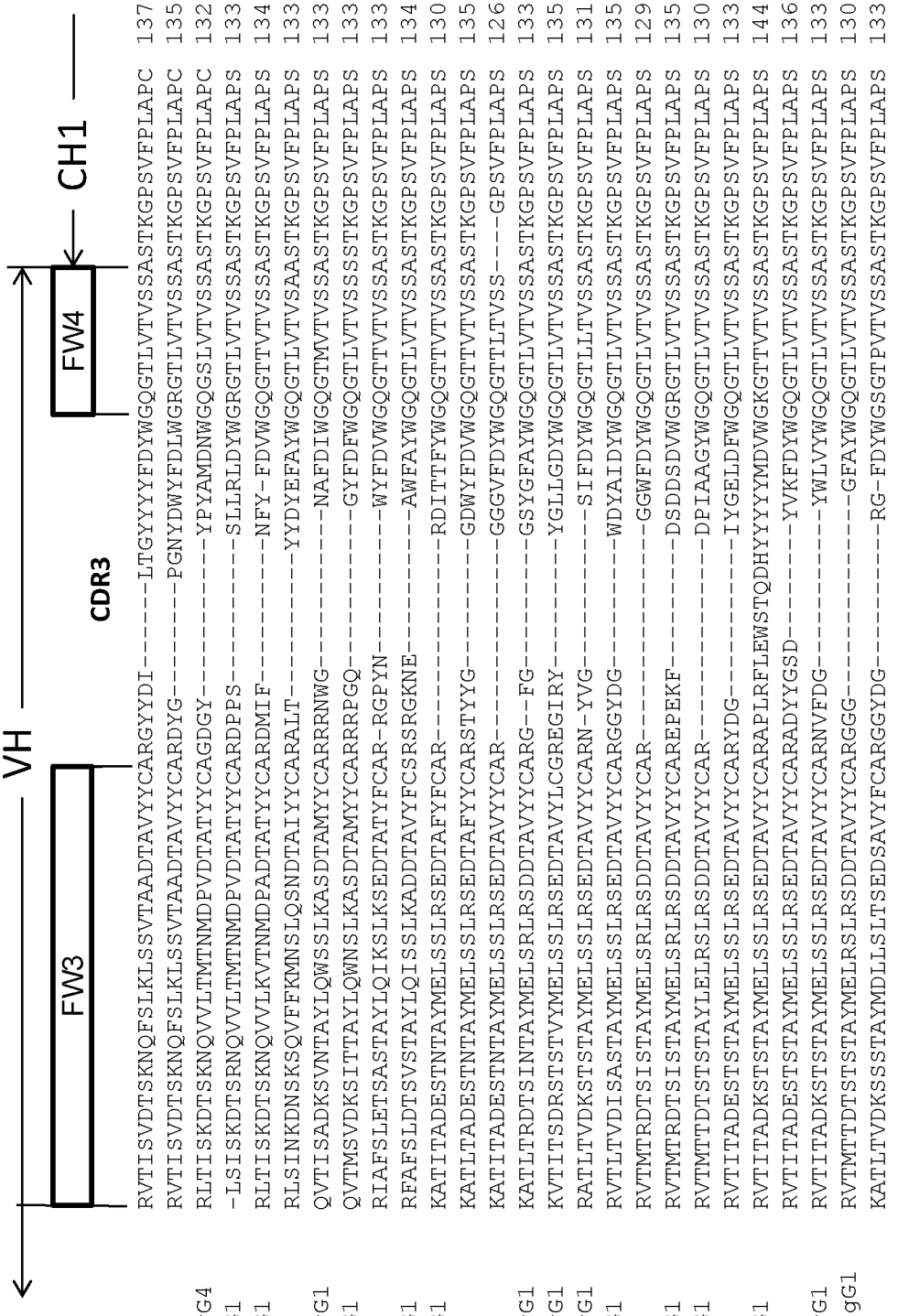


FIG. 11 continuation

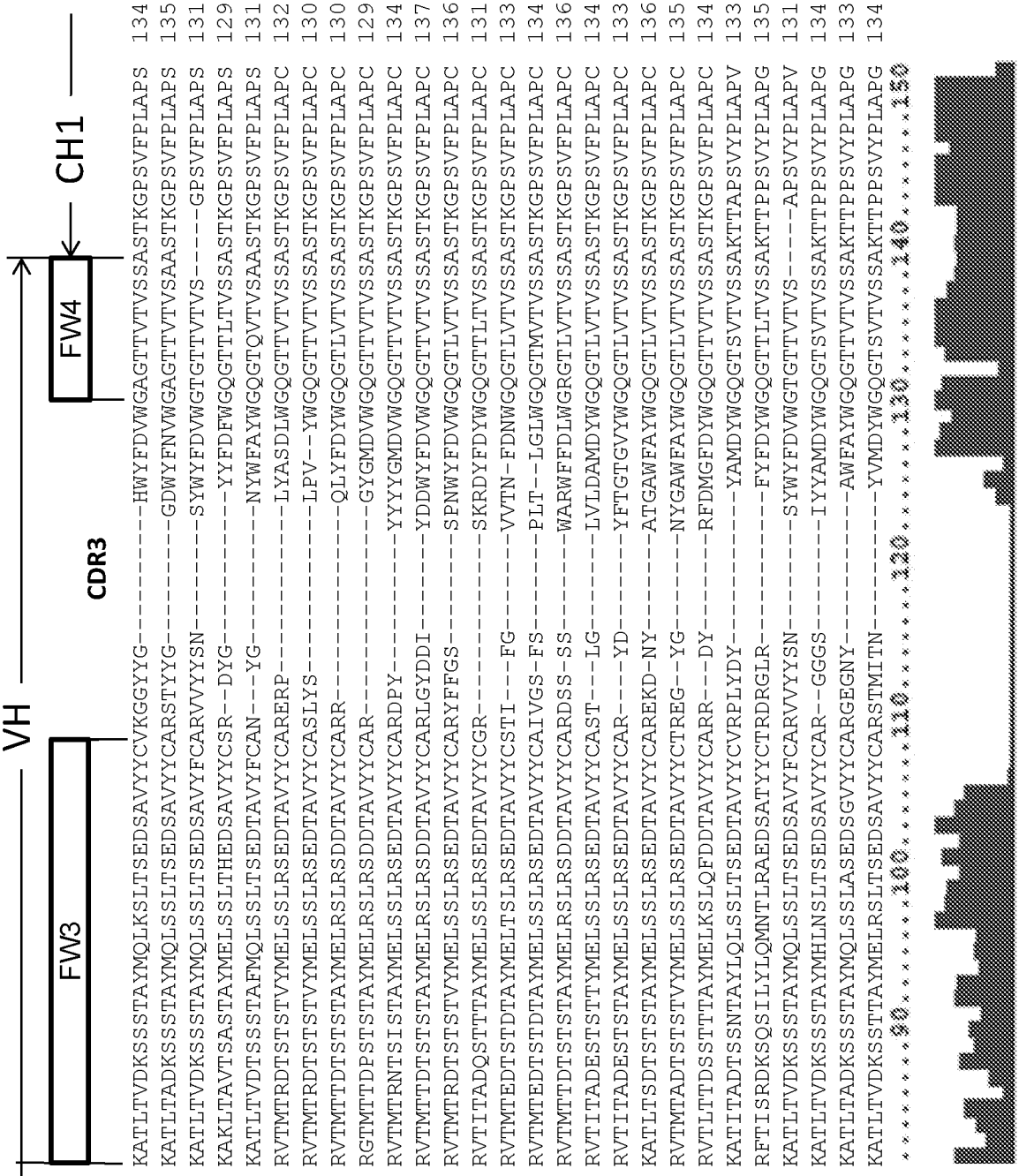


FIG. 11 continuation

CH1

```
HC_Anrukinzumab_IgG1    SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Tigatuzumab_IgG1     SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Girentuximab_IgG1    SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Siltuximab_IgG1      SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Itolizumab_IgG1      SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Solanezumab_IgG1     SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Briakinumab_IgG1     SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Icrucumab_IgG1       SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Canakinumab_IgG1     SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Intetumumab_IgG1     SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Etaracizumab_IgG1    SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Lorvotuzumab_IgG1    SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Farletuzumab_IgG1    SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Mogamulizumab_IgG1   SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Adalimumab_IgG1      SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Sarilumab_IgG1       SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Drozitumab_IgG1      SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Ramucirumab_IgG1     SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Robatumumab_IgG1     SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Sirukumab_IgG1       SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Denosumab_IgG2       SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Oxelumab_IgG1        SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Alirocumab_IgG1      SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Otelixizumab_IgG1    SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Gantenerumab_IgG1    SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Crenezumab_IgG4      SRSTSESTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTKTYTC
```

FIG. 11 continuation

CH1

```
