

VARIANT NUCLEIC ACID LIBRARIES FOR CRTH2**CROSS-REFERENCE**

[0001] This application claims the benefit of U.S. Provisional Patent Application No. 62/904,595 filed on September 23, 2019; U.S. Provisional Patent Application No. 62/935,590 filed on November 14, 2019; and U.S. Provisional Patent Application No. 62/945,752 filed on December 9, 2019, each of which is incorporated by reference in its entirety.

BACKGROUND

[0002] G protein-coupled receptors (GPCRs) such as Prostaglandin D2 receptor 2 (DP2 or CRTH2R) are implicated in a wide variety of diseases. Raising antibodies to GPCRs has been difficult due to problems in obtaining suitable antigen because GPCRs are often expressed at low levels in cells and are very unstable when purified. Thus, there is a need for improved agents for therapeutic intervention which target GPCRs.

INCORPORATION BY REFERENCE

[0003] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference.

BRIEF SUMMARY

[0004] Provided herein are nucleic acid libraries, comprising: a plurality of nucleic acids, wherein each of the nucleic acids encodes for a sequence that when translated encodes for a CRTH2R binding immunoglobulin, wherein the CRTH2R binding immunoglobulin comprises a variant of a CRTH2R binding domain, wherein the CRTH2R binding domain is a ligand for the CRTH2R, and wherein the nucleic acid library comprises at least 10,000 variant immunoglobulin heavy chains and at least 10,000 variant immunoglobulin light chains. Further provided herein are nucleic acid libraries, wherein the nucleic acid library comprises at least 50,000 variant immunoglobulin heavy chains and at least 50,000 variant immunoglobulin light chains. Further provided herein are nucleic acid libraries, wherein the nucleic acid library comprises at least 100,000 variant immunoglobulin heavy chains and at least 100,000 variant immunoglobulin light chains. Further provided herein are nucleic acid libraries, wherein the nucleic acid library comprises at least 10^5 non-identical nucleic acids. Further provided herein are nucleic acid libraries, wherein a length of the immunoglobulin heavy chain when translated is about 90 to about 100 amino acids. Further provided herein are nucleic acid libraries, wherein a length of the

immunoglobulin heavy chain when translated is about 100 to about 400 amino acids. Further provided herein are nucleic acid libraries, wherein the variant immunoglobulin heavy chain when translated comprises at least 80% sequence identity to any one of SEQ ID NO: 2338-2360 or 2403-2405. Further provided herein are nucleic acid libraries, wherein the variant immunoglobulin light chain when translated comprises at least 80% sequence identity to any one of SEQ ID NO: 2361-2381 or 2406-2408.

[0005] Provided herein are nucleic acid libraries comprising a plurality of nucleic acids, wherein each nucleic acid of the plurality of nucleic acids encodes for a sequence that when translated encodes for an antibody or antibody fragment thereof, wherein the antibody or antibody fragment thereof comprises a variable region of a heavy chain (VH) that comprises a CRTH2R binding domain, wherein each nucleic acid of the plurality of nucleic acids comprises a sequence encoding for a sequence variant of the CRTH2R binding domain, and wherein the antibody or antibody fragment binds to its antigen with a K_D of less than 100 nM. Further provided herein are nucleic acid libraries, wherein a length of the VH is about 90 to about 100 amino acids. Further provided herein are nucleic acid libraries, wherein a length of the VH is about 100 to about 400 amino acids. Further provided herein are nucleic acid libraries, wherein a length of the VH is about 270 to about 300 base pairs. Further provided herein are nucleic acid libraries, wherein a length of the VH is about 300 to about 1200 base pairs. Further provided herein are nucleic acid libraries, wherein the library comprises at least 10^5 non-identical nucleic acids.

[0006] Provided herein are nucleic acid libraries comprising: a plurality of nucleic acids, wherein each of the nucleic acids encodes for a sequence that when translated encodes for a CRTH2R single domain antibody, wherein each sequence of the plurality of sequences comprises a variant sequence encoding for a CDR1, CDR2, or CDR3 on a variable region of a heavy chain (VH); wherein the library comprises at least 30,000 variant sequences; and wherein the CRTH2R single domain antibody binds to its antigen with a K_D of less than 100 nM. Further provided herein are nucleic acid libraries, wherein a length of the VH when translated is about 90 to about 100 amino acids. Further provided herein are nucleic acid libraries, wherein a length of the VH when translated is about 100 to about 400 amino acids. Further provided herein are nucleic acid libraries, wherein a length of the VH is about 270 to about 300 base pairs. Further provided herein are nucleic acid libraries, wherein a length of the VH is about 300 to about 1200 base pairs. Further provided herein are nucleic acid libraries, wherein the VH when translated comprises at least 80% sequence identity to any one of SEQ ID NO: 2338-2360 or 2403-2405.

