(12) STANDARD PATENT (11) Application No. **AU 2018367622 B2**
(19) AUSTRALIAN PATENT OFFICE

(54) Title
mRNA display antibody library and methods

(51) International Patent Classification(s)
C07K 16/30 (2006.01) ***C12N 15/10*** (2006.01)

(21) Application No: **2018367622** (22) Date of Filing: **2018.11.16**

(87) WIPO No: **WO19/099882**

(30) Priority Data

(31) Number	(32) Date	(33) Country
62/588,914	**2017.11.20**	**US**

(43) Publication Date: **2019.05.23**
(44) Accepted Journal Date: **2022.12.22**

(71) Applicant(s)
NantBio, Inc.

(72) Inventor(s)
OLSON, Anders

(74) Agent / Attorney
FB Rice Pty Ltd, L 33 477 Collins Street, Melbourne, VIC, 3000, AU

(56) Related Art
TILLER, T. et al., "A fully synthetic human Fab antibody library based on fixed VH/VL framework pairings with favorable biophysical properties", mAbs, (2013-05), vol. 5, no. 3, pages 445 - 470
CHEN WEIZAO ET AL " Construction of a human antibody domain (VH) library"; Methods Mol Biol. 2009 ; 525: 81–xiii. doi:10.1007/978-1-59745-554-1_4.

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
23 May 2019 (23.05.2019)

WIPO | PCT

(10) International Publication Number
WO 2019/099882 A3

(51) International Patent Classification:
C07K 16/30 (2006.01) *C12N 15/10* (2006.01)

(21) International Application Number:
PCT/US2018/061592

(22) International Filing Date:
16 November 2018 (16.11.2018)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/588,914 20 November 2017 (20.11.2017) US

(71) Applicant: **NANTBIO, INC.** [US/US]; 9920 Jefferson Blvd., Culver City, California 90232 (US).

(72) Inventor: **OLSON, Anders**; 9920 Jefferson Blvd., Culver City, California 90232 (US).

(74) Agent: **FESSENMAIER, Martin**; 1920 Main Street, Suite 750, Irvine, California 92614 (US).

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

WO 2019/099882 A3

(54) Title: mRNA DISPLAY ANTIBODY LIBRARY AND METHODS

PDB#	VH type	CDR1	CDR2	CDR-H3	H3 length	κ type	CDRL3	κ55
2qsc	3-64	FNFSSYVMH	SAISSDGETTY	R[illegible]Y-NAFDV	14	k1-33	QQYDNLGDLSF	[illegible]55
4zyp	3-30	FSFSHYAMH	AVISYDGENTY	R[illegible]I-YGMDV	14	k1-33	QQYDNLPPLTF	[illegible]55
4zs6	3-11	FSLSDYYMN	AYISSSSGYTN	R[illegible]D-GYYKH	12	k1-33	QQYDKL--PTF	[illegible]55
1fgm	3-23	FTFNIFVMS	SGVFGSGGNTD	KH[illegible]V-TGFDS	12	k1-33	QQYQNLP-LTF	[illegible]55
1dfb	3-9	FTFNDYAMH	SGISWDSSSIG	KG[illegible]D-VAFDI	17	k1-5	QQYNSY--SF	[illegible]55
4ers	3-30	FTFSSYGMH	AVMYYDGSNKD	R[illegible]D-YGLDV	19	k1-17	LQHNSNP-LTF	[illegible]55
4u6v	3-13	FTFSSHDMH	GIGTAG-DTYY	R[illegible]Y-YGMDV	14	k1-5	KQYADYW--TF	[illegible]55
1dee	3-30	FTFSGYGMH	ALISYDESNKY	KV[illegible]F-APNDY	12	k1-39	QQSYSAP-RTF	Q55
2r56	3-23	FTFRHHGMT	ASLSGSGTKTH	KA[illegible]R-GYFDL	12	k1-16	QQYHSYP-WTF	Q55
2uzi	3-48	FTFSTFSMN	SYISRTSKTIY	RGR[illegible]FFDY	6	k1-39	QQSVMIP-MTF	Q55
1ddj	3-30	FTFSSYAMH	AVISSDGGNKY	RGNP-GGGDY	14	k1-33	QQNSNW--TF	Q55
4py8	3-30	FTFRMYATH	ALISYDGSNKY	R[illegible]LG-GIMDV	13	k1-12	QQANSFP-LTF	Q55
3idx	3-30	FTFRNYAMH	ALIKYDGRNKY	R[illegible]G-YGPDY	19	k1-33	QQHQNVPLTTF	Q55
4nzu	3-30	FTFGSYAPH	AFISYNGSSKY	RAPD-GAFGY	16	k1-39	QHYDDFP-ISF	A55
Randomization scheme:		[illegible]	[illegible]	[illegible]	10-16	k1-33	[illegible]	[illegible]

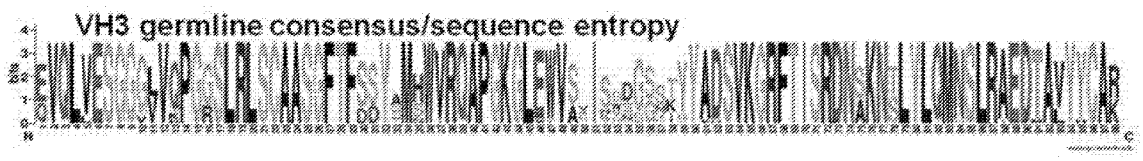


Fig. 1

(57) Abstract: Compositions, methods and uses of high-diversity nucleic acid library that encodes a plurality of antibodies or antibody fragments are presented. The high-diversity nucleic acid library comprises or is derived from (1) a V_H-CDR1/2 sub-library, (2) a plurality of V_H-CDR3 sub-libraries, and (3) a V_L sub-library, each of which comprises a plurality of members. Preferably, each member of the sub-libraries comprises at least one random cassette that has a plurality of degenerate base positions. In an especially preferred embodiment, at least portions of at least two members of the V_H-CDR1/2 sub-library, the plurality of V_H-CDR3 sub-libraries, and the V_L sub-library are recombined to form an expression library member in an expression library, where each member of the expression library encodes a distinct antibody or antibody fragment.

[Continued on next page]

EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*
— *with sequence listing part of description (Rule 5.2(a))*

(88) Date of publication of the international search report:
01 August 2019 (01.08.2019)

mRNA DISPLAY ANTIBODY LIBRARY AND METHODS

[0001] This application claims priority to our copending US provisional patent application with the serial number 62/588,914, which was filed 11/20/2017.

Field of the Invention

[0002] The field of the invention is compositions and methods for ultrahigh-diversity antibody libraries, especially as it relates to mRNA display libraries and use of mRNA display libraries for generating recombinant high-affinity binders.

Background of the Invention

[0003] The background description includes information that may be useful in understanding the present invention. It is not an admission that any of the information provided herein is prior art or relevant to the presently claimed invention, or that any publication specifically or implicitly referenced is prior art.

[0004] All publications and patent applications herein are incorporated by reference to the same extent as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference. Where a definition or use of a term in an incorporated reference is inconsistent or contrary to the definition of that term provided herein, the definition of that term provided herein applies and the definition of that term in the reference does not apply.

[0005] Targeting tumor antigens or neoepitopes with high-affinity, specific antibodies or binding molecules has been proven as effective methods for treating cancer patients. As more and more patient-specific and/or cancer specific tumor antigens and/or neoepitopes are identified *via in vivo*, *in vitro*, or *in silico* through omics data analysis, the demand of creating an antibody library or display library that provides high probabilities of selecting antibodies or binders that are stable, soluble, functional, and adaptable has grown. While high-affinity, specific antibodies or binding molecules can be identified among or derived from natural antibody pools, such identified or derived natural antibodies or binders may not be effective or specific as the diversity of such natural antibodies may be limited depending on the frequency or intensity of exposure to such antigens or neoepitopes.

[0006] In one approach to solve such problem, recombinant phage display libraries can be used. While such approach allows generation of libraries with reasonably high diversity, many rounds of enrichment for binders are often required, which is labor intensive and time consuming. Moreover, despite the relatively large diversity, the binders tend to have less than ideal affinities and stability. Still further, diversity is typically limited by practical considerations such as library volume, transfection efficiency, etc. Such and other approaches can be further optimized, for example, using multiple artificial selection pressures as is described in WO 2006/072773. While such methods may improve stability characteristics, significant amounts of library manipulation and time are required.

[0007] In yet another approach, mRNA display may be performed. Here, mRNA sequences encoding candidate binding molecules (typically scFv) are coupled with a puromycin molecule at their 3'-end, and peptides encoded by the mRNA sequences are generated via *in vitro* translation to produce a fusion product that coupled the mRNA directly to the protein encoded by the mRNA. However, while current mRNA display technology advantageously avoids problems associated with transfection limits and at least conceptually allows for higher diversity, problems with structural integrity or stability, relatively low affinity, and/or cross-reactivity still remain. To further improve at least selected binding characteristics of scFv from mRNA display, VH-CDR3 spectratyping analysis was performed (see *Protein Engineering, Design & Selection*, 2015, vol. 28 no. 10, pp. 427–435). However, such process required iterative analysis and may not be productive for all antigens.

[0008] Thus, even though methods of creating and identification candidate binders using mRNA display and other methods are known, high diversity libraries with binders having high structural integrity/stability, low affinity, and/or low cross-reactivity have remained elusive. Therefore, there is still a need for improved compositions, methods for and uses of mRNA display libraries for rapid generation of stable recombinant high-affinity binders.

Summary of The Invention

[0009] The inventive subject matter is directed to various compositions of, methods for, and use of a high-diversity nucleic acid library that encodes a plurality of antibodies or antibody fragments to allows for reliable and efficient identification of stable, soluble, and functional antibodies or binders to various biomolecules, and especially cancer antigens or neoepitope. Thus, one aspect of the subject matter includes a method of generating a high-diversity

nucleic acid library that encodes a plurality of antibodies or antibody fragments. In this method, three sub-libraries: (1) a V_H-CDR1/2 sub-library, (2) a plurality of V_H-CDR3 sub-libraries, and (3) a V_L sub-library, each having a plurality of members are generated or provided. Each member of the three sub-libraries comprises at least one random cassette that has a plurality of degenerate base positions. At least portions of at least two members of the three libraries are recombined to form an expression library member in an expression library, which has a plurality of expression library members. Each expression library member encoding a distinct antibody or antibody fragment. In a preferred embodiment, the expression library member is transcribed into an mRNA fragment, which then is coupled with a puromycin molecule at 3'-end.

[0010] In another aspect of the inventive subject matter, the inventors contemplate a composition having a plurality of nucleic acid libraries. The plurality of nucleic acid libraries includes (1) a V_H-CDR1/2 sub-library, (2) a plurality of V_H-CDR3 sub-libraries, and (3) a V_L sub-library. Each of the sub-libraries (1)-(3) comprises a plurality of members and the each member of the sub-libraries comprises at least one random cassette that has a plurality of degenerate base positions.

[0011] In still another aspect of the inventive subject matter, the inventors contemplate use of the composition above for generating a high-diversity nucleic acid library.

[0012] In still another aspect of the inventive subject matter, the inventors contemplate a high-diversity nucleic acid library composition having a plurality of library members. The high-diversity nucleic acid library member includes a recombinant nucleic acid comprising a plurality of random cassettes, each having a plurality of degenerate base positions. The plurality of random cassettes is derived from at least two members from any of two libraries from the following: (1) a V_H-CDR1/2 sub-library, (2) a plurality of V_H-CDR3 sub-libraries, and (3) a V_L sub-library.

[0013] In still another aspect of the inventive subject matter, the inventors contemplate use of the high-diversity nucleic acid library for generating a therapeutic recombinant antibody against a cancer neoepitope.

[0014] In still another aspect of the inventive subject matter, the inventors contemplate a method of generating a recombinant antibody. In this method, three sub-libraries: (1) a V_H-CDR1/2 sub-library, (2) a plurality of V_H-CDR3 sub-libraries, and (3) a V_L sub-library, each

having a plurality of members are generated or provided. Each member of the three sub-libraries comprises at least one random cassette that has a plurality of degenerate base positions. At least portions of at least two members of the three libraries are recombined to form an expression library member in an expression library, which has a plurality of expression library members. Each expression library member encoding a distinct antibody or antibody fragment. Then, the method continues with generating the recombinant antibody or fragment thereof using the expression library member.

[0015] In still another aspect of the inventive subject matter, the inventors contemplate a method of isolating a high affinity binder having an affinity of equal or less than 100nM to an antigen, by contacting the antigen to a composition constructed by the methods described above.

[0016] In still another aspect of the inventive subject matter, the inventors contemplates a recombinant nucleic acid fragment generated using an oligonucleotide selected from Table 1 or Table 2 provided below.

[0017] In still another aspect of the inventive subject matter, the inventors contemplate a synthetic nucleic acid mixture having a nucleic acid sequence selected from Table 1 or Table 2 provided below.

[0017A] In one aspect, the inventors provide a method of generating a high-diversity nucleic acid library that encodes a plurality of Fab domains, the method comprising: generating or providing (1) a VH-CDR1/2 sub-library, (2) a plurality of VH-CDR3 sub-libraries, and (3) a VL sub-library, wherein each of the sub-libraries (1)–(3) comprises a plurality of members; wherein each member of the sub-libraries comprises at least one random cassette that has a plurality of degenerate base positions, and wherein the random cassette is generated using an oligonucleotide selected from SEQ ID NOs: 1–25; and recombining at least portions of at least two members of the VH-CDR1/2 sub-library, the plurality of VH-CDR3 sub-libraries, and the VL sub-library to form an expression library member in an expression library, wherein the expression library comprises a plurality of expression library members, each expression library member encoding a distinct Fab domain.

[0017B] In another aspect, the inventors provide a method of generating a recombinant antibody, comprising: generating or providing (1) a VH-CDR1/2 sub-library, (2) a plurality of VH-CDR3 sub-libraries, and (3) a VL sub-library, wherein each of the sub-libraries (1)–

(3) comprises a plurality of members; wherein each member of the sub-libraries comprises at least one random cassette that has a plurality of degenerate base positions, and wherein the random cassette is generated using an oligonucleotide selected from SEQ ID NOs: 1–25; recombining at least portions of at least two members of the VH-CDR1/2 sub-library, the plurality of VH-CDR3 sub-libraries, and the VL sub-library to form an expression library member in an expression library, wherein the expression library comprises a plurality of expression library members, each expression library member encoding a distinct antibody or antibody fragment; and generating the recombinant viral vector comprising at least one expression library member.

[0017C] In another aspect, the inventors provide a method of generating a recombinant Fab domain, comprising: generating or providing (1) a VH-CDR1/2 sub-library, (2) a plurality of VH-CDR3 sub-libraries, and (3) a VL sub-library, wherein each of the sub-libraries (1)–(3) comprises a plurality of members; wherein each member of the sub-libraries comprises at least one random cassette that has a plurality of degenerate base positions, and wherein the random cassette is generated using an oligonucleotide selected from SEQ ID NOs: 1–25; recombining at least portions of at least two members of the VH-CDR1/2 sub-library, the plurality of VH-CDR3 sub-libraries, and the VL sub-library to form an expression library member in an expression library, wherein the expression library comprises a plurality of expression library members, each expression library member encoding a distinct Fab domain; and generating the recombinant Fab domain using the expression library member.

[0018] Various objects, features, aspects and advantages of the inventive subject matter will become more apparent from the following detailed description of preferred embodiments, along with the accompanying drawing.

[0018A] Any discussion of documents, acts, materials, devices, articles or the like which has been included in the present specification is not to be taken as an admission that any or all of these matters form part of the prior art base or were common general knowledge in the field relevant to the present disclosure as it existed before the priority date of each of the appended claims.

[0018B] Throughout this specification the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated element, integer or step, or group of elements, integers or steps, but not the exclusion of any other element, integer or step, or group of elements, integers or steps.

Brief Description of the Drawing

[0019] **Fig. 1** illustrates one exemplary randomization strategy using VH3/Vk1 pairs.

[0020] **Fig. 2** illustrates exemplary locations for sequence randomization in heavy chain CDR1 and CDR2.

[0021] **Fig. 3** illustrates exemplary sequence randomization in heavy chain CDR3.

[0022] **Fig. 4** illustrates exemplary sequence randomization in light chain CDR3 with nucleic acid sequences to the left and amino acid choices to the right.