HC_Dupilumab_IgG4          SRSTSESTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTKTYTC
HC_Nivolumab_IgG4          SRSTSESTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTKTYTC
HC_Figitumumab_IgG2        SRSTSESTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSNFGTQTYTC
HC_Fulranumab_IgG2         SRSTSESTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSNFGTQTYTC
HC_Elotuzumab_IgG1         SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Secukinumab_IgG1        SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Omalizumab_IgG1         SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Trastuzumab_IgG1        SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Dacetuzumab_IgG1        SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Ocrelizumab_IgG1        SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Pateclizumab_IgG1       SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Rontalizumab_IgG1       SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Pertuzumab_IgG1         SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Bevacizumab_IgG1        SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Teplizumab_IgG1         SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Alemtuzumab_IgG1        SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Tocilizumab_IgG1        SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Conatumumab_IgG1        SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Necitumumab_IgG1        SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Dalotuzumab_IgG1        SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Olaratumab_IgG1         SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Ganitumab_IgG1          SKSTSGGTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYIC
HC_Glembatumumab_IgG2      SRSTSESTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSNFGTQTYTC
HC_Rilotumumab_IgG2        SRSTSESTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSNFGTQTYTC
HC_Tanezumab_IgG2          SRSTSESTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSNFGTQTYTC
HC_Gevokizumab_IgG2        SRSTSESTAALGCLVKD------YFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVTSSNFGTQTYTC
```