[0007] Provided herein are antibodies or antibody fragments that bind CRTH2R, comprising an immunoglobulin heavy chain and an immunoglobulin light chain: a. wherein the immunoglobulin

heavy chain comprises an amino acid sequence at least about 90% identical to that set forth in any one of SEQ ID NO: 2338-2360 or 2403-2405; and b. wherein the immunoglobulin light chain comprises an amino acid sequence at least about 90% identical to that set forth in any one of SEQ ID NO: 2361-2381 or 2406-2408. Further provided are antibodies or antibody fragments that bind CRTH2R, wherein the immunoglobulin heavy chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2338; and wherein the immunoglobulin light chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2361. Further provided are antibodies or antibody fragments that bind CRTH2R, wherein the immunoglobulin heavy chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2339; and wherein the immunoglobulin light chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2362. Further provided are antibodies or antibody fragments that bind CRTH2R, wherein the immunoglobulin heavy chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2340; and wherein the immunoglobulin light chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2363. Further provided are antibodies or antibody fragments that bind CRTH2R, wherein the immunoglobulin heavy chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2341; and wherein the immunoglobulin light chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2364. Further provided are antibodies or antibody fragments that bind CRTH2R, wherein the immunoglobulin heavy chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2342; and wherein the immunoglobulin light chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2365. Further provided are antibodies or antibody fragments that bind CRTH2R, wherein the immunoglobulin heavy chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2343; and wherein the immunoglobulin light chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2366. Further provided are antibodies or antibody fragments that bind CRTH2R, wherein the immunoglobulin heavy chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2344; and wherein the immunoglobulin light chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2367. Further provided are antibodies or antibody fragments that bind CRTH2R, wherein the immunoglobulin heavy chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2345; and wherein the immunoglobulin light chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2368. Further provided are antibodies or antibody fragments that bind

CRTH2R, wherein the immunoglobulin heavy chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2346; and wherein the immunoglobulin light chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2369. Further provided are antibodies or antibody fragments that bind CRTH2R, wherein the immunoglobulin heavy chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2347; and wherein the immunoglobulin light chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2370. Further provided are antibodies or antibody fragments that bind CRTH2R, wherein the immunoglobulin heavy chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2348; and wherein the immunoglobulin light chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2371. Further provided are antibodies or antibody fragments that bind CRTH2R, wherein the immunoglobulin heavy chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2349; and wherein the immunoglobulin light chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2372. Further provided are antibodies or antibody fragments that bind CRTH2R, wherein the immunoglobulin heavy chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2350; and wherein the immunoglobulin light chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2373. Further provided are antibodies or antibody fragments that bind CRTH2R, wherein the immunoglobulin heavy chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2351; and wherein the immunoglobulin light chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2374. Further provided are antibodies or antibody fragments that bind CRTH2R, wherein the immunoglobulin heavy chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2352; and wherein the immunoglobulin light chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2375. Further provided are antibodies or antibody fragments that bind CRTH2R, wherein the immunoglobulin heavy chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2353; and wherein the immunoglobulin light chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2376. Further provided are antibodies or antibody fragments that bind CRTH2R, wherein the immunoglobulin heavy chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2354; and wherein the immunoglobulin light chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2377. Further provided are antibodies or antibody fragments that bind CRTH2R, wherein the