[0023] Fig. 5 illustrates an exemplary generation of hybrid nucleic acid elements by isolating and combining random cassettes of multiple recombinant nucleic acid segments.

[0024] Fig. 6 shows a size exclusion chromatography result showing a single peak indicating a stable protein expression of αB7-H4$_{801}$.

[0025] Fig. 7 shows a capillary electrophoresis sodium dodecyl sulfate (CE-SDS) data indicating similar molecular behavior of αB7-H4$_{801}$ compared to commercial antibodies.

[0026] Fig. 8 shows graphs indicating binding of *in vitro* selected αB7-H4 antibodies to B7-H4.

[0027] Fig. 9 shows graphs of functional analysis of *in vitro* selected αB7-H4 binders and αPD-L1 binders.

[0028] Fig. 10 shows graphs indicating binding affinities of αB7-H4 scFv and αB7-H4 IgG1.

[0029] Fig. 11 shows an IL-8 activity assay and its result by measuring neutrophil size changes.

[0030] Fig. 12 shows bar graphs indicating neutralization effect of αIL-8 antibody to IL-8 activity of increasing neutrophil size.

[0031] Fig. 13 shows IL-8 activity assay and its results shown in bar graph indicating neutralization effect of αIL-8 antibody to IL-8 activity by inhibiting neutrophil migration.

[0032] Fig. 14 shows exemplary results using mRNA display library compositions presented herein with respect to selected antigen targets.

[0033] Fig. 15 shows an exemplary graph depicting affinities of selected binders configured as scFv versus IgG where the binders were identified using mRNA display library compositions presented herein.

Detailed Description

[0034] The inventors now discovered that specific and effective recombinant antibodies or fragments thereof can be generated or identified by constructing a high-diversity nucleic acid library using targeted diversification of selected domains of the antibodies or fragments thereof encoded by members of the high-diversity nucleic acid library. In order to achieve

such goal, the inventors have now discovered that one or more domains or subdomains of antibody/binder can be pre-selected and a plurality of nucleic acid sub-libraries can be generated using random cassettes in a pre-selected domain or subdomain. The inventors further discovered that the members of the sub-libraries can be recombined to construct the high-diversity nucleic acid library that allows high diversity among library members, yet provides higher probabilities of identifying antibodies/binders that are stable, soluble, functional, and adaptable when used *in vivo* against the cancer antigens or neoepitopes (preferably cancer-specific, patient-specific neoepitopes or neoantigens).

[0035] Indeed, and as shown in more detail below, the libraries presented herein allow for isolation of at least one binder to any arbitrary antigen, typically in a single or two-pass enrichment, where the binder has a K_d of equal or less than 100nM, and more typically equal or less than 10nM. Moreover, contemplated systems and methods allow for scFv libraries having a diversity of at least 10^9, at least 10^{10}, at least 10^{11}, at least 10^{12}, at least 10^{13}, at least 10^{14}, at least 10^{15}, or at least 10^{16} distinct library members, all in a time frame that is significantly reduced as compared to conventional library construction. Thus, it should be appreciated that the speed of antibody discovery is substantially increased.

[0036] As used herein, the term "tumor" refers to, and is interchangeably used with one or more cancer cells, cancer tissues, malignant tumor cells, or malignant tumor tissue, that can be placed or found in one or more anatomical locations in a human body.

[0037] As used herein, the term "bind" refers to, and can be interchangeably used with a term "recognize" and/or "detect", an interaction between two molecules with a high affinity with a K_D of equal or less than 10^{-6}M, or equal or less than 10^{-7}M.

[0038] As used herein, the term "provide" or "providing" refers to and includes any acts of manufacturing, generating, placing, enabling to use, or making ready to use.

Construction of Nucleic Acid Sub-libraries

[0039] Generally, structural components (heavy chain, light chain, constant domains, variable domains) of antibodies are closely related to their functions. For example, the variable domains in the heavy chain (V_H) and light chain (V_L) constitute, together, the epitope binding domain, which provides specificity to the antibodies. Each of the V_H and V_L includes three complementarity determining regions (CDRs, CDR1-3) with unique amino acid

sequences based on their specificity to an antigen. Thus, it had previously been contemplated that a recombinant nucleic acid library for generating or identifying antibodies can be created by randomizing the sequences encoding the CDRs of V_H and V_L. However, the inventors found that while complete randomization of all CDRs of V_H and V_L may provide great diversity to the library, it also creates inefficiency in generating all combinations of random sequences and screening all randomized combinations as not all randomized V_H and V_L can be soluble or stably expressed when it is recombined to form an antibody (e.g., IgG1, etc.). Moreover, covering the entire diversity space is not practical due to the extremely large number of possible library members.

[0040] Thus, the inventors contemplate that subdomains of V_H and V_L can be divided into two categories: a framework region that are generally common among V_H or V_L of different antibodies (or genes encoding the antibodies) and a targeted diversification region that can be at least partially or completely randomized without significantly affecting the stability and/or solubility of the final peptide product (e.g., scFv, IgG1, etc.). Preferably, the targeted diversification region of V_H includes at least a portion of CDR1, CDR2-n (N-terminus side of CDR2), CDR2-c (C-terminus side of CDR2), and CDR3. In further preferred aspects, the targeted diversification region of V_L includes at least a portion of CDR3.

[0041] As such, in one exemplary and especially preferred aspect of the inventive subject matter, a nucleic acid library can be created by generating recombinant nucleic acids that include one or more random sequence cassettes in one or more targeted diversification region of V_H and/or V_L. In one preferred embodiment, the inventors contemplate three different sub-libraries having different sets of random sequence cassettes in different targeted diversification regions such that each sub-library retains the diversity within randomized targeted diversification regions while avoiding too many randomized recombinant sequences in a single sub-library that may render the volume of the single sub-library impractical or inefficient to handle for quick or timely screenings. Furthermore, conserved areas between the targeted diversification regions are selected or designed for maximum stability and solubility.

[0042] In one embodiment, the sub-libraries include a V_H-CDR1/2 sub-library. The V_H-CDR1/2 sub-library comprises a plurality of recombinant nucleic acids (e.g., recombinant DNA) having one or more random sequence cassettes corresponding to at least a portion of V_H CDR1 and/or at a portion of V_H CDR2. As used herein, the random cassette

corresponding to a portion of V_H CDR1 means that the random cassette is located in an area of the recombinant nucleic acid, in which sequences encoding CDR1 portion should be present in order to encode a portion of V_H domain which is at least structurally or functionally similar to V_H domains of natural antibodies. For example, recombinant nucleic acids in a V_H-CDR1/2 sub-library may have a structure as below (randomized region is underlined, and fixed sequenced region is parenthesized):

5'- (Promoter –5' UTR – FW1) + <u>CDR1</u> + (FW2) + <u>CDR2</u> + (FW3 – CDR3 – FW4)

As used herein, UTR refers to untranslated region and FW refers framework region (e.g., FW1 is the first framework region that may be distinct from the second framework region (FW2)). In this structure, the random sequence cassettes can be inserted in areas of CDR1 or CDR2, or preferably, both CDR1 and CDR2. In some embodiments, more than one random sequence cassettes, preferably two random sequence cassettes can be inserted in the area of CDR2: CDR2-n (for 5'-end side of CDR2) and CDR-c (for 3'-end side of CDR2).

[0043] The sub-libraries can also include a plurality of V_H-CDR3 sub-libraries. Each of V_H-CDR3 sub-library comprises a plurality of recombinant nucleic acids (e.g., recombinant DNA) having one or more random sequence cassettes corresponding to at least a portion of V_H CDR3. Similar to the V_H-CDR1/2 sub-library, a recombinant nucleic acids in V_H-CDR1/2 sub-library may have a structure as below (randomized region is underlined, and fixed sequenced region is parenthesized):

5'- (Promoter – 5' UTR – FW1 + CDR1 + FW2 + CDR2 + FW3) – <u>CDR3</u> – (FW4)

Preferably, the fixed sequences (e.g., Promoter – 5' UTR – FW1 + CDR1 + FW2 + CDR2 + FW3, FW4) of the recombinant nucleic acids of the V_H-CDR1/2 sub-library and/or the V_H-CDR3 sub-library are selected to use the most common and/or conserved sequences among the natural antibodies (e.g., IgG1s against various antigens) such that the fixed sequences are most expressable and adaptable to multiple formats including peptides expressed as a single chain variable fragment (scFv), a modified form of scFv, full length immunoglobulin, or a portion of immunoglobulin. Thus, in preferred embodiments, the fixed sequences of the recombinant nucleic acids of V_H-CDR1/2 sub-library and of the recombinant nucleic acids of V_H-CDR3 sub-library are at least 70%, preferably at least 80%, more preferably at least 90% identical (shared) with each other.

[0044] The sub-libraries can also include a V_L sub-library. The V_L sub-library comprises a plurality of recombinant nucleic acids (e.g., recombinant DNA) having one or more random sequence cassettes corresponding to at least a portion of V_L CDR3. Similar to the V_H-CDR1/2 sub-library, recombinant nucleic acids in V_H-CDR1/2 sub-library may have a structure as below (randomized region is underlined, and fixed sequenced region is parenthesized):

5'- (Promoter – 5' UTR – FW1 + CDR1 + FW2 + CDR2 + FW3) – <u>CDR3</u> – (FW4)

Preferably, the fixed sequences of the recombinant nucleic acids of the V_L sub-library are at least 70%, preferably at least 80%, more preferably at least 90% identical (shared) to those of recombinant nucleic acids of the V_H-CDR1/2 sub-library or V_H-CDR3 sub-library.

[0045] While any randomized sequences can be considered to generate the random sequence cassettes, the inventors contemplate that strategized random sequence cassettes for CDR1, CDR2, CDR3 of the V_H and CDR3 of the V_L domain would render a high complexity and large potential binding surface when expressed as a binding peptide (e.g., scFv, etc.). For example, the strategized random sequence cassettes for CDR1, CDR2 of the V_H-CDR1/2 sub-library may be semi-random sequence cassettes having 3 or less, preferably 2 or less, or more preferably, one random sequence (encoding 3 or less, 2 or less, or one random amino acid per cassette) per cassette. The location of the random sequence in the random cassette may vary depending on the random amino acid in the cassette. In another example, the strategized random sequence cassettes for CDR3 of V_H-CDR3 sub-library may include more randomized sequences such that 4 or more, preferably 5 or more, or more preferably 6 or more random sequences (encoding 4 or more, preferably 5 or more, or more preferably 6 or more random amino acids per cassette) are present per cassette. In yet another example, the strategized random sequence cassettes for CDR3 of V_L sub-library may include more randomized sequences such that 4 or more, preferably 5 or more, or more preferably 6 or more random sequences (encoding 4 or more, preferably 5 or more, or more preferably 6 or more random amino acid per cassette) are present per cassette.

[0046] In an especially preferred aspect of the inventive subject matter, the inventors contemplate that preferred random sequence cassettes for sub-libraries can be generated using oligonucleotides presented in Table 1 (for V_H-CDR1/2 sub-library and V_H-CDR3 sub-library), and Table 2 (for V_L sub-library). As shown in **Tables 1 and 2**, each oligonucleotide includes a random sequences (highlighted) having degenerate code, shown as IUPAC

ambiguity codes. For example, one oligonucleotide for CDR1 random sequence cassette includes a random sequence "RVT", which represents "A/G,A/C/G,T", whose combination can encode one of threonine (T), alanine (A), asparagine (N), aspartic acid (D), serine (S) or glycine (G). The choice of amino acids encoded by the degenerate codons are depicted to the right and are indicated with X.

Additionally and preferably, the random sequence cassettes for V_H-CDR3 sub-library may include nucleic acid sequences in different length. For example, the random sequence cassettes for V_H-CDR3 sub-library may be in any length between 10-30 amino acids, preferably between 10-25 amino acids, more preferably between 10-20 amino acids. Thus, as shown in Table 1, the oligonucleotides for generating random sequence cassette for V_H-CDR3 sub-library may include a various repeats (e.g., 4-10 repeats) of "NNK" (which represents G/A/T/C, G/A/T/C, G/T) between sequences encoding D/G-R/L and A/G (see also Fig. 3). Generation and diversity of light chain sequences are exemplarily shown in Fig. 4.

V_H CDR1	SEQ ID NO.1: GGCTTAGGTCTCATTTC**RVT**AGTTACGCTATGCATTGGGCGAGACGAGGTCTGAACGG	X=T,A,N,D,**S**,G
	SEQ ID NO.2: GGCTTAGGTCTCATTTCTCT**RVK**TACGCTATGCATTGGGCGAGACGAGGTCTGAACGG	X=T,A,N,K,D,E, **S**,R,G
	SEQ ID NO.3: GGCTTAGGTCTCATTTCTCTAGTTAC**KKG**ATGCATTGGGCGAGACGAGGTCTGAACGG	X=G,W,L,V
	SEQ ID NO.4: GGCTTAGGTCTCATTTCTCTAGTTAC**WMT**ATGCATTGGGCGAGACGAGGTCTGAACGG	X=S,Y,T,N
	SEQ ID NO.5: GGCTTAGGTCTCATTTCTCTAGTTACGCTATG**RYT**TGGGCGAGACGAGGTCTGAACGG	X=S,T,N
V_H CDR2-n	SEQ ID NO.6: GGCTTAGGTCTCGTTCA**THC**ATTAGTGGTAGTGGACGAGACGAGGTCTGAACGG	X=Y,F,S
	SEQ ID NO.7: GGCTTAGGTCTCGTTCA**VKT**ATTAGTGGTAGTGGACGAGACGAGGTCTGAACGG	X=V,G,I,S,L,R
	SEQ ID NO.8:	X=W,R

	GGCTTAGGTCTCGTTCAGCTATTYGGGGTAGTGGACGAGACGAGGTCTGAACGG	
	SEQ ID NO.9: GGCTTAGGTCTCGTTCAGCTATTDATGGTAATGGACGAGACGAGGTCTGAACGG	X=Y,N,D+N53
	SEQ ID NO.101: GGCTTAGGTCTCGTTCAGCTATTAGTWMTAGTGGACGAGACGAGGTCTGAACGG	X=Y,S,T,N
	SEQ ID NO.11: GGCTTAGGTCTCGTTCAGCTATTAGTKGGAGTGGACGAGACGAGGTCTGAACGG	X=W,G
	SEQ ID NO.12: GGCTTAGGTCTCGTTCAGCTATTAGTGGTRRTGGACGAGACGAGGTCTGAACGG	X=D,G,S,N
V_H CDR2-c	SEQ ID NO.13: GGCTTAGGTCTCGTGGARVKAGTACTTACTACGCGAGACGAGGTCTGAACGG	X=S,T,**G,**A,N,K,D,E
	SEQ ID NO.14: GGCTTAGGTCTCGTGGAGGTNATACTTACTACGCGAGACGAGGTCTGAACGG	X=Y,N,D,H
	SEQ ID NO.15: GGCTTAGGTCTCGTGGAGGTRTAACTTACTACGCGAGACGAGGTCTGAACGG	X=T,K,R,E,A,G
	SEQ ID NO.16: GGCTTAGGTCTCGTGGAGGTAGTACTYRTTACGCGAGACGAGGTCTGAACGG	X=D,G,N,S,H,R
V_H CDR3	SEQ ID NO.17: GGCTTAGGTCTCTCCGTGRTCKC(NNK)mGSTTTCGCGAGACGAGGTCTGAACGG	**(D,G)-(R,L)-(Xaa=4-10)-(A,G)**

Table 1

V_L CDR3	SEQ ID NO.18: GGCTTAGGTCTCTGCAGDSGDMTRYTDSGCCTTWCACTTCGAGACGAGGTCTGAACGG	Q-X_1-X_2-X_3-X_4-P-X_5 X_1=Y,D,L,A,H,S, F,R,T,W,G X_2=Y,N,D,S,T,A X_3=S,N,T,A,D,G X_4=Y,F,A,L,T,S, H,W,I,N,R,V,D,G
	SEQ ID NO.19: GGCTTAGGTCTCTGCAGEWTDMTRVTDSGCCTTWCACTTCGAGACGAGGTCTGAACGG	
	SEQ ID NO.20:	

	GGCTTAGGTCTCTGCAGDSGDMTRVTNWTCCTTWCACTTCGAGACGAGGTCTGAACGG	X_5=L,Y,W,F,R
	SEQ ID NO.21: GGCTTAGGTCTCTGCAGBWTDMTRVTNWTCCTTWCACTTCGAGACGAGGTCTGAACGG	
	SEQ ID NO.22: GGCTTAGGTCTCTGCAGDSGDMTRVTDSGCCTYKGACTTCGAGACGAGGTCTGAACGG	
	SEQ ID NO.23: GGCTTAGGTCTCTGCAGBWTDMTRVTDSGCCTYKGACTTCGAGACGAGGTCTGAACGG	
	SEQ ID NO.24: GGCTTAGGTCTCTGCAGDSGDMTRVTNWTCCTYKGACTTCGAGACGAGGTCTGAACGG	
	SEQ ID NO.25: GGCTTAGGTCTCTGCAGBWTDMTRVTNWTCCTYKGACTTCGAGACGAGGTCTGAACGG	

Table 2

[0047] Most typically, the oligonucleotides presented in Table 1 and 2 are provided in a single strand DNA, which can be converted using DNA polymerase I (Klenow fragment) into double-stranded DNA fragment to so be inserted into a backbone comprising the fixed sequenced region (e.g., 5'- (Promoter – 5' UTR – FW1 + CDR1 + FW2 + CDR2 + FW3) – (FW4) for recombinant nucleic acids of V_L sub-library, etc.). Yet, it is also contemplated that the oligonucleotides presented in Table 1 and 2 are also present with the complementary oligonucleotides to form a double stranded nucleic acids without using polymerase enzymes.