FIG. 11 continuation

FIG. 11 continuation

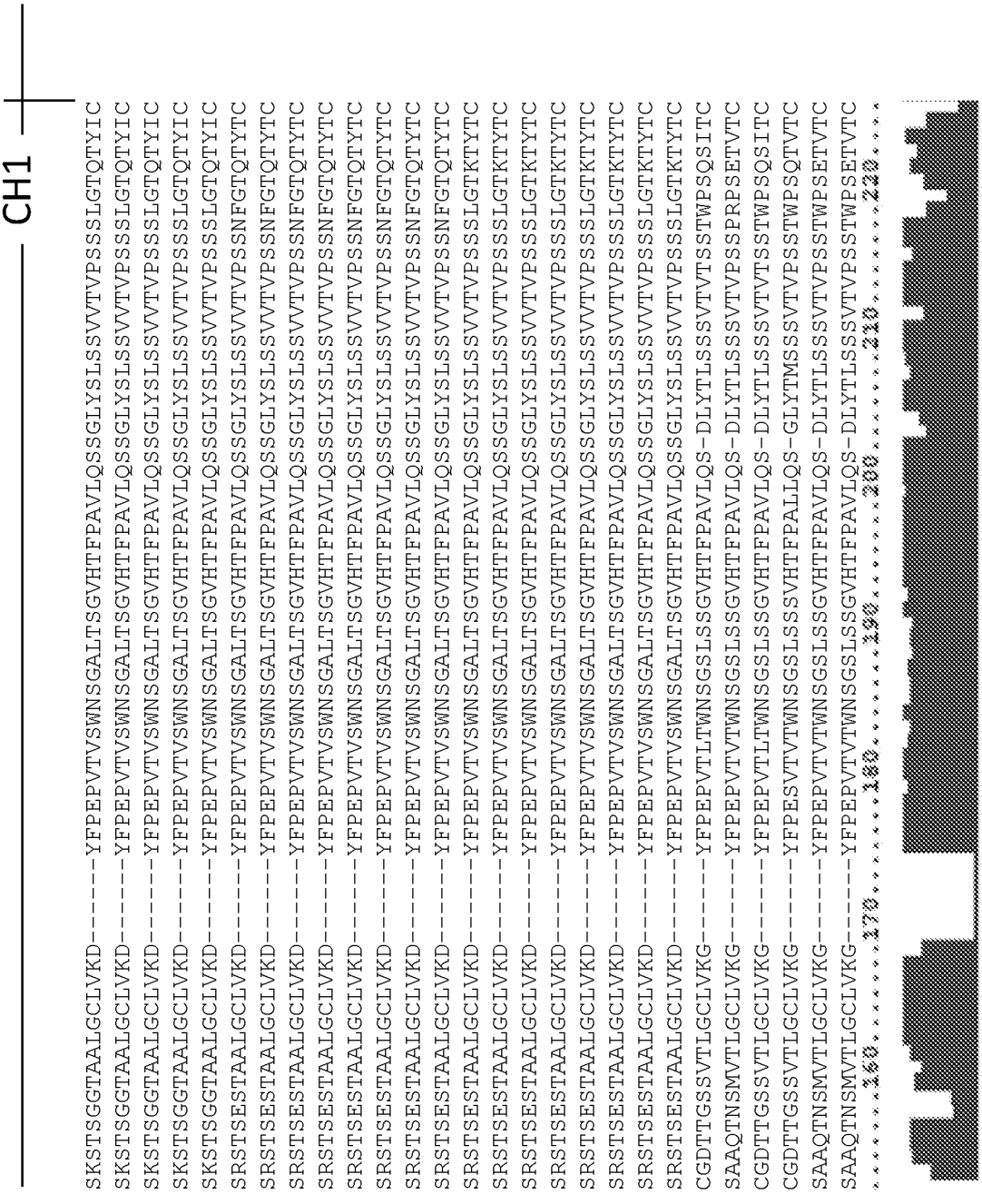


FIG. 11 continuation

FIG. 11 continuation

FIG. 11 continuation

FIG. 11 continuation

FIG. 11 continuation

FIG. 11 continuation

CH2 CH3

HC_Dupilumab_IgG4	EDP-EVQFNWYVDGVEVHNAKTKPREEQFNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKGLPSSIEKTISKAK
HC_Nivolumab_IgG4	EDP-EVQFNWYVDGVEVHNAKTKPREEQFNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKGLPSSIEKTISKAK
HC_Figitumumab_IgG2	EDP-EVQFNWYVDGVEVHNAKTKPREEQFNSTFRVVSVLTVVHQDWLNGKEYKCKVSNKGLPAPIEKTISKTK
HC_Fulranumab_IgG2	EDP-EVQFNWYVDGVEVHNAKTKPREEQFNSTFRVVSVLTVVHQDWLNGKEYKCKVSNKGLPAPIEKTISKTK
HC_Elotuzumab_IgG1	EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Secukinumab_IgG1	EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Omalizumab_IgG1	EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Trastuzumab_IgG1	EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Dacetuzumab_IgG1	EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Ocrelizumab_IgG1	EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Pateclizumab_IgG1	EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Rontalizumab_IgG1	EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Pertuzumab_IgG1	EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Bevacizumab_IgG1	EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Teplizumab_IgG1	EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Alemtuzumab_IgG1	EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Tocilizumab_IgG1	EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Conatumumab_IgG1	EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Necitumumab_IgG1	EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Dalotuzumab_IgG1	EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Olaratumab_IgG1	EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Ganitumab_IgG1	EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Glembatumumab_IgG2	EDP-EVQFNWYVDGVEVHNAKTKPREEQFNSTFRVVSVLTVVHQDWLNGKEYKCKVSNKGLPAPIEKTISKTK
HC_Rilotumumab_IgG2	EDP-EVQFNWYVDGVEVHNAKTKPREEQFNSTFRVVSVLTVVHQDWLNGKEYKCKVSNKGLPAPIEKTISKTK
HC_Tanezumab_IgG2	EDP-EVQFNWYVDGVEVHNAKTKPREEQFNSTFRVVSVLTVVHQDWLNGKEYKCKVSNKGLPSSIEKTISKTK
HC_Gevokizumab_IgG2	EDP-EVQFNWYVDGMEVHNAKTKPREEQFNSTFRVVSVLTVVHQDWLNGKEYKCKVSNKGLPAPIEKTISKTK

FIG. 11 continuation

CH2 CH3

```
HC_Tabalumab_IgG4          EDP-EVQFNWYVDGVEVHNAKTKPREEQFNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKGLPSSIEKTISKAK
HC_Urelumab_IgG4           EDP-EVQFNWYVDGVEVHNAKTKPREEQFNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKGLPSSIEKTISKAK
HC_Lebrikizumab_IgG4       EDP-EVQFNWYVDGVEVHNAKTKPREEQFNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKGLPSSIEKTISKAK
HC_Mepolizumab_IgG1        EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Motavizumab_IgG1        EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Cetuximab_IgG1          EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Bezlotoxumab_IgG1       EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Ustekinumab_IgG1        EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Cantuzumab_IgG1         EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Milatuzumab_IgG1        EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Epratuzumab_IgG1        EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Veltuzumab_IgG1         EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Daclizumab_IgG1         EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Clivatuzumab_IgG1       EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Benralizumab_IgG1       EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Ficlatuzumab_IgG1       EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Vedolizumab_IgG1        EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Bimagrumab_IgG1         EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Fezakinumab_IgG1        EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Sifalimumab_IgG1        EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Carlumab_IgG1           EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Cixutumumab_IgG1        EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Enokizumab_IgG1         EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Obinutuzumab_IgG1       EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Tildrakizumab_IgG1      EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
HC_Amatuximab_IgG1         EDP-EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
```

FIG. 11 continuation

FIG. 11 continuation

CH3

```
HC_Anrukinzumab_IgG1      GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 415
HC_Tigatuzumab_IgG1       GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 416
HC_Girentuximab_IgG1      GQPREPQVYTLP-PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 416
HC_Siltuximab_IgG1        GQPREPQVYTLP-PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 416
HC_Itolizumab_IgG1        GQPREPQVYTLP-PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 416
HC_Solanezumab_IgG1       GQPREPQVYTLP-PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 409
HC_Briakinumab_IgG1       GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 412
HC_Icrucumab_IgG1         GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 423
HC_Canakinumab_IgG1       GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 415
HC_Intetumumab_IgG1       GQPREPQVYTLP-PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 416
HC_Etaracizumab_IgG1      GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 414
HC_Lorvotuzumab_IgG1      GQPREPQVYTLP-PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 415
HC_Farletuzumab_IgG1      GQPREPQVYTLP-PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 416
HC_Mogamulizumab_IgG1     GQPREPQVYTLP-PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 416
HC_Adalimumab_IgG1        GQPREPQVYTLP-PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 418
HC_Sarilumab_IgG1         GQPREPQVYTLP-PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 413
HC_Drozitumab_IgG1        GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 418
HC_Ramucirumab_IgG1       GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 413
HC_Robatumumab_IgG1       GQPREPQVYTLP-PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 415
HC_Sirukumab_IgG1         GQPREPQVYTLP-PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 416
HC_Denosumab_IgG2         GQPREPQVYTLP-PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 419
HC_Oxelumab_IgG1          GQPREPQVYTLP-PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 417
HC_Alirocumab_IgG1        GQPREPQVYTLP-PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 415
HC_Otelixizumab_IgG1      GQPREPQVYTLP-PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 416
HC_Gantenerumab_IgG1      GQPREPQVYTLP-PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 423
HC_Crenezumab_IgG4        GQPREPQVYTLP-PSQEEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--RLTVDK 406
```