immunoglobulin heavy chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2355; and wherein the immunoglobulin light chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2378. Further provided are antibodies or antibody fragments that bind CRTH2R, wherein the immunoglobulin heavy chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2356; and wherein the immunoglobulin light chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2379. Further provided are antibodies or antibody fragments that bind CRTH2R, wherein the immunoglobulin heavy chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2357; and wherein the immunoglobulin light chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2380. Further provided are antibodies or antibody fragments that bind CRTH2R, wherein the immunoglobulin heavy chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2358; and wherein the immunoglobulin light chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2381. Further provided are antibodies or antibody fragments that bind CRTH2R, wherein the immunoglobulin heavy chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2403; and wherein the immunoglobulin light chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2406. Further provided are antibodies or antibody fragments that bind CRTH2R, wherein the immunoglobulin heavy chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2404; and wherein the immunoglobulin light chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2407. Further provided are antibodies or antibody fragments that bind CRTH2R, wherein the immunoglobulin heavy chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2405; and wherein the immunoglobulin light chain comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NO: 2408. Further provided are antibodies or antibody fragments that bind CRTH2R, wherein the antibody is a monoclonal antibody, a polyclonal antibody, a bi-specific antibody, a multispecific antibody, a grafted antibody, a human antibody, a humanized antibody, a synthetic antibody, a chimeric antibody, a camelized antibody, a single-chain Fvs (scFv), a single chain antibody, a Fab fragment, a F(ab')₂ fragment, a Fd fragment, a Fv fragment, a single-domain antibody, an isolated complementarity determining region (CDR), a diabody, a fragment comprised of only a single monomeric variable domain, disulfide-linked Fvs (sdFv), an intrabody, an anti-idiotypic (anti-Id) antibody, or ab antigen-binding fragments thereof. Further provided are antibodies or antibody fragments that bind CRTH2R, wherein the antibody or antibody fragment

thereof is chimeric or humanized. Further provided are antibodies or antibody fragments that bind CRTH2R, wherein the antibody has an EC₅₀ less than about 25 nanomolar in a cAMP assay. Further provided are antibodies or antibody fragments that bind CRTH2R, wherein the antibody has an EC₅₀ less than about 20 nanomolar in a cAMP assay. Further provided are antibodies or antibody fragments that bind CRTH2R, wherein the antibody has an EC₅₀ less than about 10 nanomolar in a cAMP assay.

[0008] Provided herein are antibodies or antibody fragments, wherein the antibody or antibody fragment comprises a complementarity determining region (CDR) comprising an amino acid sequence at least about 90% identical to that set forth in any one of SEQ ID NOs: 2382-2402.

[0009] Provided herein are antibodies or antibody fragments, wherein the antibody or antibody fragment comprises a sequence of any one of SEQ ID NOs: 2382-2402 and wherein the antibody is a monoclonal antibody, a polyclonal antibody, a bi-specific antibody, a multispecific antibody, a grafted antibody, a human antibody, a humanized antibody, a synthetic antibody, a chimeric antibody, a camelized antibody, a single-chain Fvs (scFv), a single chain antibody, a Fab fragment, a F(ab')₂ fragment, a Fd fragment, a Fv fragment, a single-domain antibody, an isolated complementarity determining region (CDR), a diabody, a fragment comprised of only a single monomeric variable domain, disulfide-linked Fvs (sdFv), an intrabody, an anti-idiotypic (anti-Id) antibody, or ab antigen-binding fragments thereof.

[0010] Provided herein are methods of treating a disease or disorder of the central nervous system, kidney, intestine, lung, hair, skin, bone, or cartilage, comprising administering the antibody or antibody fragment described herein.

[0011] Provided herein are methods of treating a disease or disorder characterized by an inflammatory response, comprising administering the antibody or antibody fragment described herein.

[0012] Provided herein are methods of treating an allergic reaction, comprising administering the antibody or antibody fragment described herein. Further provided herein are methods, wherein the allergic reaction is chronic idiopathic urticaria or allergic rhinitis.

[0013] Provided herein are methods of treating asthma, comprising administering the antibody or antibody fragment described herein.

[0014] Provided herein are methods of treating alopecia or baldness, comprising administering the antibody or antibody fragment described herein.

[0015] Provided herein are methods for generating a nucleic acid library encoding for a CRTH2R antibody or antibody fragment thereof comprising: (a) providing predetermined sequences encoding for: i. a first plurality of polynucleotides, wherein each polynucleotide of the