[0048] In some embodiments, the recombinant nucleic acids of sub-libraries also include a nucleic acid sequence encoding a protein tag such that the peptide encoded by the recombinant nucleic acids can be isolated using the binder against the protein tag. For example, preferred proteins tag include a FLAG tag (with a sequence motif DYKDDDDK), a Myc tag (with a sequence motif EQKLISEEDL), and an HA-tag. In some embodiments, the protein tags can be repeated to strengthen the signal or increase the detection (e.g., three repetitions of FLAG tag (3X FLAG), etc.)

[0049] It is contemplated that some random sequence cassettes inserted in the recombinant nucleic acids of sub-libraries, may introduce frame shifts, nonsense mutations, and

sequence(s) that are destabilizing the structure of the peptide encoded by the recombinant nucleic acids. Thus, in some embodiments, the inventors contemplate that the recombinant nucleic acids of sub-libraries are *in vitro* tested so that any recombinant nucleic acids encoding unstable or misfolded peptides can be removed from the library. For example, the recombinant nucleic acids of the V_H-CDR3 sub-libraries or the V_L sub-library can be tested for their binding affinity to protein A of *Staphylococcus aureus* or protein L of *Finegoldia magna*, which binds to structured epitopes of V_H3 domain or V_L (Vκ) domain of immunoglobulin independently to CDR sequences, respectively.

[0050] Any suitable methods to screen the recombinant nucleic acids by their binding affinities to protein A or protein L are contemplated. In one exemplary embodiment the recombinant nucleic acids of sub-libraries are transcribed into mRNAs by *in vitro* transcription and the 3'-end of the mRNAs are coupled (covalently linked) to puromycin. The puromycin-coupled mRNAs are *in vitro* translated such that the peptides transcribed from the puromycin-coupled mRNAs are coupled with the mRNAs via the puromycin. Next, the peptides are contacted with protein A or protein L to identify peptides effectively binding to the protein A or protein L. Preferably, peptides binding to protein A or protein L with an affinity with a K_D of equal or less than 10^{-6}M, preferably equal or less than 10^{-7}M are selected and isolated. Once the peptides with high affinity to protein A or protein L are isolated, cDNAs of the isolated peptides can be generated via *in vitro* reverse-transcription of the mRNAs coupled with the puromycin and the peptides. The so generated cDNAs of the isolated peptides can be then inserted as random sequence cassettes to generate selected recombinant nucleic acids of V_H-CDR3 sub-libraries or the V_L sub-library. Alternatively, it is also contemplated that the recombinant nucleic acids of sub-libraries can be present in a form of mRNAs, which is optionally pre-coupled with puromycin molecule such that the *in vitro* transcription step for the recombinant nucleic acids (in DNA format) may not be needed.

Construction of scFv Library from the sub-libraries

[0051] The inventors further contemplate that at least two recombinant nucleic acids (members) of the sub-libraries can be recombined to form recombinant scFv nucleic acids. In a preferred embodiment, each of the at least two recombinant nucleic acids (members) is selected from different sub-libraries. For example, one recombinant nucleic acid may be selected from each of the V_H-CDR1/2 sub-library, the plurality of V_H-CDR3 sub-libraries, and the V_L sub-library. For other example, one recombinant nucleic acid may be selected

from each of two of V_H-CDR1/2 sub-library, the plurality of V_H-CDR3 sub-libraries, and the V_L sub-library. Preferably, at least one of, more preferably all of, the recombinant nucleic acid(s) selected from the sub-libraries are pre-selected via affinity binding screening as described above.

[0052] Most typically, the recombinant scFv nucleic acids can be constructed by recombining a portion of the recombinant nucleic acids from sub-libraries. In this embodiment, the portion of the recombinant nucleic acids includes the random sequence cassettes inserted into the recombinant nucleic acids. Thus, for example, as a first step, the portion of the recombinant nucleic acids of the V_H-CDR1/2 sub-library can be 5'-[CDR1 + (FW2) + CDR2]-3' (random sequence cassettes are underlined), preferably 5'-(portion of FW1)-[CDR1 + (FW2) + CDR2]-(portion of FW3)-3', more preferably 5'- (Promoter –5' UTR – FW1) + CDR1 + (FW2) + CDR2 + (portion of FW3)-3' or 5'- (Promoter –5' UTR – FW1) + CDR1 + (FW2) + CDR2 + (a small linker)-3'. Similarly, for example, the portion of the recombinant nucleic acids of the V_H-CDR3 sub-libraries can be 5'- [CDR3] – 3' (random sequence cassettes are underlined), preferably 5'- (portion of FW3) – CDR3 – (portion of FW4)-3', more preferably, 5'- (portion of FW3) – CDR3 – (FW4)-3', or 5'- (a small linker) – CDR3 – (FW4)-3'. The portions of the recombinant nucleic acids from the V_H-CDR1/2 sub-library and the V_H-CDR3 sub-libraries are then isolated (e.g., by PCR) and can be recombined (e.g., fused via restriction-ligation methods, generated via a recombinant-PCR, etc.) to form a V_H domain recombinant nucleic acid. Thus, typically, the V_H domain recombinant nucleic acid would be in a structure of 5'- Promoter – 5' UTR – FW1 + CDR1 + FW2 + CDR2 + FW3 – CDR3 – FW4 –3'(random sequence cassettes are underlined). Optionally, the V_H domain recombinant nucleic acid may also include a nucleic acid sequence encoding a protein tag (e.g., FLAG tag, Myc tag, HA tag, etc.) in its 3'-end as described above. In addition, such generated V_H domain recombinant nucleic acids can be placed in a V_H domain library as V_H domain library members.

[0053] The so formed V_H domain recombinant nucleic acids can be further recombined with recombinant nucleic acids of the V_L sub-library to form the recombinant scFv nucleic acids. Fig. 5 shows one exemplary method of recombining the sequences from sub-libraries. As shown, and also typically, a portion of the V_H domain recombinant nucleic acid and a portion of the recombinant nucleic acid of the V_L sub-library are fused into one the recombinant scFv nucleic acids. For example, the portion of V_H domain recombinant nucleic acid may include

5'- Promoter – [5' UTR – FW1 + CDR1 + FW2 + CDR2 + FW3 – CDR3 –FW4 –3' (preferably without any nucleic acid encoding a protein tag in its 3'-end), and the portion of the recombinant nucleic acid of the V_L sub-library may include FW1' + CDR1 + FW2' + CDR2 + FW3' – CDR3 – FW4' (without promoter and 5'-UTR) such that the recombinant nucleic acid of the V_L sub-library can be fused to the 3'-end of the portion of V_H domain recombinant nucleic acid. Thus, the typical recombinant scFv nucleic acid would be in a structure of 5'- Promoter – [5' UTR – FW1 + CDR1 + FW2 + CDR2 + FW3 – CDR3 – FW4]V_H- [FW1' + CDR1 + FW2' + CDR2 + FW3' – CDR3 – FW4']V_L-3'. It is highly preferred that the portion of V_H domain recombinant nucleic acid and the portion of the recombinant nucleic acid of the V_L sub-library are placed in the same reading frame such that they encode a single polypeptide.

[0054] Preferably, the portion of V_H domain recombinant nucleic acid and the portion of the recombinant nucleic acid of the V_L sub-library are fused via a nucleic acid encoding a linker (a short peptide spacer fragment) between two portions. Any suitable length and order of peptide sequence for the linker or the spacer can be used. However, it is preferred that the length of the linker peptide is between 3-30 amino acids, preferably between 5-20 amino acids, more preferably between 5-15 amino acids. For example, the inventors contemplate that glycine-rich sequences (*e.g.,* gly-gly-ser-gly-gly, etc.) are employed to provide flexibility of scFv between the V_H and V_L domains.

[0055] Optionally, the recombinant scFv nucleic acids may also include a nucleic acid sequence encoding a protein tag (e.g., FLAG tag, Myc tag, HA tag, etc.) in its 3'-end as described above. In addition, such generated recombinant scFv nucleic acids can be placed in an expression library as expression library members.

[0056] In some embodiments, the so formed recombinant scFv nucleic acids are further screened and/or ranked based on their binding affinities to one or more ligands of interest (e.g., cancer antigens, neoepitopes, etc.), stability, pH sensitivity, and/or species cross-reactivity. For example, the stability of the scFv peptides encoded by the recombinant scFv nucleic acids can be analyzed by size exclusion chromatography measuring the size of the peptide over time. For other example, pH sensitivity and binding affinity of the scFv peptides encoded by the recombinant scFv nucleic acids can be analyzed by contacting the scFv peptides with one or more ligands in different buffer conditions (pH, temperature, etc.).

[0057] For those analysis and further isolation of desired recombinant scFv nucleic acids from the expression library, the inventors contemplate that the recombinant scFv nucleic acids can be present in a form of mRNAs, which is optionally pre-coupled with puromycin molecule at the 3'-end of the mRNAs. The puromycin-coupled mRNAs can then be *in vitro* translated such that the peptides transcribed from the puromycin-coupled mRNAs are coupled with the mRNAs via the puromycin. Then, the peptides are contacted with one or more ligands, optionally in different buffer conditions (pH, temperature, etc.). Preferably, peptides binding to the ligand with an affinity with a K_D of equal or less than 10^{-6}M, preferably equal or less than 10^{-7}M, between pH 5.0-8.0, preferably between pH 6.0-8.0, more preferably between pH 6.5-8.0 are selected and isolated. Once the peptides with high affinity to the ligand(s) are isolated, cDNAs of the isolated peptides can be generated via *in vitro* reverse-transcription of the mRNAs coupled with the puromycin and the peptides.

[0058] Additionally, the so generated cDNAs of the isolated peptides encoded by recombinant scFv nucleic acids can be grafted on and replaced the portion of the immunoglobulin to form a recombinant immunoglobulin or fragments thereof. For example, the so generated cDNA can be fused with the backbone of the immunoglobulin heavy chain constant region such that the variable region of heavy and light chain of the immunoglobulin can be replaced with the scFv formed by the isolated peptide. Alternatively, the inventors also contemplate that the V_H portion (or derived from V_H domain recombinant nucleic acid) and V_L portion (or derived from of the recombinant scFv nucleic acid) of the recombinant scFv nucleic acids can be grafted on and replaced the portion of the immunoglobulin to form a recombinant immunoglobulin or fragments thereof. For example, the V_H portion (or derived from V_H domain recombinant nucleic acid) and V_L portion (or derived from of the recombinant scFv nucleic acid) of the recombinant scFv nucleic acids are fused with the backbone of the immunoglobulin heavy chain constant region or light chain constant region, respectively, to form an immunoglobulin with variable regions specific to the desired ligand.

[0059] In these examples, it is contemplated that the immunoglobulin can include any type (e.g., IgG, IgE, IgM, IgD, IgA and IgY) and any class (e.g., IgG1, IgG2, IgG3, IgG4, IgA1 and IgA2) of heavy chain or constant domain to constitute different types of immunoglobulin. In addition, the "antibody" can include, but not limited to a human antibody, a humanized antibody, a chimeric antibody, a monoclonal antibody, a polyclonal antibody. In this context, it should be noted that contemplated systems and methods allow for the generation of

species-specific antibodies by grafting the isolated V_H and V_L domains onto the remainder of the antibody of a desired species (e.g., human). In another example, the so generated cDNA can be fused with nucleic acids encoding other portion of the immunoglobulin to form a fragment of the immunoglobulin. In this example, it is contemplated that the fragment of the immunoglobulin can be Fab fragments, Fab' fragments, F(ab')2, disulfide linked Fvs (sdFvs), and Fvs. The inventors further contemplate that a portion of the so generated cDNA can be fused with nucleic acids encoding other portion of the immunoglobulin to form any fragment comprising either V_H segment and/or V_L segment.

[0060] Additionally, the inventors contemplate that the scFv portions may also be used as targeting entities for various proteins and non-protein molecules. For example, the scFv portions may be coupled (typically as chimeric protein) to an ALT-803 type molecule to form a TxM entity that has specific targeting capability (see *e.g.*, *J Biol Chem.* 2016 Nov 11;291(46):23869-23881). In another example, the scFv portion may be coupled to a carrier protein (*e.g.*, albumin) to allow target specific delivery of one or more drugs to a specific location in a tumor microenvironment where the drugs are coupled to the carrier.

[0061] The inventors further contemplate that by construction the sub-libraries via targeted diversification of random sequences, and/or preselecting the members of the sub-libraries, the expression library can achieve approximately 10^{12} complexity with minimal sacrifice of diversity by removing unstable, non-binding, or misfolded sequences. Thus, the above described approach to generate expression library provides meaningful size of sequence complexity, yet is practical to screen binders/antibodies in a small volume. In addition, the above described approach to generate expression library simplified the screening procedure of the binders/antibodies. Traditionally, *in vitro* validation of any nucleic acid sequences (e.g., randomized sequences) encoding binding domain (or motif) required the nucleic acid sequences converted to F_{ab} domain, then the binding affinity could be tested via pull-down assay with the ligand of interest. The methods presented herein allows *in vitro* validation of nucleic acid sequences encoding binding domain (or motif) via ranking by affinity (e.g., Kd value), pH sensitivity, and species cross-reactivity (e.g., via surface plasmon resonance assay, etc.) without converting the nucleic acid sequences into F_{ab} domain. Further, pre-selection of members from each library based on stability and sensitivity reduces the pool to be tested in the library such that the desired binders/scFv/antibody domains can be identified more quickly and efficiently. Therefore, the inventors also contemplate methods for isolation of

high-affinity binders (*e.g.*, with nano- and picomolar K_d) from a high-diversity pool using mRNA display techniques in which library members after *in vitro* translation are screened against a solid phase bound antigen. Once binders are identified, they can be further characterized by surface plasmon resonance spectroscopy with respect to affinity and K_{on}/K_{off} characteristics as is further described below. Viewed form a different perspective, contemplated systems and methods allow for rapid detection of binders and generation of scFv or antibodies in a process that is entirely independent from an *in vivo* immune system.

Examples

[0062] While any suitable diversification scheme to identify targeted diversification region(s) can be contemplated to maximize diversity while maintaining efficiency, the inventors found that VH3/Vk1 can be one of the good candidate regions for randomization among the various domains of immunoglobulin, VH3 is considered by far most stable and soluble VH domain, and Vk1 of light chain is stable and soluble. Thus, it is contemplated that the VH3/Vk1 randomized pairs would convert to a full size immunoglobulin more efficiently. Accordingly, the inventors developed pre-selection strategy using VH3 and Vk1 frameworks. **Fig. 1** shows one exemplary randomization strategy using VH3/Vk1 pairs. Protein sequences of at least 14 immunoglobulin molecules specific to one antigen are compared and analyzed. The most stable and conserved sequences among 14 immunoglobulin molecules are used as frameworks and locus of variable sequences are analyzed to use as randomized sequences and the degree of randomization (e.g., complete random, partially random, etc.).