FIG. 11 continuation

CH3

```
HC_Dupilumab_IgG4          GQPREPQVYTLP-PSQEEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--RLTVDK 419
HC_Nivolumab_IgG4          GQPREPQVYTLP-PSQEEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--RLTVDK 407
HC_Figitumumab_IgG2        GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPMLDSDGSFFLYS--KLTVDK 418
HC_Fulranumab_IgG2         GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPMLDSDGSFFLYS--KLTVDK 416
HC_Elotuzumab_IgG1         GQPREPQVYTLP-PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 416
HC_Secukinumab_IgG1        GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 424
HC_Omalizumab_IgG1         GQPREPQVYTLP-PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 414
HC_Trastuzumab_IgG1        GQPREPQVYTLP-PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 418
HC_Dacetuzumab_IgG1        GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 411
HC_Ocrelizumab_IgG1        GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 419
HC_Pateclizumab_IgG1       GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 415
HC_Rontalizumab_IgG1       GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 414
HC_Pertuzumab_IgG1         GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 416
HC_Bevacizumab_IgG1        GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 420
HC_Teplizumab_IgG1         GQPREPQVYTLP-PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 416
HC_Alemtuzumab_IgG1        GQPREPQVYTLP-PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 418
HC_Tocilizumab_IgG1        GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 416
HC_Conatumumab_IgG1        GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 419
HC_Necitumumab_IgG1        GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 418
HC_Dalotuzumab_IgG1        GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 414
HC_Olaratumab_IgG1         GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 424
HC_Ganitumab_IgG1          GQPREPQVYTLP-PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 416
HC_Glembatumumab_IgG2      GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPMLDSDGSFFLYS--KLTVDK 412
HC_Rilotumumab_IgG2        GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPMLDSDGSFFLYS--KLTVDK 413
HC_Tanezumab_IgG2          GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPMLDSDGSFFLYS--KLTVDK 414
HC_Gevokizumab_IgG2        GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPMLDSDGSFFLYS--KLTVDK 413
```

FIG. 11 continuation

CH3

```
HC_Tabalumab_IgG4          GQPREPQVYTLP-PSQEEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--RLTVDK 417
HC_Urelumab_IgG4           GQPREPQVYTLP-PSQEEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--RLTVDK 415
HC_Lebrikizumab_IgG4       GQPREPQVYTLP-PSQEEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--RLTVDK 412
HC_Mepolizumab_IgG1        GQLP-PSRDELPREPQVYTTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 416
HC_Motavizumab_IgG1        GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 417
HC_Cetuximab_IgG1          GQPREPQVYTLP-PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 419
HC_Bezlotoxumab_IgG1       GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 416
HC_Ustekinumab_IgG1        GQPREPQVYTLP-PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 416
HC_Cantuzumab_IgG1         GQPREPQVYTLP-PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 416
HC_Milatuzumab_IgG1        GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 417
HC_Epratuzumab_IgG1        GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 413
HC_Veltuzumab_IgG1         GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 418
HC_Daclizumab_IgG1         GQPREPQVYTLP-PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 409
HC_Clivatuzumab_IgG1       GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 416
HC_Benralizumab_IgG1       GQPREPQVYTLP-PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 418
HC_Ficlatuzumab_IgG1       GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 414
HC_Vedolizumab_IgG1        GQPREPQVYTLP-PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 418
HC_Bimagrumab_IgG1         GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 412
HC_Fezakinumab_IgG1        GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 418
HC_Sifalimumab_IgG1        GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 413
HC_Carlumab_IgG1           GQPREPQVYTLP-PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 416
HC_Cixutumumab_IgG1        GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 427
HC_Enokizumab_IgG1         GQPREPQVYTLP-PSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 419
HC_Obinutuzumab_IgG1       GQPREPQVYTLP-PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 416
HC_Tildrakizumab_IgG1      GQPREPQVYTLP-PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 413
HC_Amatuximab_IgG1         GQPREPQVYTLP-PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYS--KLTVDK 416
```

FIG. 11 continuation

FIG. 11 continuation

CH3 →

HC_Anrukinzumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--	448
HC_Tigatuzumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--	449
HC_Girentuximab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--	449
HC_Siltuximab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--	449
HC_Itolizumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--	449
HC_Solanezumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--	442
HC_Briakinumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--	445
HC_Icrucumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--	456
HC_Canakinumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--	448
HC_Intetumumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--	449
HC_Etaracizumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--	447
HC_Lorvotuzumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--	448
HC_Farletuzumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--	449
HC_Mogamulizumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--	449
HC_Adalimumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--	451
HC_Sarilumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--	446
HC_Drozitumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--	451
HC_Ramucirumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--	446
HC_Robatumumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--	448
HC_Sirukumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--	449
HC_Denosumab_IgG2	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--	452
HC_Oxelumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPG----	449
HC_Alirocumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPG----	447
HC_Otelixizumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--	449
HC_Gantenerumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--	456
HC_Crenezumab_IgG4	SRWQEGNVFSCSVMHEALHN-HYTQKSLSLSLG----	438

FIG. 11 continuation

CH3

```
HC_Dupilumab_IgG4          SRWQEGNVFSCSVMHEALHN-HYTQKSLSLSLG---  451
HC_Nivolumab_IgG4          SRWQEGNVFSCSVMHEALHN-HYTQKSLSLSLGK--  440
HC_Figitumumab_IgG2        SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPG----  450
HC_Fulranumab_IgG2         SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--  449
HC_Elotuzumab_IgG1         SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--  449
HC_Secukinumab_IgG1        SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--  457
HC_Omalizumab_IgG1         SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--  447
HC_Trastuzumab_IgG1        SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--  451
HC_Dacetuzumab_IgG1        SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--  444
HC_Ocrelizumab_IgG1        SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--  452
HC_Pateclizumab_IgG1       SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPG----  447
HC_Rontalizumab_IgG1       SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--  447
HC_Pertuzumab_IgG1         SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPG----  448
HC_Bevacizumab_IgG1        SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--  453
HC_Teplizumab_IgG1         SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--  449
HC_Alemtuzumab_IgG1        SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--  451
HC_Tocilizumab_IgG1        SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--  449
HC_Conatumumab_IgG1        SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--  452
HC_Necitumumab_IgG1        SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--  451
HC_Dalotuzumab_IgG1        SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--  447
HC_Olaratumab_IgG1         SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--  457
HC_Ganitumab_IgG1          SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--  449
HC_Glembatumumab_IgG2      SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--  445
HC_Rilotumumab_IgG2        SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--  446
HC_Tanezumab_IgG2          SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK--  447
HC_Gevokizumab_IgG2        SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPG---  445
```

FIG. 11 continuation

CH3 →

HC_Tabalumab_IgG4	SRWQEGNVFSCSVMHEALHN-HYTQKSLSLSLGK---	450
HC_Urelumab_IgG4	SRWQEGNVFSCSVMHEALHN-HYTQKSLSLSLGK---	448
HC_Lebrikizumab_IgG4	SRWQEGNVFSCSVMHEALHN-HYTQKSLSLSLG----	444
HC_Mepolizumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK---	449
HC_Motavizumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK---	450
HC_Cetuximab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK---	452
HC_Bezlotoxumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK---	449
HC_Ustekinumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK---	449
HC_Cantuzumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK---	449
HC_Milatuzumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK---	450
HC_Epratuzumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK---	446
HC_Veltuzumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK---	451
HC_Daclizumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK---	442
HC_Clivatuzumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK---	449
HC_Benralizumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK---	451
HC_Ficlatuzumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK---	447
HC_Vedolizumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK---	451
HC_Bimagrumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK---	445
HC_Fezakinumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPG----	450
HC_Sifalimumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK---	446
HC_Carlumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK---	449
HC_Cixutumumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK---	460
HC_Enokizumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK---	452
HC_Obinutuzumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK---	449
HC_Tildrakizumab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK---	446
HC_Amatuximab_IgG1	SRWQQGNVFSCSVMHEALHN-HYTQKSLSLSPGK---	449