[0063] Based on the randomization strategy, the inventors further generated targeted diversified sequences (randomized sequences, random oligos) for CDR1, CDR2-n, CDR2-c of V_H domain (see **Fig. 2**) and for CDR3 of V_H domain (see **Fig. 3**). The process of generating recombinant scFv nucleic acids using the random oligos of CDR1, CDR2-n, CDR2-c, CDR3 of V_H domain, and CDR3 of V_L domain is described above and also shown in the schematic diagram in **Fig. 4**. A high-diversity library was constructed as exemplarily shown in **Fig. 5** and discussed in more detail above.

[0064] Using the targeted diversification scheme and methods of generating recombinant scFv nucleic acids as described in Figs. 1-5, the inventors generated a high-diversity library and isolated thereform a recombinant α-B7-$H4_{801}$ (α-B7-H4, clone number 801) binder. The stability of the recombinant α-B7-$H4_{801}$ was determined by analytical size exclusion

chromatography over 15 min to evaluate any degradation or deformation of the antibody. As shown in **Fig. 6**, the eluate of α-B7-H4$_{801}$ shows a single peak without any significant smaller peaks, indicating the α-B7-H4$_{801}$ binder generated by methods described above could produce scFv or an antibody with high stability.

[0065] The inventors found that the recombinant α-B7-H4$_{801}$ comprises antibody components of substantially similar to other commercially available α-B7-H4 antibodies (Rituxan®, LEAF®). The fragments of the recombinant α-B7-H4$_{801}$ and two commercially available α-B7-H4 antibodies (Rituxan®, LEAF®) were analyzed via Capillary electrophoresis sodium dodecyl sulfate (CE-SDS). As shown in **Fig. 7**, CE-SDS separation of recombinant α-B7-H4$_{801}$ antibody and two commercially available α-B7-H4 antibodies (Rituxan®, LEAF®) fragments show two profound peaks, each corresponds to light chain (middle peak) and glycosylated heavy chain (right peak). Left peak indicates the location of a 10 Kd standard marker for the CE-SDS analysis.

[0066] The inventors further found that various recombinant α-B7-H4 antibodies may show different binding characters (e.g., affinities, specificities, etc.) to the target ligand. **Fig. 8** shows two recombinant α-B7-H4 antibodies, α-B7-H4$_{801}$ and α-B7-H4$_{817}$ that are tested for binding with B7-H4 expressing 293T cells, measured by mean fluorescence intensity (MFI). The results show that α-B7-H4$_{801}$ antibodies have higher binding affinity to B7-H4 expressing 293T cells compared to α-B7-H4$_{817}$ antibodies, indicating differently randomized CDR domains may render different binding affinities to the ligand. The right most panels show the control experiment with nonspecific human IgG1 (hIgG1).

[0067] The recombinant α-B7-H4 antibodies were further tested to determine specific and effective binding to the ligands (B7-H4) expressed on the antigen presenting cells (APCs) using flow cytometry. As shown in **Fig. 9**, the recombinant α-B7-H4 antibodies could specifically bind to B7-H4 ligands (separating the peak out from nonspecific isotype binding), indicating that the recombinant α-B7-H4 antibodies are fully functional.

[0068] The inventors also found that scFv peptide against B7-H4 (scFv B7-H4$_{801}$) and recombinant α-B7-H4 antibodies (IgG α-B7-H4$_{801}$) generated by the same scFv peptide with the scFv B7-H4$_{801}$ are functionally compatible using the surface plasmon resonance assay. In this assay, Flag-tagged scFv B7-H4$_{801}$ are immobilized on the surface via α-Flag biotinylated antibody, which is coupled with surface-linked neutravidin. The surface immobilized scFv

B7-H4$_{801}$ peptides are then contacted with analyte including B7-H4. Similar assay was performed with α-B7-H4 antibodies. As shown in **Fig. 10** and **Table 3**, scFv B7-H4$_{801}$ and IgG α-B7-H4$_{801}$ shows substantially similar affinity and binding characteristics to B7-H4, indicating that they are functionally compatible. Further, as the binding affinity of in vitro translated peptide (scFv) can be directly measured without grafting the peptide into an antibody backbone, more recombinant scFv nucleic acids in the expression library can be screened efficiently.

	Ka	Kd	KD	Res sd
IgG	$1.2e^{6}$	$2.0e^{-4}$	175 pm	0.391
scFv	$1.2e^{6}$	$1.7e^{-4}$	141 pm	0.353

Table 3

[0069] Among a plurality of scFv peptides against B7-H4 having various random sequence cassettes in CDR1-3 of V_H and CDR3 of V_L, the inventors examined whether similarities in specific domains (specific random sequence cassettes) may render the scFv peptides to have similar binding characteristics to the ligand. Five scFv peptides (801, 802, 905, 906, and 817) were examined for their binding affinities to B7-H4. Among those, as shown in Table 4, four scFv peptides (clone 801, 802, 905, 906) have similar CDR3 sequences. Those four scFv peptides having similar random sequence cassettes in CDR3 of V_H show similar binding affinities to B7-H4 (as shown in Table 5) in both 25°C and 37°C, indicating that at least in scFv peptides against B7-H4, sequences in CDR3 of V_H may be critical in binding to the ligand.

Clone	CDR1	CDR2	CDR3	CDR L3
801	NSYAMH (SEQ ID NO:26)	AISGNGGSTR (SEQ ID NO:27)	DRFRKVHG (SEQ ID NO:28)	DATFPL (SEQ ID NO:29)
802	GSYAMH (SEQ ID NO:30)	AISGSGGSTR (SEQ ID NO:31)	DLYRRVHG (SEQ ID NO:32)	DYGFPL (SEQ ID NO:33)
905	SSYLMH (SEQ ID NO:34)	VISGSGGSTR (SEQ ID NO:35)	DLYRRVAG (SEQ ID NO:36)	DYALPL (SEQ ID NO:37)
906	SNYAMH (SEQ ID NO:38)	AISGNGGSTH (SEQ ID NO:39)	DRFRRVYG (SEQ ID NO:40)	DYTFPL (SEQ ID NO:41)
817	SSYAMH (SEQ ID NO:42)	AISGSGGSTR (SEQ ID NO:43)	GRWSKWG (SEQ ID NO:44)	TDNFPY (SEQ ID NO:45)

Table 4

Temp	scFv	ka	kd	KD
25°C	801	1.20E+06	2.00E-04	174 pM
	802	4.50E+05	2.40E-05	54 pM

	905	4.10E+05	1.20E-04	290 pM
	906	1.70E+05	1.00E-05	59 pM
37°C	**801**	**6.10E+05**	**7.30E-04**	**1.2 nM**
	802	5.70E+05	5.50E-04	1.0 nM
	905	5.80E+05	9.70E-04	1.7 nM
	906	2.80E+05	3.80E-04	1.4 nM

Table 5

[0070] The inventors also generated a plurality of scFv peptides binding to interleukin-8 (IL-8) (scFv IL-8) using the sub-libraries and expression library, and examined the affinity to IL-8 in different conditions (temperatures and pH). Exemplary scFv IL-8 peptides and their binding affinities measured in various conditions are shown in Table 6. Among the clones shown in Table 6, clones 49-7, 49-1 and 49-12 contain similar V_H CDR3 sequences, and clones 49-19, 49-37, and 49-25 contain similar V_H CDR3 sequences. In addition, clones 49-3 and 43-2 contain similar V_H CDR3 sequences. In contrast to the scFv peptides against B7-H4, the inventors found that the binding affinity of scFv IL-8 peptides may not be critically dependent on the similarities in random sequences in CDR3 of V_H. For example, while clone 49-18, 49-37, and 49-25 contain similar V_H CDR3 sequences, the binding affinity (unit measured in K_D X 10^{-9} M) of those sequences varies between 0.894 X 10^{-9} M and 25 X 10^{-9} M.

clone	**count**	**25°C**	**pH 6**	**25°C**	**pH 6**	**37°C**
49-31	1/36	0.012	0.0025			
49-22	3/36	0.113	0.328			
49-7	1/36	0.166	0.462			
49-32	1/36	0.239	0.714			
49-34	1/36	0.618	0.342			
49-18	1/36	0.894	2.23			
49-3	4/36	1.26	6.68	2.14	3.14	9.19
43-2	5/16	1.41	1.3	0.79	0.96	0.89
49-37	1/36	1.46	4.01			
43-12	3/16	1.5	11.04			
49-10	6/36	1.65	8.58	2.21	8.7	3.45
49-1	1/36	2.66	6.13			
49-6	1/36	4.8	17.6			
49-12	3/36	10.1	11.9			
49-25	2/36	25	7.26			

Table 6

[0071] The inventors further tested whether the scFv IL-8 can effectively trap IL-8 to thereby neutralize the effect of IL-8 by measuring neutrophil size. Generally, neutrophils are enlarged (*e.g.*, having a larger diameter, etc.) upon being stimulated by IL-8 (as shown in **Fig. 11**). The inventors found that such IL-8 effect on neutrophil enlargement could be largely abolished upon addition of the recombinant α-IL-8 antibody (mAb αIL-8_{201}, as shown in **Fig. 12**, upper-left graph) or several scFv IL-8 peptides (αIL-$8_{\#2}$, αIL-$8_{49\text{-}3}$, αIL-$8_{49\text{-}10}$, as shown in Fig. 12, lower graphs), indicating that the scFv IL-8 peptides could effectively neutralize the effect of IL-8 by binding to free IL-8 in the media.

[0072] IL-8 is a neutrophil chemotactic factor that causes neutrophils to migrate toward the site of IL-8 release (e.g., site of infection). In order to evaluate the functional effect of scFv IL-8 peptides, neutrophils were placed on the bottom of the insert having a porous membrane and placed in the media including various concentration of IL-8 such that attracted neutrophils by IL-8 can trans-migrate out of the insert through the porous membrane toward the media. As shown in **Fig. 13**, number of migrated neutrophils increased by increasing IL-8 concentration in the media. Interestingly, such IL-8 effect has almost completely abolished upon addition of the scFv IL-8 peptide (αIL-$8_{43\text{-}2}$) or the recombinant IL-8 antibody derived from a scFv IL-8 peptide (mAb αIL-8_{201}).

[0073] **Fig.14** depicts further experimental data for a variety of scFvs isolated using the mRNA display library as presented herein. More specifically, each data point represents an scFv for the target indicated at the bottom, and affinity values for each scFv was determined. As can be readily seen, the (same) library yielded multiple high-affinity binders for a variety of distinct targets, with all of the bonders in the sub-microM, and many in the sub-nanoM affinity range. Moreover, the inventors also studies whether the affinity of the scFvs could be preserved upon CDR grafting onto a human IgG. Fig.15 depicts exemplary results for 29 CDR grafting experiments for selected scFv that were grafted into a human IgG1 scaffold. As can be seen from the results in Fig. 15, the humanized IgG1 antibodies retained high specificity and affinity (typically within one order of magnitude).

[0074] It should be apparent to those skilled in the art that many more modifications besides those already described are possible without departing from the inventive concepts herein. The inventive subject matter, therefore, is not to be restricted except in the scope of the appended claims. Moreover, in interpreting both the specification and the claims, all terms should be interpreted in the broadest possible manner consistent with the context. In

particular, the terms “comprises” and “comprising” should be interpreted as referring to elements, components, or steps in a non-exclusive manner, indicating that the referenced elements, components, or steps may be present, or utilized, or combined with other elements, components, or steps that are not expressly referenced. As used in the description herein and throughout the claims that follow, the meaning of “a,” “an,” and “the” includes plural reference unless the context clearly dictates otherwise. Also, as used in the description herein, the meaning of “in” includes “in” and “on” unless the context clearly dictates otherwise. Where the specification claims refers to at least one of something selected from the group consisting of A, B, C …. and N, the text should be interpreted as requiring only one element from the group, not A plus N, or B plus N, etc.

CLAIMS

What is claimed is:

1. A method of generating a high-diversity nucleic acid library that encodes a plurality of Fab domains, the method comprising:

generating or providing (1) a V_H-CDR1/2 sub-library, (2) a plurality of V_H-CDR3 sub-libraries, and (3) a V_L sub-library, wherein each of the sub-libraries (1)–(3) comprises a plurality of members;

wherein each member of the sub-libraries comprises at least one random cassette that has a plurality of degenerate base positions, and wherein the random cassette is generated using an oligonucleotide selected from SEQ ID NOs: 1–25; and

recombining at least portions of at least two members of the V_H-CDR1/2 sub-library, the plurality of V_H-CDR3 sub-libraries, and the V_L sub-library to form an expression library member in an expression library, wherein the expression library comprises a plurality of expression library members, each expression library member encoding a distinct Fab domain.

2. The method of claim 1, wherein the plurality of members of the V_H-CDR1/2 sub-library comprises a random cassette corresponding to at least one of a portion of V_H CDR1 and a portion of V_H CDR2.

3. The method of claim 1, wherein the plurality of members of the V_H-CDR1/2 sub-library comprises a plurality of random cassettes corresponding to at least one of a portion of V_H CDR1 and a portion of V_H CDR2.

4. The method of claim 3, wherein the plurality of members of the V_H-CDR1/2 sub-library comprises a plurality of random cassettes corresponding to at least the portion of V_H CDR2.

5. The method of claim 1, wherein the plurality of the members of the V_H-CDR3 sub-libraries comprises a random cassette corresponding to at least a portion of V_H CDR3.

6. The method of claim 1, wherein at least two random cassettes of members of the V_H-CDR3 sub-libraries encodes peptides with different lengths.

7. The method according to any one of claims 1 to 6, wherein the plurality of the members of the V_L sub-library comprises a random cassette at a portion of V_L CDR3.

8. The method according to any one of claims 1 to 7, wherein the recombining comprises isolating the at least portions of the members of the V_H-CDR1/2 sub-library and one of the plurality of V_H-CDR3 sub-libraries and fusing together to form a V_H domain library member in a V_H domain library, wherein the V_H domain library comprises a plurality of V_H domain library members.

9. The method of claim 8, further comprising isolating at least a portion of the member of the V_L sub-library and fusing the portion of the member of the V_L sub-library with one of the V_H domain library members to form the expression library member.

10. The method according to any one of claims 1 to 7, wherein the recombining comprises isolating the at least portions of the members of the V_H-CDR1/2 sub-library and one of the plurality of V_H-CDR3 sub-libraries and fusing together to form a first group of expression library members.

11. The method of claim 10, further comprising a second group of expression library members, wherein the second group comprises at least portions of the members of the V_L sub-library.

12. The method according to any one of claims 1 to 11, further comprising:

transcribing the expression library member into an mRNA fragment; and

coupling a puromycin molecule at 3'-end of the mRNA fragment .

13. A method of generating a recombinant antibody, comprising:

generating or providing (1) a V_H-CDR1/2 sub-library, (2) a plurality of V_H-CDR3 sub-libraries, and (3) a V_L sub-library, wherein each of the sub-libraries (1)–(3) comprises a plurality of members;

wherein each member of the sub-libraries comprises at least one random cassette that has a plurality of degenerate base positions, and wherein the random cassette is generated using an oligonucleotide selected from SEQ ID NOs: 1–25;

recombining at least portions of at least two members of the V_H-CDR1/2 sub-library, the plurality of V_H-CDR3 sub-libraries, and the V_L sub-library to form an expression library member in an expression library, wherein the expression library comprises a plurality of expression library members, each expression library member encoding a distinct antibody or antibody fragment; and

generating the recombinant viral vector comprising at least one expression library member.

14. A method of generating a recombinant Fab domain, comprising:

generating or providing (1) a V_H-CDR1/2 sub-library, (2) a plurality of V_H-CDR3 sub-libraries, and (3) a V_L sub-library, wherein each of the sub-libraries (1)–(3) comprises a plurality of members;

wherein each member of the sub-libraries comprises at least one random cassette that has a plurality of degenerate base positions, and wherein the random cassette is generated using an oligonucleotide selected from SEQ ID NOs: 1–25;

recombining at least portions of at least two members of the V_H-CDR1/2 sub-library, the plurality of V_H-CDR3 sub-libraries, and the V_L sub-library to form an expression library member in an expression library, wherein the expression library comprises a plurality of expression library members, each expression library member encoding a distinct Fab domain; and

generating the recombinant Fab domain using the expression library member.