FIG. 11 continuation

FIG. 11 continuation

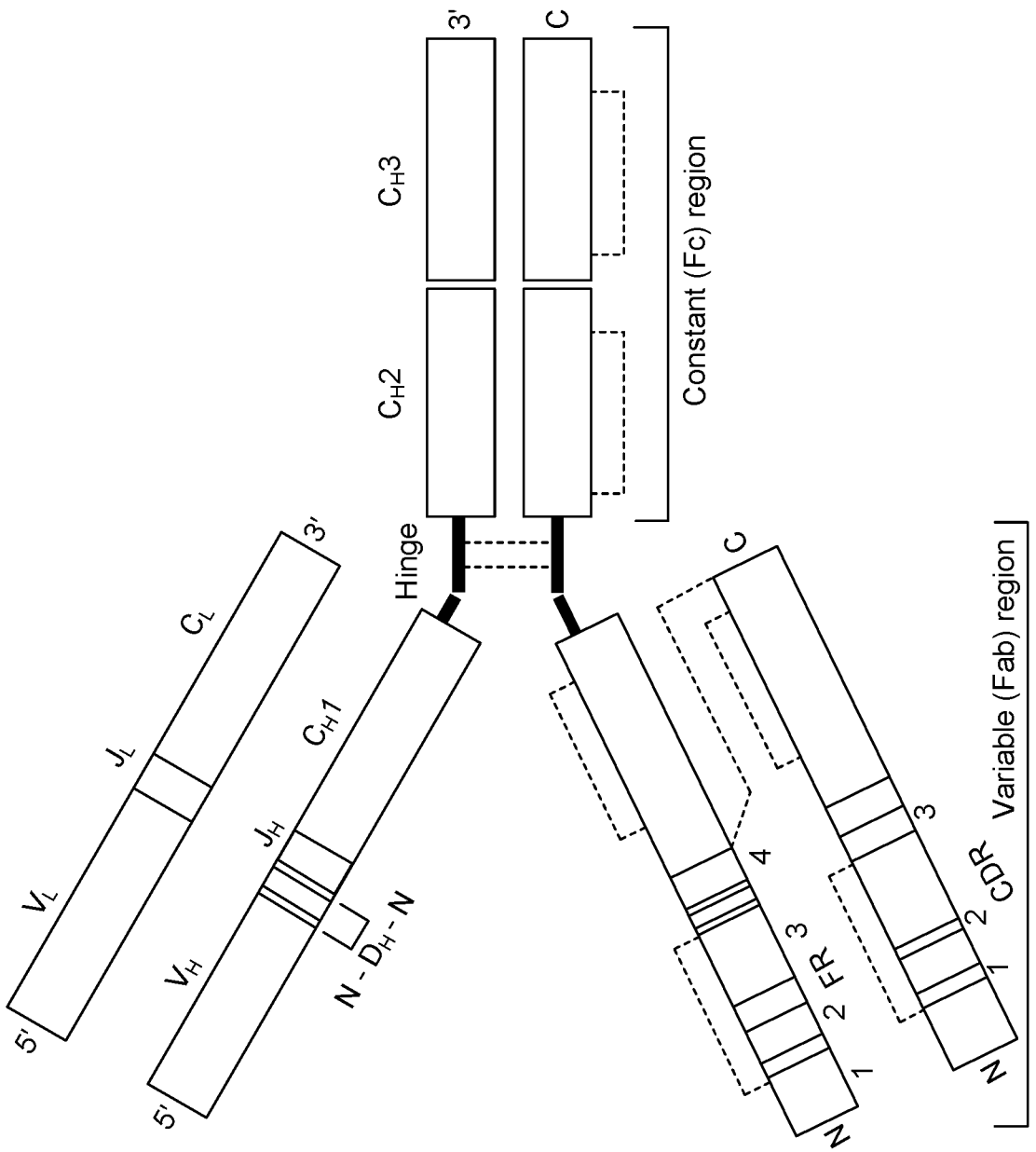

3'
C
C_H3
C_H2
Constant (Fc) region
Hinge
C_L
J_L
V_L
5'
3'
C_H1
J_H
V_H
5'
N - D_H - N
C
4
3
2 FR
1
N
3
2 CDR
1
N
Variable (Fab) region