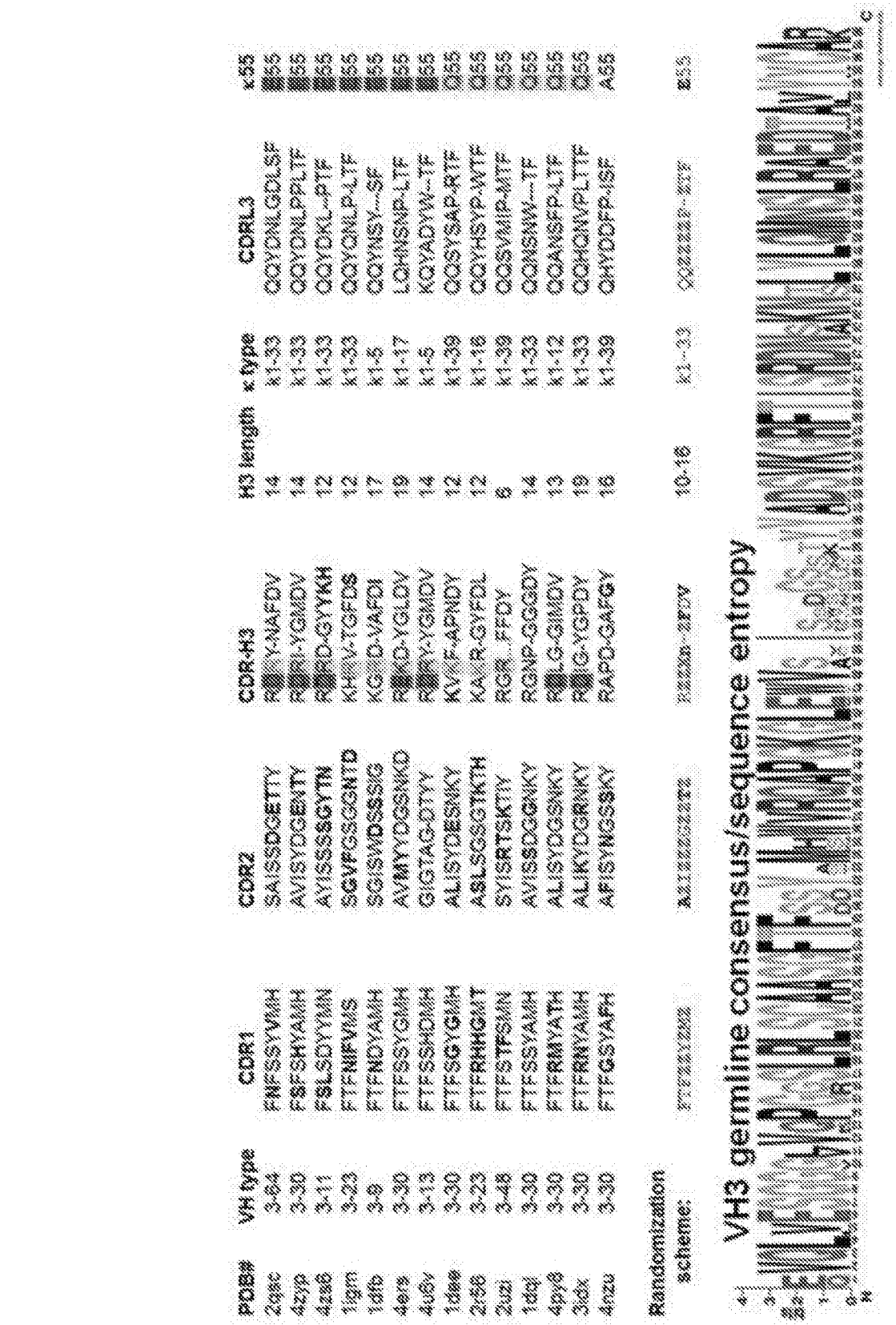
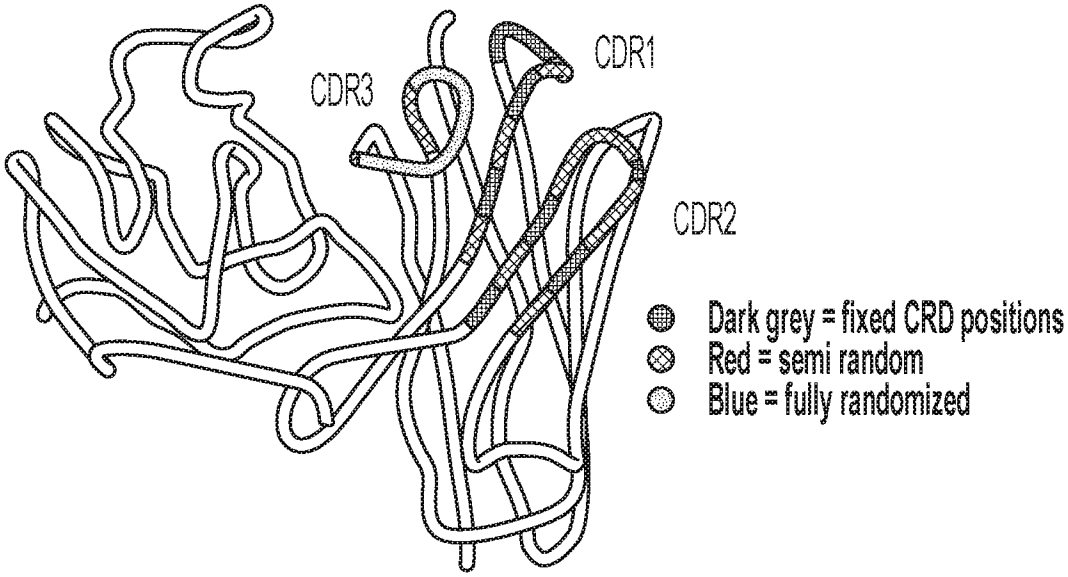

PDB# VH type CDR1 CDR2 CDR-H3 H3 length κ type CDRL3 κ55
2qsc 3-64 FNFSSYVMH SAISSDGETTY R Y-NAFDV 14 k1-33 QQYDNLGDLSF 55
4zyp 3-30 F SFSHYAMH AVISYDGENTY R RI-YGMDV 14 k1-33 QQYDNLPPLTF 55
4za6 3-11 FSLSDYYMN AYISSSSGYTN R RD-GYYKH 12 k1-33 QQYDKL--PTF 55
1igm 3-23 FTFNMFVMS SGVPGSGGMTD KH V-TGFDS 12 k1-33 QQYGNLP-LTF 55
1dfb 3-9 FTFNDYAMH SGISWDSSSIG KG D-VAFDI 17 k1-5 CQYNSY--SF 55
4ers 3-30 FTFSSYGMH AVMYYDGSNKD R D-YGLDV 19 k1-17 LQHNSNP-LTF 55
4u6v 3-13 FTFSSHDMH GIGTAGDTYY R RY-YGMDV 14 k1-5 KQYADYW--TF 55
1dee 3-30 FTFSGYGMH ALISYDESNKY KV F-APNDY 12 k1-39 CQSYSAP-RTF 55
2r56 3-23 FTFRHHGMT ASLSGSGTKTH KA R-GYFDL 12 k1-18 CQYHSYP-WTF 55
2uzi 3-48 FTFSTFSMN SYISRTSKTIY RGR FFDY 8 k1-39 CQSVMIP-MTF 55
1dql 3-30 FTFSSYAMH AVISSDGGNKY RGNP-GGGDY 14 k1-33 CQNSNW---TF 55
4py8 3-30 FTFRMYATH ALISYDGSNKY R LG-GMDV 13 k1-12 CQANSFP-LTF 55
3idx 3-30 FTFRNYAMH ALIKYDGRNKY R G-YGPDY 19 k1-33 CQHQNVPLTTF 55
4nzu 3-30 FTFGSYAPH AFISYWGSSKY RAPD-GAFGY 16 k1-39 QHYDDFP-ISF A55
Randomization scheme:
VH3 germline consensus/sequence entropy

Fig. 1



CDR1 and 2

• CDR1

GGCTTA GGTCTC A TTTC RVTAGTTACGCTATGCAT TGGG C GAGACG AGGTCTGAACGG X=TANDSG
GGCTTA GGTCTC A TTTC TCTRVKTACGCTATGCAT TGGG C GAGACG AGGTCTGAACGG X=TANKDESRG
GGCTTA GGTCTC A TTTC TCTAGTTACKKGATGCAT TGGG C GAGACG AGGTCTGAACGG X=GWLV
GGCTTA GGTCTC A TTTC TCTAGTTACWMTATGCAT TGGG C GAGACG AGGTCTGAACGG X=SYTN
GGCTTA GGTCTC A TTTC TCTAGTTACGCTATGAVT TGGG C GAGACG AGGTCTGAACGG X=STN

• CDR2-n

GGCTTA GGTCTC G TTCA THCATTAGTGGTAGTGGA C GAGACG AGGTCTGAACGG X=YFS
GGCTTA GGTCTC G TTCA VKTATTAGTGGTAGTGGA C GAGACG AGGTCTGAACGG X=VGISLR
GGCTTA GGTCTC G TTCA GCTATTYGGGGTAGTGGA C GAGACG AGGTCTGAACGG X=WR
GGCTTA GGTCTC G TTCA GCTATTDATGGTAATGGA C GAGACG AGGTCTGAACGG X=YND+N53
GGCTTA GGTCTC G TTCA GCTATTAGTWMTAGTGGA C GAGACG AGGTCTGAACGG X=YSTN
GGCTTA GGTCTC G TTCA GCTATTAGTKGGAGTGGA C GAGACG AGGTCTGAACGG X=WG
GGCTTA GGTCTC G TTCA GCTATTAGTGGTRRTGGA C GAGACG AGGTCTGAACGG X=DGSN

• CDR2-c

GGCTTA GGTCTC G TGGA RVKAGTACTTAC TACG C GAGACG AGGTCTGAACGG X=STGANKDE
GGCTTA GGTCTC G TGGA GGTNATACTTAC TACG C GAGACG AGGTCTGAACGG X=YNDH
GGCTTA GGTCTC G TGGA GGTRVAACTTAC TACG C GAGACG AGGTCTGAACGG X=TKREAG
GGCTTA GGTCTC G TGGA GGTAGTACTVRT TACG C GAGACG AGGTCTGAACGG X=DGNSHR

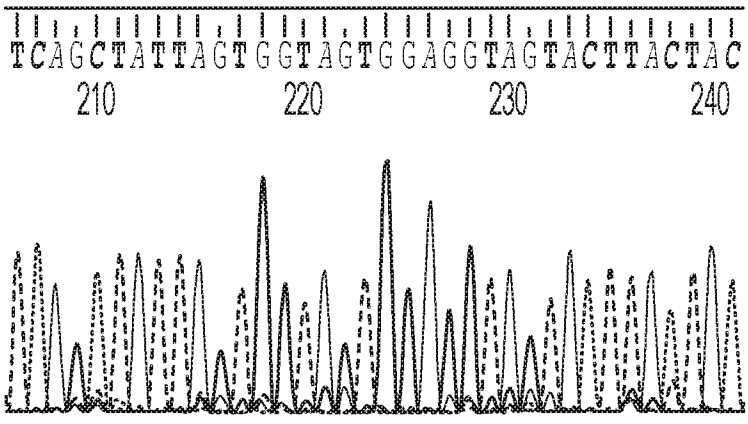


Fig. 2

CDR3 Length	BsaI	G/D R/L	Xaa	G/A	BsmBI	Complexity DNA	Complexity Protein
10	GGCTTA GGTCTC T CCGT	GRT CKC	(NNK)-4	GST	TTCG C GAGACG AGGTCTGAACGG	$8x10^6$	$1x10^6$
11	GGCTTA GGTCTC T CCGT	GRT CKC	(NNK)-5	GST	TTCG C GAGACG AGGTCTGAACGG	$3x10^8$	$3x10^7$
12	GGCTTA GGTCTC T CCGT	GRT CKC	(NNK)-6	GST	TTCG C GAGACG AGGTCTGAACGG	$9x10^9$	$5x10^8$
13	GGCTTA GGTCTC T CCGT	GRT CKC	(NNK)-7	GST	TTCG C GAGACG AGGTCTGAACGG	$3x10^{11}$	$1x10^{10}$
14	GGCTTA GGTCTC T CCGT	GRT CKC	(NNK)-8	GST	TTCG C GAGACG AGGTCTGAACGG	$9x10^{12}$	$2x10^{11}$
15	GGCTTA GGTCTC T CCGT	GRT CKC	(NNK)-9	GST	TTCG C GAGACG AGGTCTGAACGG	$3x10^{14}$	$4x10^{12}$
16	GGCTTA GGTCTC T CCGT	GRT CKC	(NNK)-10	GST	TTCG C GAGACG AGGTCTGAACGG	$9x10^{15}$	$8x10^{13}$

Fig. 3

CDR L3
RANDL3-1
GGCTTA GGTCTC T GCAG DSG DMT RVT DSG CCT TWC ACTT C GAGACG AGGTCTGAACGG
RANDL3-2
GGCTTA GGTCTC T GCAG BWT DMT RVT DSG CCT TWC ACTT C GAGACG AGGTCTGAACGG
RANDL3-3
GGCTTA GGTCTC T GCAG DSG DMT RVT NWT CCT TWC ACTT C GAGACG AGGTCTGAACGG
RANDL3-4
GGCTTA GGTCTC T GCAG BWT DMT RVT NWT CCT TWC ACTT C GAGACG AGGTCTGAACGG
RANDL3-5
GGCTTA GGTCTC T GCAG DSG DMT RVT DSG CCT YKG ACTT C GAGACG AGGTCTGAACGG
RANDL3-6
GGCTTA GGTCTC T GCAG BWT DMT RVT DSG CCT YKG ACTT C GAGACG AGGTCTGAACGG
RANDL3-7
GGCTTA GGTCTC T GCAG DSG DMT RVT NWT CCT YKG ACTT C GAGACG AGGTCTGAACGG
RANDL3-8
GGCTTA GGTCTC T GCAG BWT DMT RVT NWT CCT YKG ACTT C GAGACG AGGTCTGAACGG
Q Q
Y D L A H S F R T W G
Y L N D S T A
S N T A D G
Y F A L T S H W I N R V D G
P L Y W F R T
100
110
120
(reverse complement)

Fig. 4

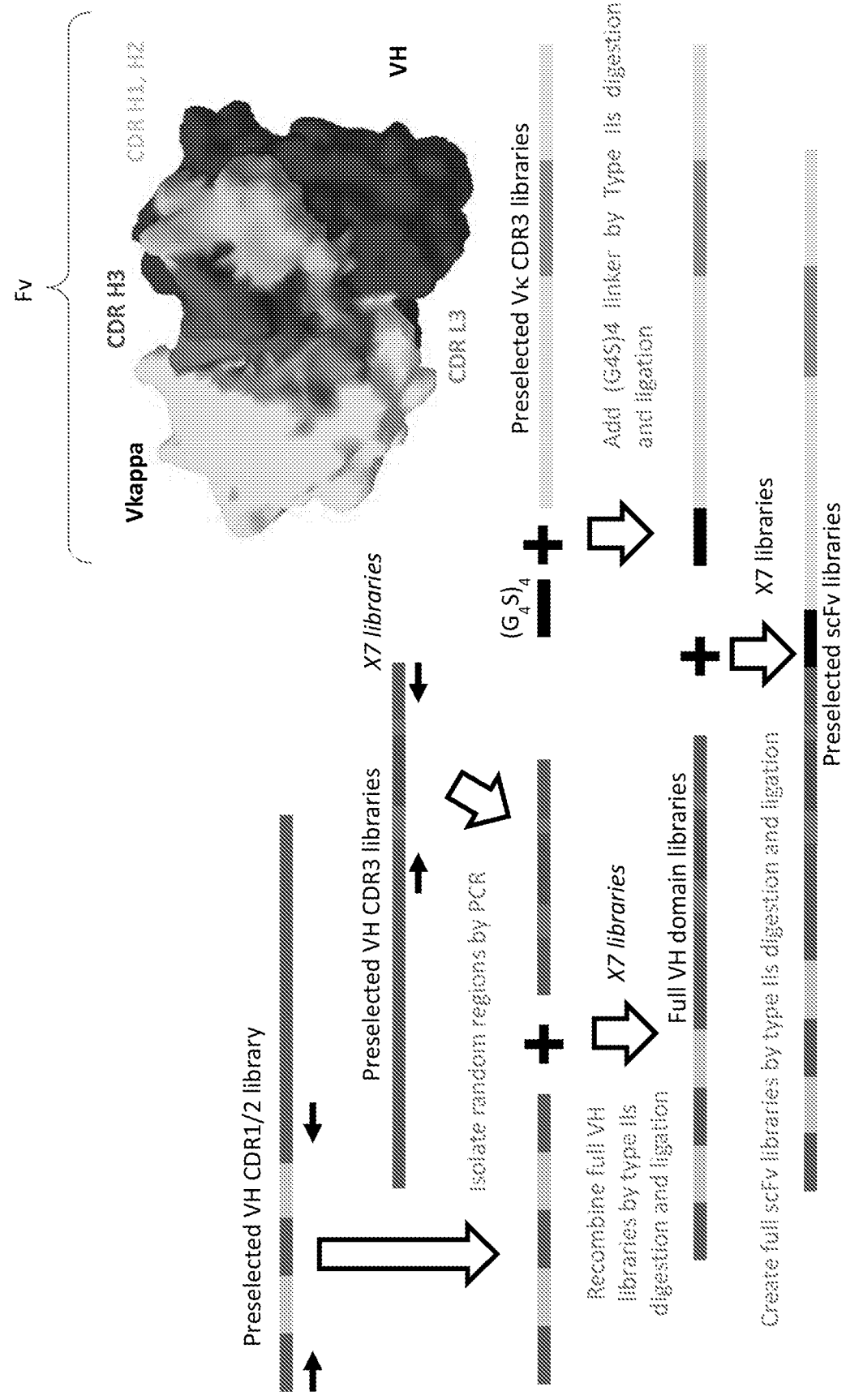
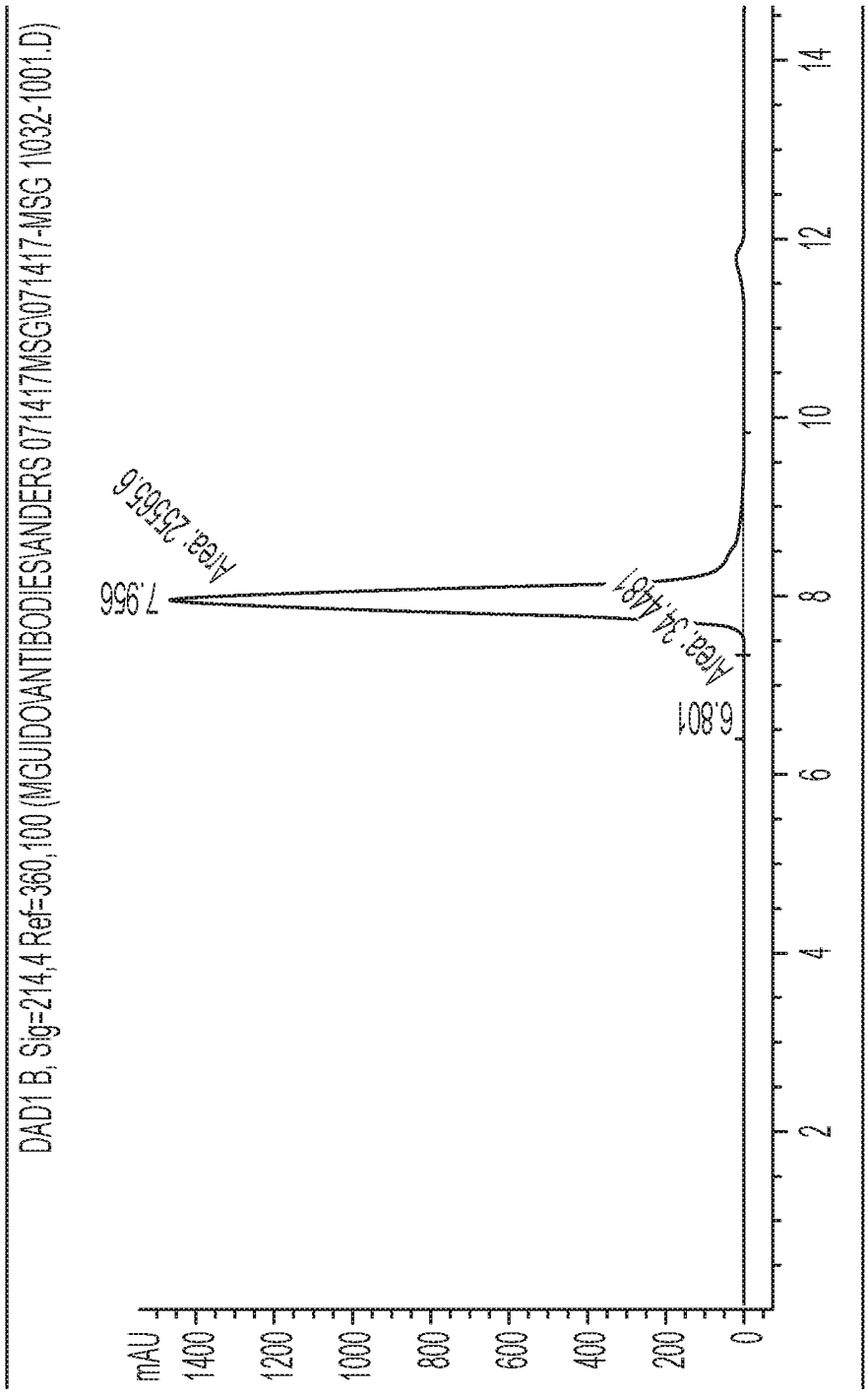

Fv
CDR H3
CDR H1, H2
VH
Vkappa
CDR L3
Preselected VH CDR1/2 library
Preselected VH CDR3 libraries
X7 libraries
Isolate random regions by PCR
Recombine full VH libraries by type IIs digestion and ligation
X7 libraries
Full VH domain libraries
$(G_4S)_4$
Preselected Vκ CDR3 libraries
Add (G4S)4 linker by Type IIs digestion and ligation
X7 libraries
Create full scFv libraries by type IIs digestion and ligation
Preselected scFv libraries

Fig. 5



PEAK #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	6.719	MF	0.4504	49.97571	1.84921	0.1062
2	7.911	FM	0.3160	4.69964e4	2478.48633	99.8938

Fig. 6

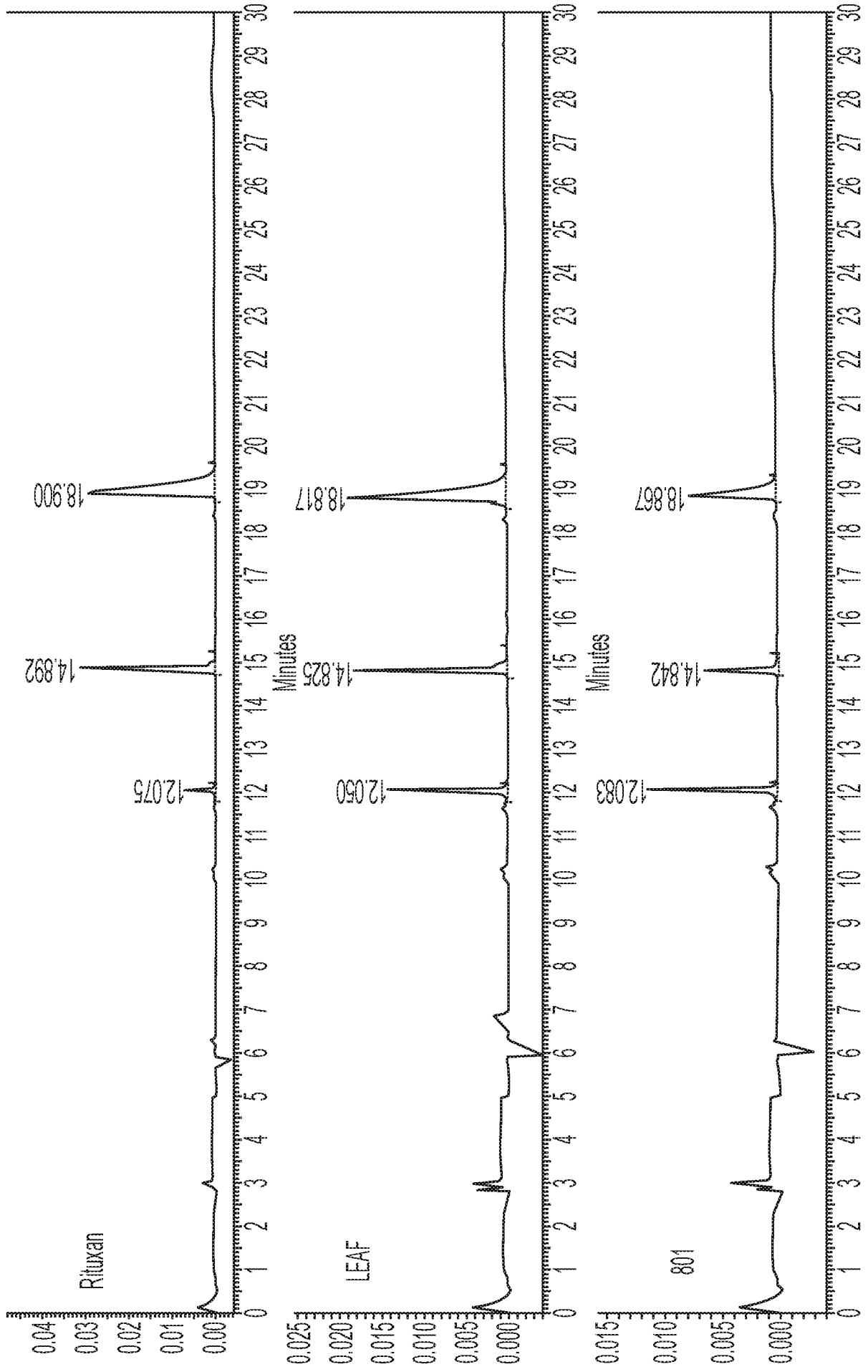

Rituxan
LEAF
801
12.075
14.892
18.900
12.060
14.825
18.817
12.083
14.842
18.867
Minutes
Minutes
Minutes

Fig. 7

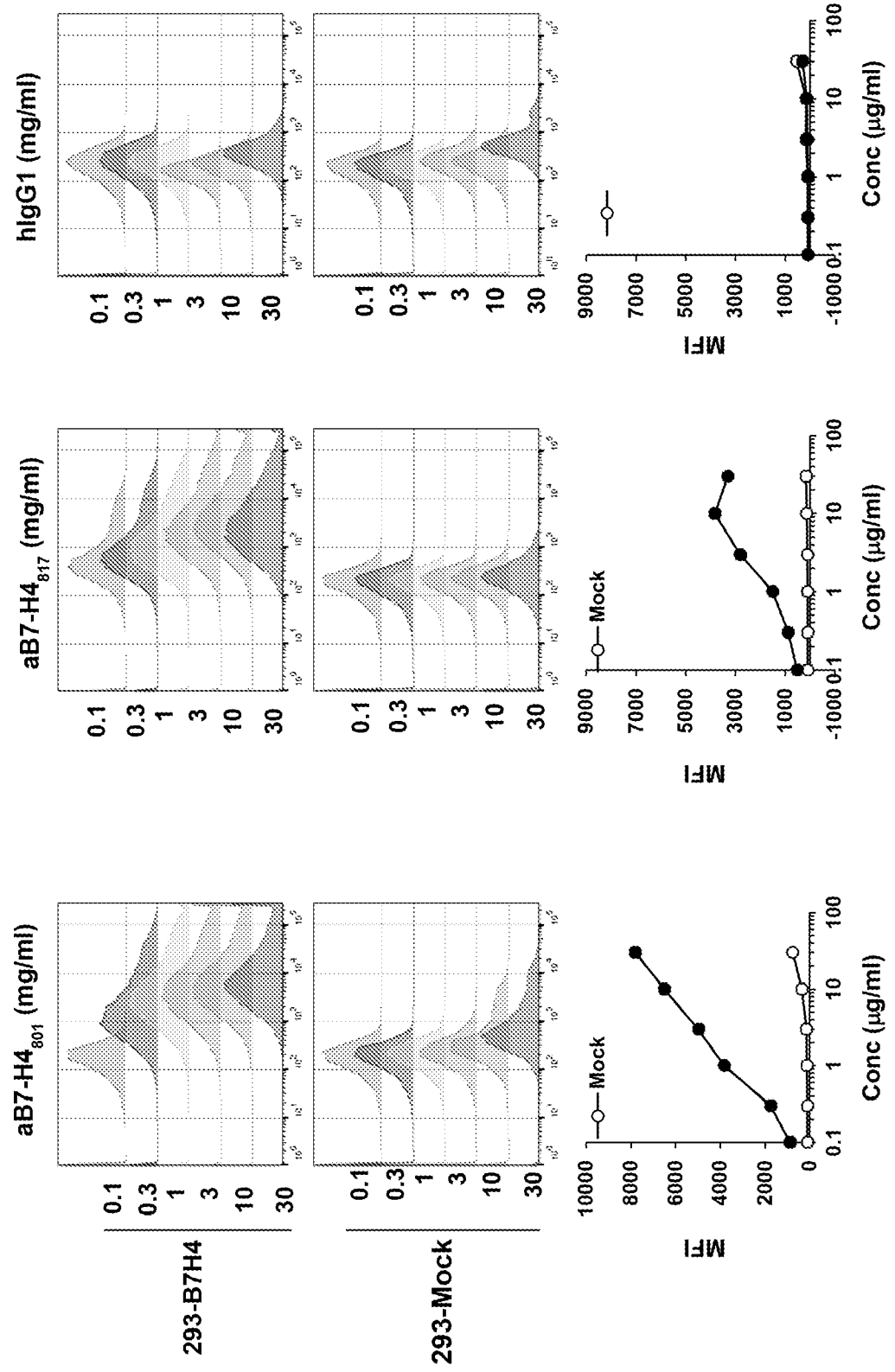

aB7-H4801 (mg/ml)
aB7-H4817 (mg/ml)
hIgG1 (mg/ml)
0.1
0.3
1
3
10
30
293-B7H4
293-Mock
Mock
MFI
Conc (µg/ml)