FIG. 12

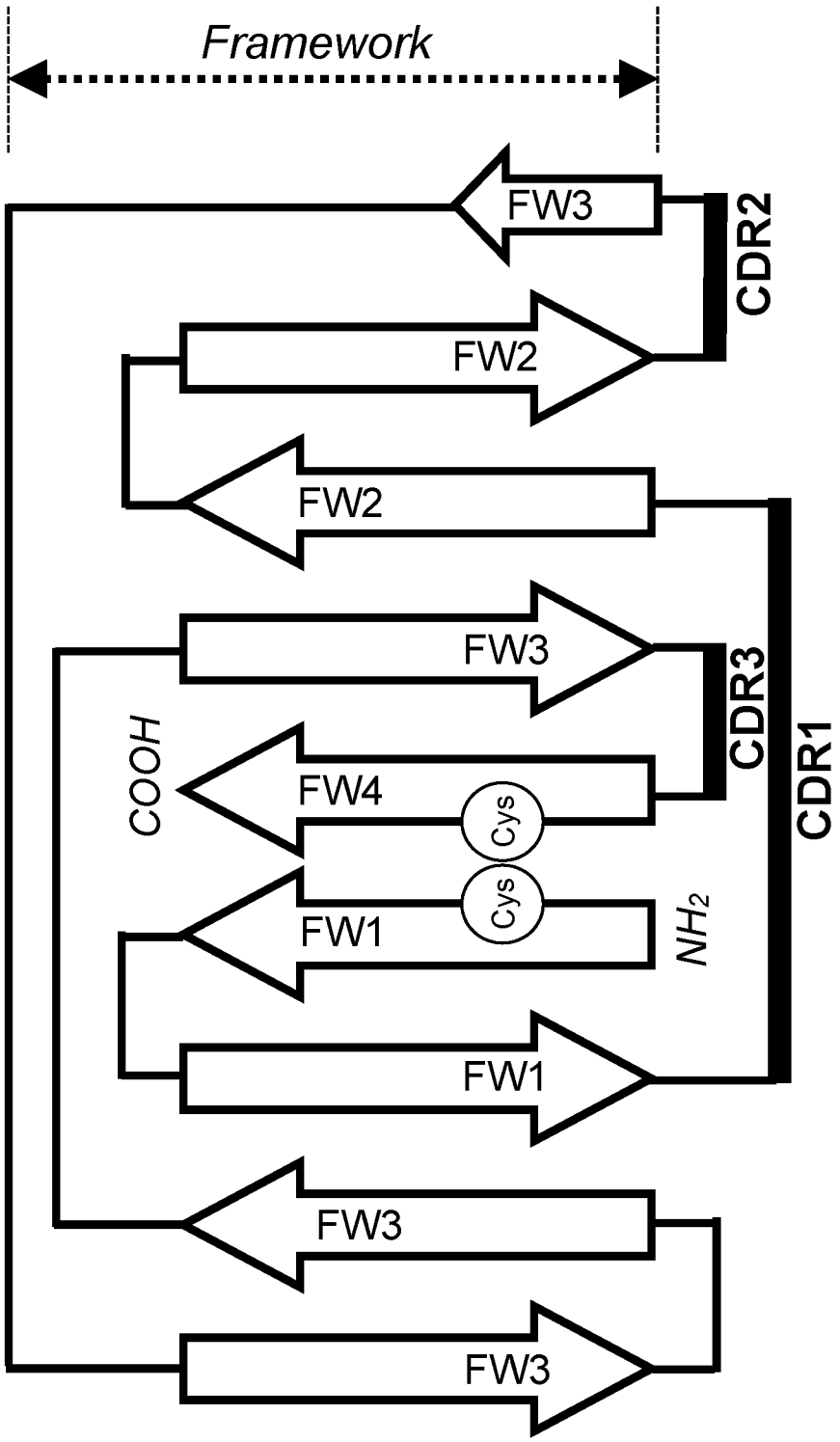

Framework
FW3
CDR2
FW2
FW2
FW3
CDR3
CDR1
COOH
FW4
Cys
Cys
NH2
FW1
FW1
FW3
FW3

FIG. 13

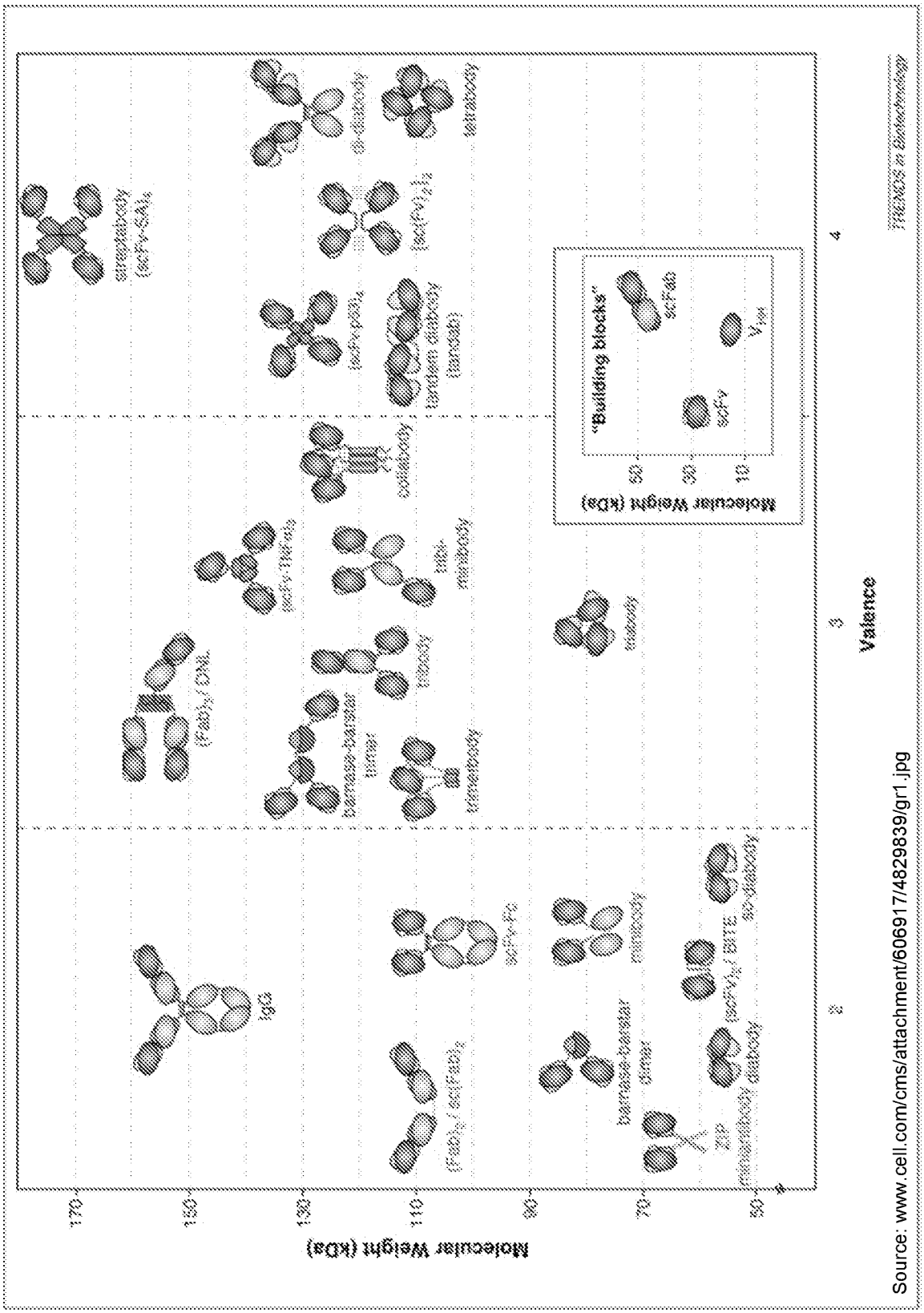

Molecular Weight (kDa)
170
150
130
110
90
70
50
IgG
(Fab)2 / scFab2
scFv-Fc
barnase-barstar dimer
minibody
ZIP
(scFv)2 / BiTE
sc-diabody
miniantibody diabody
(Fab)2 / DNL
barnase-barstar trimer
scFv-Trimer
triabody
triple-antibody
collabody
streptabody (scFv-SA)4
(scFv-pDC)4
tandem diabody (tandab)
sc-diabody
(scFv)2-L2
tetrabody
"Building blocks"
scFab
scFv
VHH
Molecular Weight (kDa)
50
30
10
2
3
4
Valance
TRENDS in Biotechnology
Source: www.cell.com/cms/attachment/606917/4829839/gr1.jpg

FIG. 14