Fig. 8

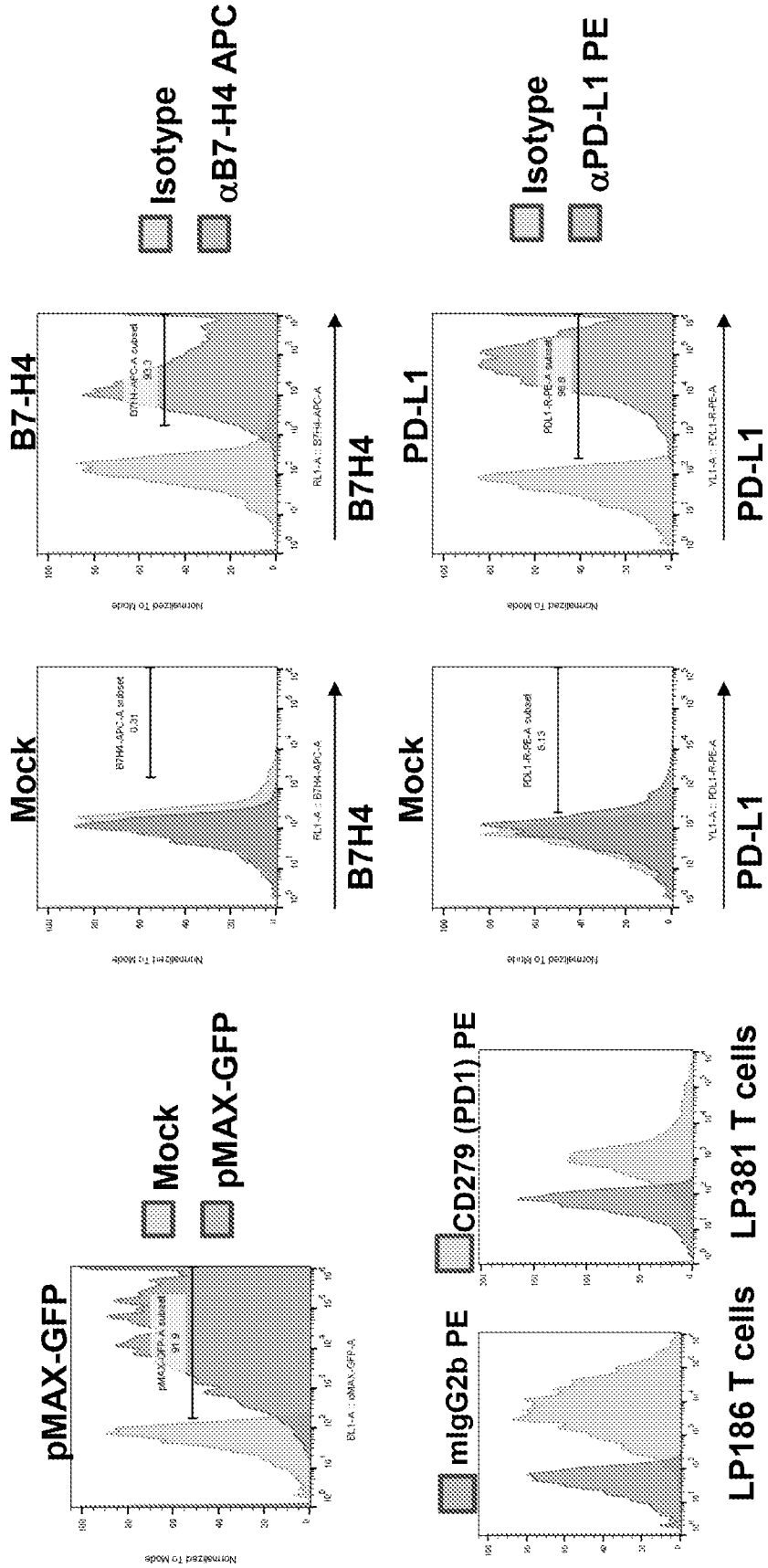
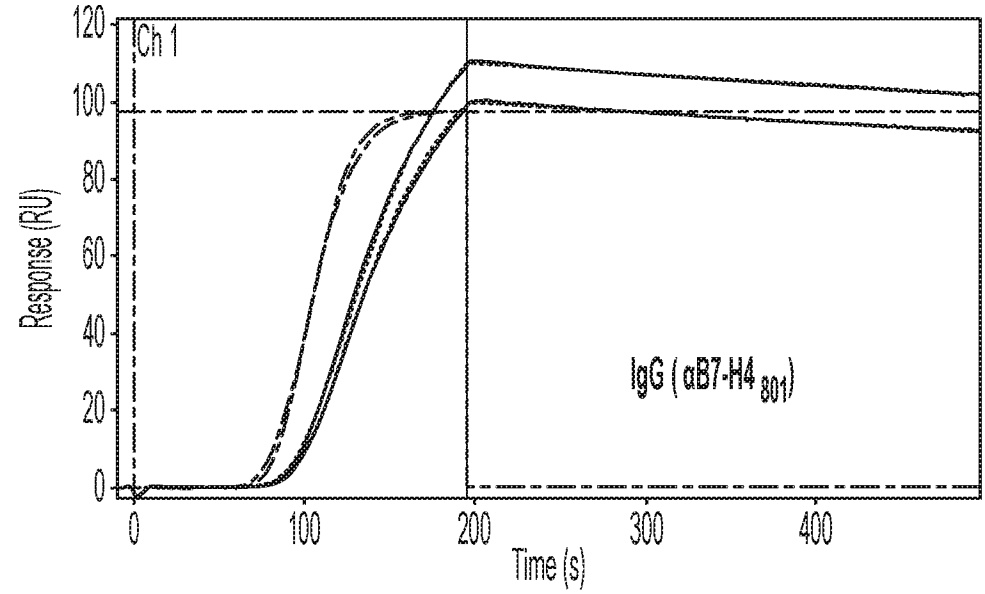

Isotype
αB7-H4 APC
Isotype
αPD-L1 PE
B7-H4
PD-L1
Mock
Mock
B7H4
PD-L1
B7H4
PD-L1
pMAX-GFP
Mock
pMAX-GFP
mIgG2b PE
CD279 (PD1) PE
LP186 T cells
LP381 T cells

Fig. 9


Ch 1
Response (RU)
120
100
80
60
40
20
0
0
100
200
300
400
Time (s)
IgG (αB7-H4 801)

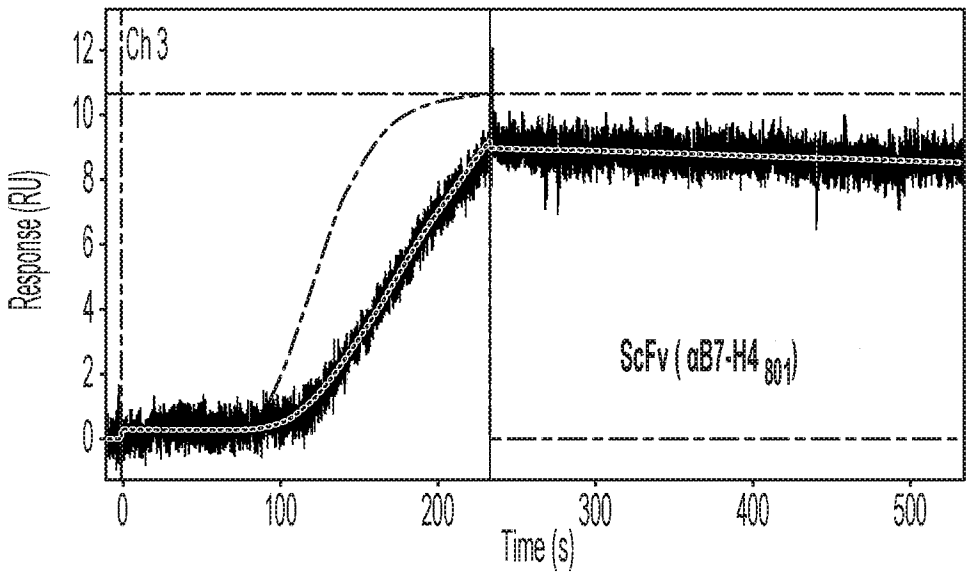
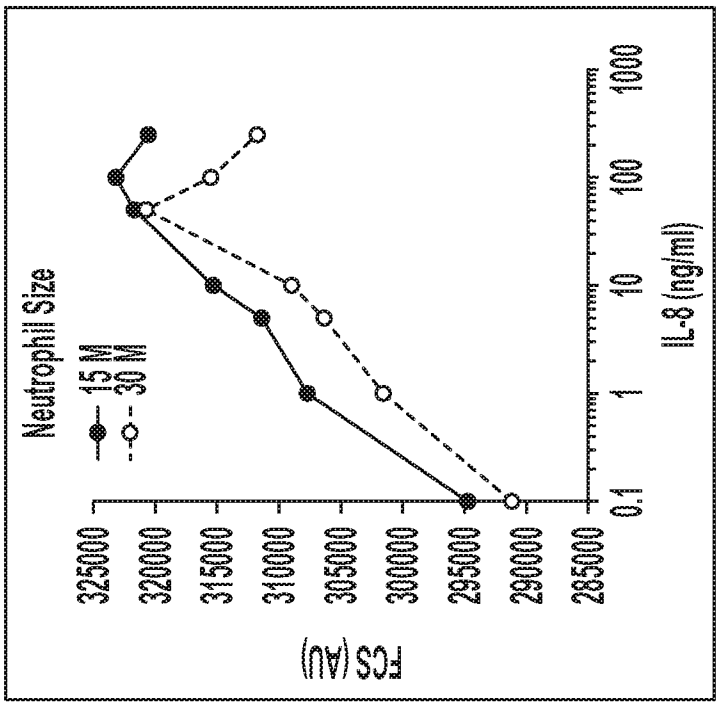
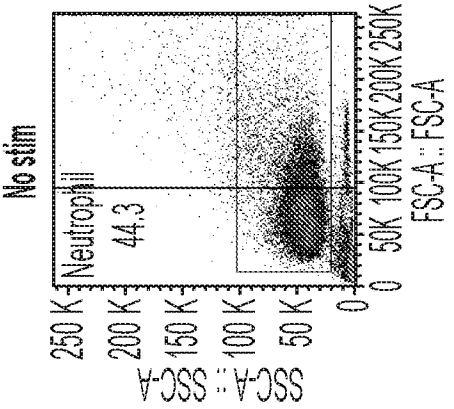
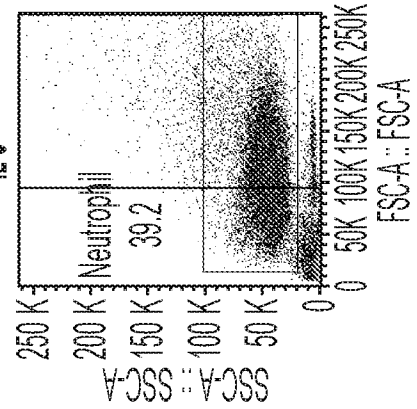

Ch 3
Response (RU)
12
10
8
6
4
2
0
0
100
200
300
400
500
Time (s)
ScFv (αB7-H4 801)

Fig. 10


Neutrophil Size
15 M
30 M
FCS (AU)
325000
320000
315000
310000
305000
300000
295000
290000
285000
0.1
1
10
100
1000
IL-8 (ng/ml)


No stim
Neutrophil
44.3
SSC-A :: SSC-A
250 K
200 K
150 K
100 K
50 K
0
0 50K 100K 150K 200K 250K
FSC-A :: FSC-A


IL-8
Neutrophil
39.2
SSC-A :: SSC-A
250 K
200 K
150 K
100 K
50 K
0
0 50K 100K 150K 200K 250K
FSC-A :: FSC-A

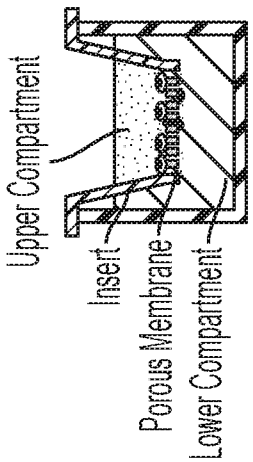

Upper Compartment
Insert
Porous Membrane
Lower Compartment

Fig. 11

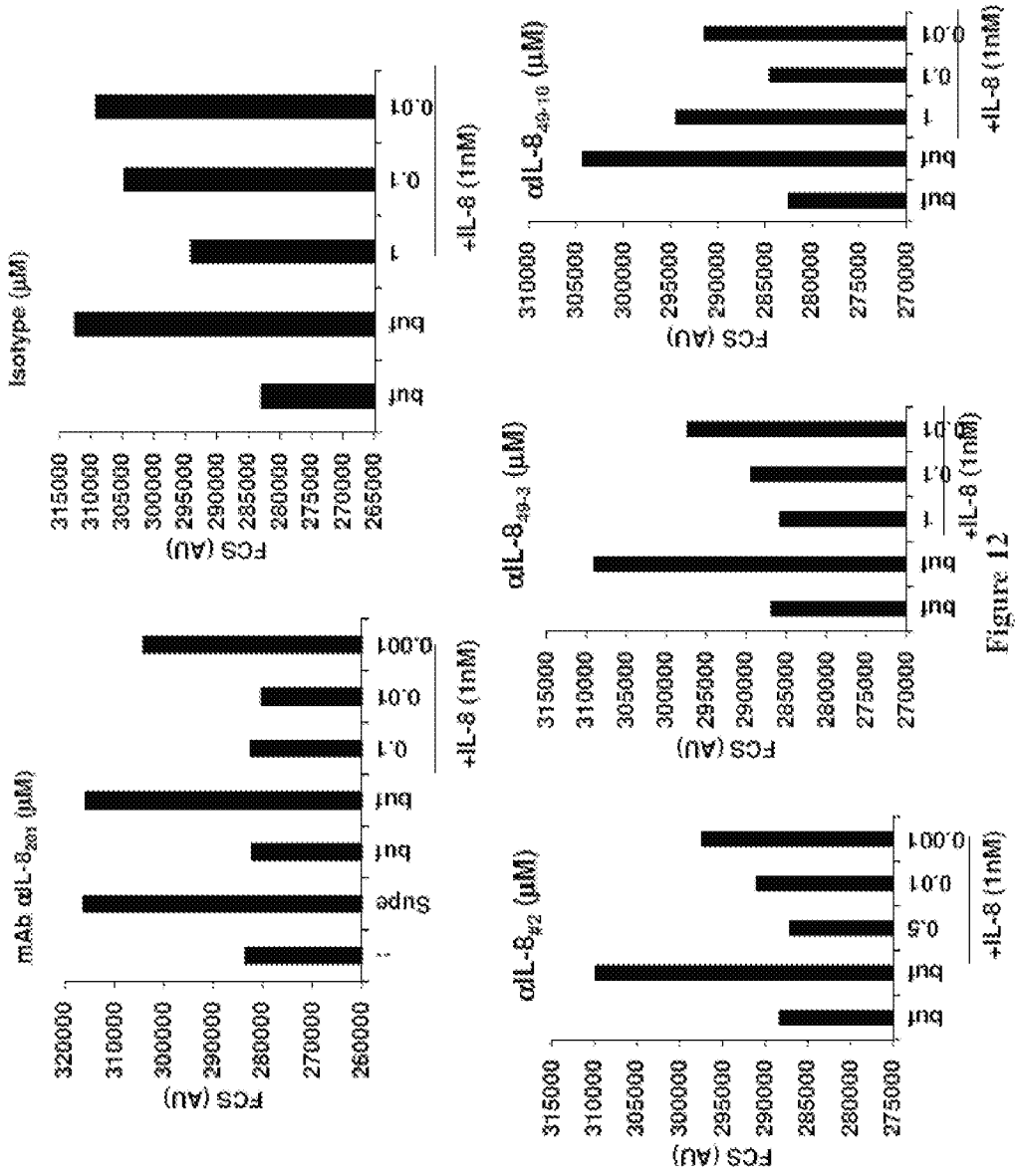

Isotype (μM)
mAb αIL-8$_{281}$ (μM)
αIL-8$_{#2}$ (μM)
αIL-8$_{49-3}$ (μM)
αIL-8$_{49-10}$ (μM)
FCS (AU)
+IL-8 (1nM)
but
Supe
1
0.5
0.1
0.01
0.001

Figure 12

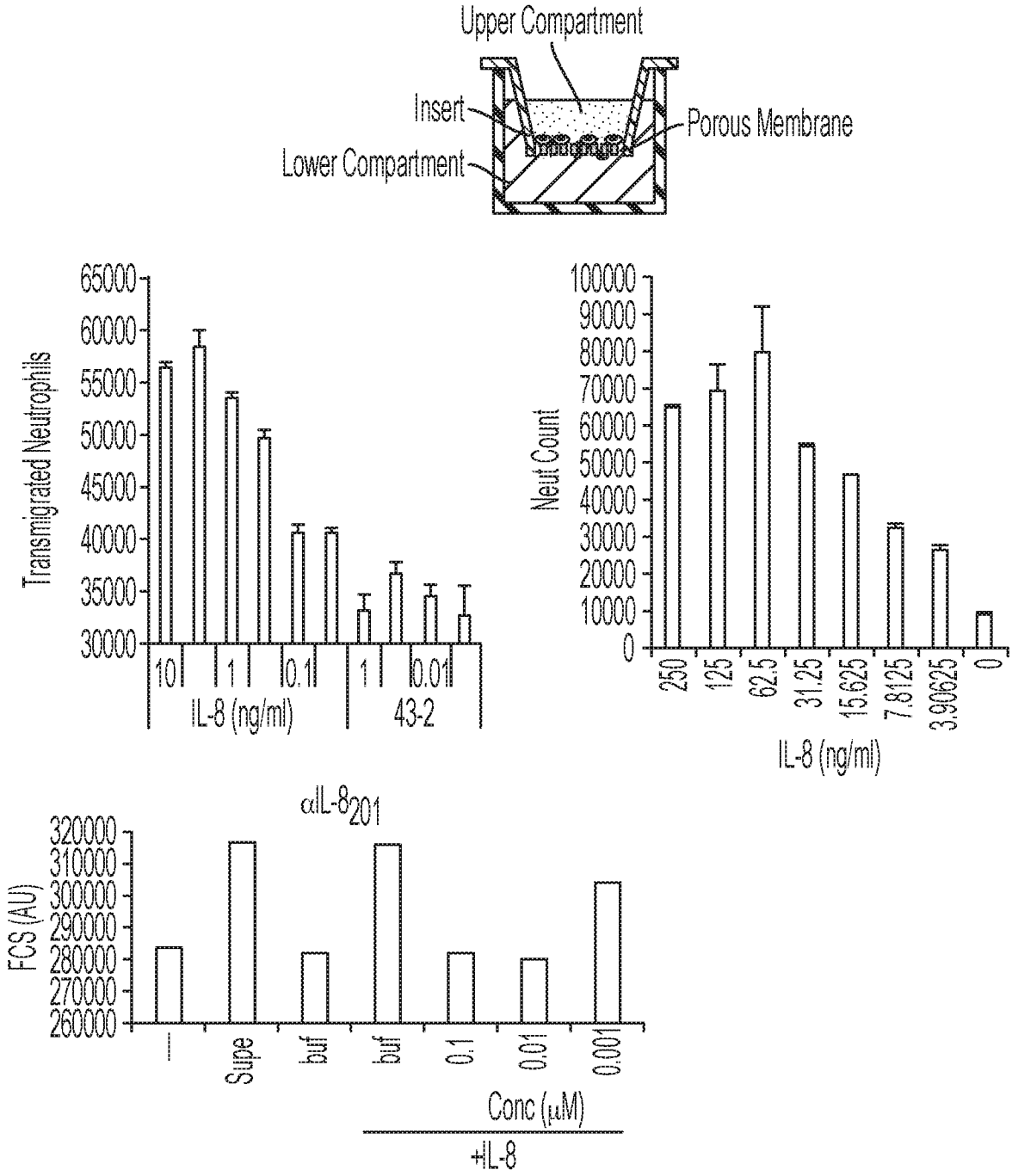

Upper Compartment
Insert
Porous Membrane
Lower Compartment
Transmigrated Neutrophils
65000
60000
55000
50000
45000
40000
35000
30000
10
1
0.1
1
0.01
IL-8 (ng/ml)
43-2
Neut Count
100000
90000
80000
70000
60000
50000
40000
30000
20000
10000
0
250
125
62.5
31.25
15.625
7.8125
3.90625
0
IL-8 (ng/ml)
αIL-8201
FCS (AU)
320000
310000
300000
290000
280000
270000
260000
—
Supe
buf
buf
0.1
0.01
0.001
Conc (μM)
+IL-8

Fig. 13

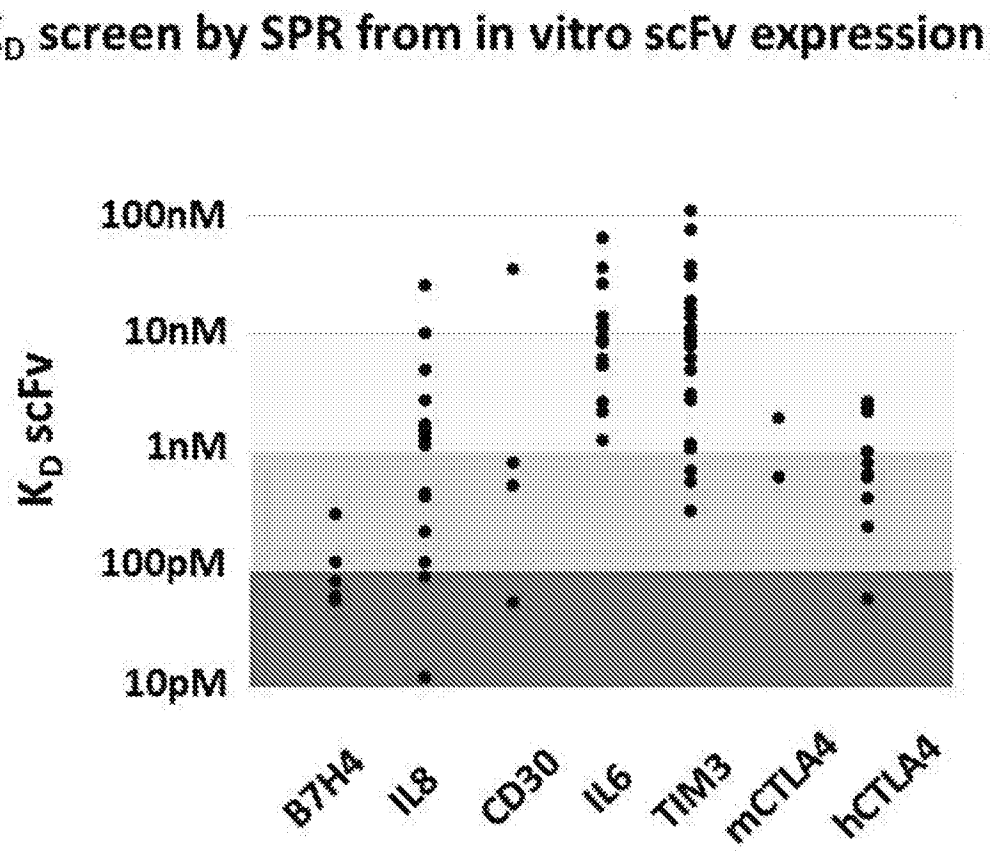

K_D screen by SPR from in vitro scFv expression
100nM
10nM
1nM
100pM
10pM
K_D scFv
B7H4
IL8
CD30
IL6
TIM3
mCTLA4
hCTLA4

Fig. 14

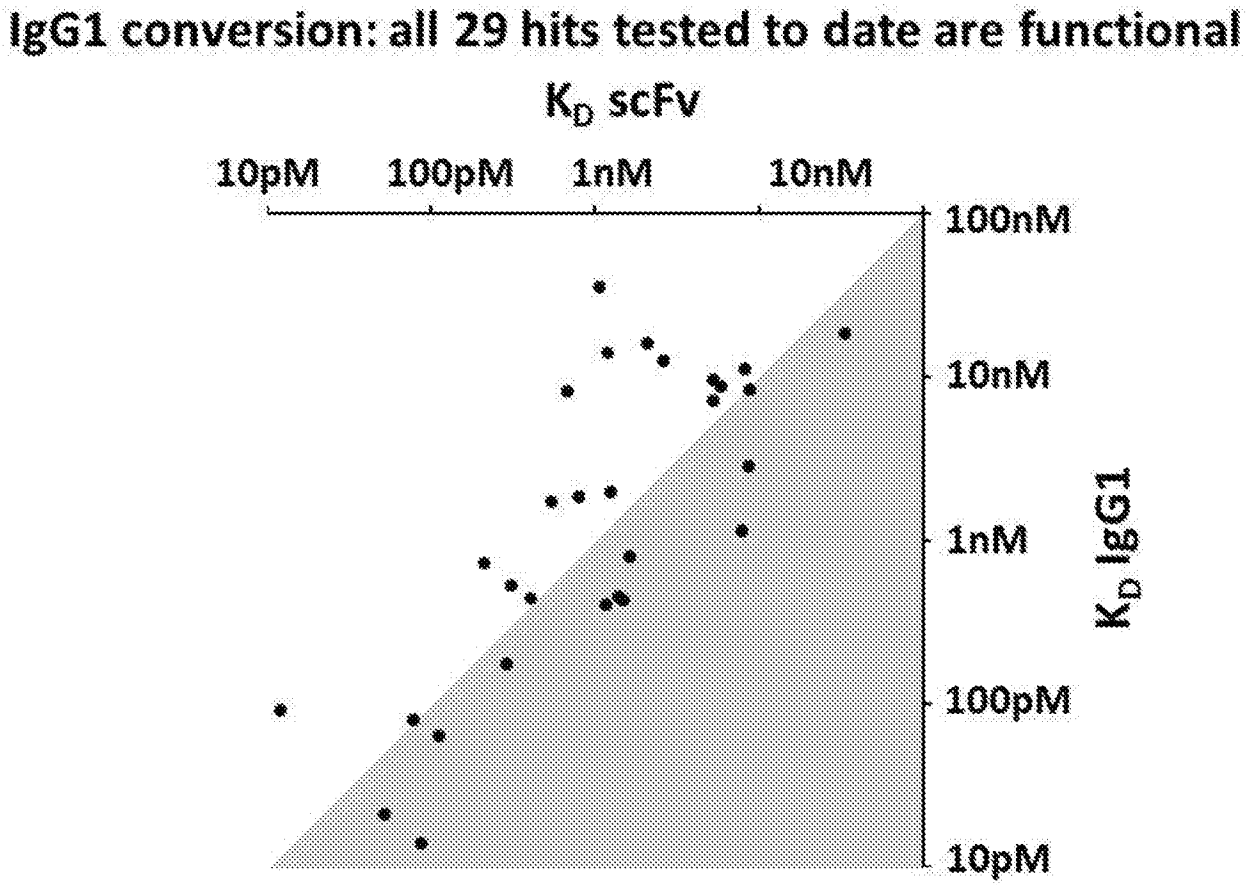

IgG1 conversion: all 29 hits tested to date are functional
K_D scFv
10pM
100pM
1nM
10nM
100nM
10nM
1nM
100pM
10pM
K_D IgG1

Fig.15

SEQUENCE LISTING

<110> NantBio, Inc.

<120> mRNA DISPLAY ANTIBODY LIBRARY AND METHODS

<130> 102538.0050PCT

<150> US62/588,914
<151> 2017-11-20

<160> 45

<170> PatentIn version 3.5

<210> 1
<211> 58
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence for VH CDR1

<400> 1
ggcttaggtc tcatttcrvt agttacgcta tgcattgggc gagacgaggt ctgaacgg 58

<210> 2
<211> 58
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence for VH CDR1

<400> 2
ggcttaggtc tcatttctct rvktacgcta tgcattgggc gagacgaggt ctgaacgg 58

<210> 3
<211> 58
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence for VH CDR1

<400> 3
ggcttaggtc tcatttctct agttackkga tgcattgggc gagacgaggt ctgaacgg 58

<210> 4
<211> 58

<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence for VH CDR1

<400> 4
ggcttaggtc tcatttctct agttacwmta tgcattgggc gagacgaggt ctgaacgg 58

<210> 5
<211> 58
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence for VH CDR1

<400> 5
ggcttaggtc tcatttctct agttacgcta tgavttgggc gagacgaggt ctgaacgg 58

<210> 6
<211> 54
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence for VH CDR2-n

<400> 6
ggcttaggtc tcgttcathc attagtggta gtggacgaga cgaggtctga acgg 54

<210> 7
<211> 54
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence for VH CDR2-n

<400> 7
ggcttaggtc tcgttcavkt attagtggta gtggacgaga cgaggtctga acgg 54

<210> 8
<211> 54
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence for VH CDR2-n

<400> 8
ggcttaggtc tcgttcagct attyggggta gtggacgaga cgaggtctga acgg 54

<210> 9
<211> 54
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence for VH CDR2-n

<400> 9
ggcttaggtc tcgttcagct attdatggta atggacgaga cgaggtctga acgg 54

<210> 10
<211> 54
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence for VH CDR2-n

<400> 10
ggcttaggtc tcgttcagct attagtwmta gtggacgaga cgaggtctga acgg 54

<210> 11
<211> 54
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence for VH CDR2-n

<400> 11
ggcttaggtc tcgttcagct attagtkgga gtggacgaga cgaggtctga acgg 54

<210> 12
<211> 54
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence for VH CDR2-n

<400> 12
ggcttaggtc tcgttcagct attagtggtr rtggacgaga cgaggtctga acgg 54

<210> 13
<211> 52

<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence for VH CDR2-c

<400> 13
ggcttaggtc tcgtggarvk agtacttact acgcgagacg aggtctgaac gg 52

<210> 14
<211> 52
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence for VH CDR2-c

<220>
<221> misc_feature
<222> (21)..(21)
<223> n is a, c, g, or t

<400> 14
ggcttaggtc tcgtggaggt natacttact acgcgagacg aggtctgaac gg 52

<210> 15
<211> 52
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence for VH CDR2-c

<400> 15
ggcttaggtc tcgtggaggt rvaacttact acgcgagacg aggtctgaac gg 52

<210> 16
<211> 52
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence for VH CDR2-c

<400> 16
ggcttaggtc tcgtggaggt agtactvrtt acgcgagacg aggtctgaac gg 52

<210> 17
<211> 53

<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence for VH CDR3

<220>
<221> misc_feature
<222> (24)..(25)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (27)..(27)
<223> n is a, c, g, or t

<400> 17
ggcttaggtc tctccgtgrt ckcnnkngst ttcgcgagac gaggtctgaa cgg 53

<210> 18
<211> 58
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence for VL CDR3

<400> 18
ggcttaggtc tctgcagdsg dmtrvtdsgc cttwcacttc gagacgaggt ctgaacgg 58

<210> 19
<211> 58
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence for VL CDR3

<400> 19
ggcttaggtc tctgcagbwt dmtrvtdsgc cttwcacttc gagacgaggt ctgaacgg 58

<210> 20
<211> 58
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence for VL CDR3

<220>
<221> misc_feature
<222> (27)..(27)
<223> n is a, c, g, or t

<400> 20
ggcttaggtc tctgcagdsg dmtrvtnwtc cttwcacttc gagacgaggt ctgaacgg 58

<210> 21
<211> 58
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence for VL CDR3

<220>
<221> misc_feature
<222> (27)..(27)
<223> n is a, c, g, or t

<400> 21
ggcttaggtc tctgcagbwt dmtrvtnwtc cttwcacttc gagacgaggt ctgaacgg 58

<210> 22
<211> 58
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence for VL CDR3

<400> 22
ggcttaggtc tctgcagdsg dmtrvtdsgc ctykgacttc gagacgaggt ctgaacgg 58

<210> 23
<211> 58
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence for VL CDR3

<400> 23
ggcttaggtc tctgcagbwt dmtrvtdsgc ctykgacttc gagacgaggt ctgaacgg 58

<210> 24
<211> 58
<212> DNA

<213> Artificial Sequence

<220>
<223> Artificial Sequence for VL CDR3

<220>
<221> misc_feature
<222> (27)..(27)
<223> n is a, c, g, or t

<400> 24
ggcttaggtc tctgcagdsg dmtrvtnwtc ctykgacttc gagacgaggt ctgaacgg 58

<210> 25
<211> 58
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial Sequence for VL CDR3

<220>
<221> misc_feature
<222> (27)..(27)
<223> n is a, c, g, or t

<400> 25
ggcttaggtc tctgcagbwt dmtrvtnwtc ctykgacttc gagacgaggt ctgaacgg 58

<210> 26
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> Clone 801 CDRH1

<400> 26

Asn Ser Tyr Ala Met His
1 5

<210> 27
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Clone 801 CDRH2

<400> 27

Ala Ile Ser Gly Asn Gly Gly Ser Thr Arg
1 5 10

<210> 28
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> Clone 801 CDRH3

<400> 28

Asp Arg Phe Arg Lys Val His Gly
1 5

<210> 29
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> Clone 801 CDRL3

<400> 29

Asp Ala Thr Phe Pro Leu
1 5

<210> 30
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> Clone 802 CDRH1

<400> 30

Gly Ser Tyr Ala Met His
1 5

<210> 31
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Clone 802 CDRH2

<400> 31

Ala Ile Ser Gly Ser Gly Gly Ser Thr Arg
1 5 10

<210> 32
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> Clone 802 CDRH3

<400> 32

Asp Leu Tyr Arg Arg Val His Gly
1 5

<210> 33
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> Clone 802 CDRL3

<400> 33

Asp Tyr Gly Phe Pro Leu
1 5

<210> 34
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> Clone 905 CDRH1

<400> 34

Ser Ser Tyr Leu Met His
1 5

<210> 35
<211> 10
<212> PRT

<213> Artificial Sequence

<220>
<223> Clone 905 CDRH2

<400> 35

Val Ile Ser Gly Ser Gly Gly Ser Thr Arg
1 5 10

<210> 36
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> Clone 905 CDRH3

<400> 36

Asp Leu Tyr Arg Arg Val Ala Gly
1 5

<210> 37
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> Clone 905 CDRL3

<400> 37

Asp Tyr Ala Leu Pro Leu
1 5

<210> 38
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> Clone 906 CDRH1

<400> 38

Ser Asn Tyr Ala Met His
1 5

<210> 39

<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Clone 906 CDRH2

<400> 39

Ala Ile Ser Gly Asn Gly Gly Ser Thr His
1 5 10

<210> 40
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> Clone 906 CDRH3

<400> 40

Asp Arg Phe Arg Arg Val Tyr Gly
1 5

<210> 41
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> Clone 906 CDRL3

<400> 41

Asp Tyr Thr Phe Pro Leu
1 5

<210> 42
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> Clone 817 CDRH1

<400> 42

Ser Ser Tyr Ala Met His
1 5

<210> 43
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Clone 817 CDRH2

<400> 43

Ala Ile Ser Gly Ser Gly Gly Ser Thr Arg
1 5 10

<210> 44
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> Clone 817 CDRH3

<400> 44

Gly Arg Trp Ser Lys Trp Gly
1 5

<210> 45
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> Clone 817 CDRL3

<400> 45

Thr Asp Asn Phe Pro Tyr
1 5