first plurality of polynucleotides encodes for at least 1000 variant sequence encoding for CDR1 on a heavy chain; ii. a second plurality of polynucleotides, wherein each polynucleotide of the second plurality of polynucleotides encodes for at least 1000 variant sequence encoding for CDR2 on a heavy chain; iii. a third plurality of polynucleotides, wherein each polynucleotide of the third plurality of polynucleotides encodes for at least 1000 variant sequence encoding for CDR3 on a heavy chain; and (b) mixing the first plurality of polynucleotides, the second plurality of polynucleotides, and the third plurality of polynucleotides to form the nucleic acid library of variant nucleic acids encoding for the CRTH2R antibody or antibody fragment thereof, and wherein at least about 70% of the variant nucleic acids encode for an antibody or antibody fragment that binds to its antigen with a K_D of less than 100 nM. Further provided herein are methods for generating a nucleic acid library, wherein the CRTH2R antibody or antibody fragment thereof is a single domain antibody. Further provided herein are methods for generating a nucleic acid library, wherein the single domain antibody comprises one heavy chain variable domain. Further provided herein are methods for generating a nucleic acid library, wherein the single domain antibody is a VHH antibody. Further provided herein are methods for generating a nucleic acid library, wherein the nucleic acid library comprises at least 50,000 variant sequences. Further provided herein are methods for generating a nucleic acid library, wherein the nucleic acid library comprises at least 100,000 variant sequences. Further provided herein are methods for generating a nucleic acid library, wherein the nucleic acid library comprises at least 10^5 non-identical nucleic acids. Further provided herein are methods for generating a nucleic acid library, wherein the nucleic acid library comprises at least one sequence encoding for the CRTH2R antibody or antibody fragment that binds to CRTH2R with a K_D of less than 75 nM. Further provided herein are methods for generating a nucleic acid library, wherein the nucleic acid library comprises at least one sequence encoding for the CRTH2R antibody or antibody fragment that binds to CRTH2R with a K_D of less than 50 nM. Further provided herein are methods for generating a nucleic acid library, wherein the nucleic acid library comprises at least one sequence encoding for the CRTH2R antibody or antibody fragment that binds to CRTH2R with a K_D of less than 10 nM. Further provided herein are methods for generating a nucleic acid library, wherein the nucleic acid library comprises at least 500 variant sequences. Further provided herein are methods for generating a nucleic acid library, wherein the nucleic acid library comprises at least five sequences encoding for the CRTH2R antibody or antibody fragment that binds to CRTH2R with a K_D of less than 75 nM.

[0016] Provided herein are protein libraries encoded by the nucleic acid library described herein, wherein the protein library comprises peptides. Further provided herein are protein libraries, wherein the protein library comprises immunoglobulins. Further provided herein are

protein libraries, wherein the protein library comprises antibodies. Further provided herein are protein libraries, wherein the protein library is a peptidomimetic library.

[0017] Provided herein are vector libraries comprising the nucleic acid library described herein.

[0018] Provided herein are cell libraries comprising the nucleic acid library described herein.

[0019] Provided herein are cell libraries comprising the protein library described herein.

BRIEF DESCRIPTION OF THE DRAWINGS

[0020] **Figure 1** depicts a schematic of G protein-coupled receptor (GPCR) ligand interaction surfaces.

[0021] **Figure 2A** depicts a first schematic of an immunoglobulin scaffold.

[0022] **Figure 2B** depicts a second schematic of an immunoglobulin scaffold.

[0023] **Figure 3** depicts a schematic of a motif for placement in a scaffold.

[0024] **Figure 4** depicts a schematic of a GPCR.

[0025] **Figure 5** depicts schematics of segments for assembly of clonal fragments and non-clonal fragments.

[0026] **Figure 6** depicts schematics of segments for assembly of clonal fragments and non-clonal fragments.

[0027] **Figure 7** presents a diagram of steps demonstrating an exemplary process workflow for gene synthesis as disclosed herein.

[0028] **Figure 8** illustrates an example of a computer system.

[0029] **Figure 9** is a block diagram illustrating an architecture of a computer system.

[0030] **Figure 10** is a diagram demonstrating a network configured to incorporate a plurality of computer systems, a plurality of cell phones and personal data assistants, and Network Attached Storage (NAS).

[0031] **Figure 11** is a block diagram of a multiprocessor computer system using a shared virtual address memory space.

[0032]

[0033] **Figure 12A** depicts a schematic of an immunoglobulin scaffold comprising a VH domain attached to a VL domain using a linker.

[0034] **Figure 12B** depicts a schematic of a full-domain architecture of an immunoglobulin scaffold comprising a VH domain attached to a VL domain using a linker, a leader sequence, and pIII sequence.

[0035] **Figure 12C** depicts a schematic of four framework elements (FW1, FW2, FW3, FW4) and the variable 3 CDR (L1, L2, L3) elements for a VL or VH domain.

- [0036] **Figure 13** depicts a graph of yield of purified GPCR immunoglobulins.
- [0037] **Figure 14** depicts FACS data of GPCR binding protein expression.
- [0038] **Figures 15A-15C** depict cell-binding affinity of purified immunoglobulins.
- [0039] **Figure 15D** depicts cAMP activity of purified immunoglobulins.
- [0040] **Figure 16** depicts BVP ELISA data of purified immunoglobulins.
- [0041] **Figures 17A-17B** depict graphs of FACS analysis (**FIG. 17A**) and graphs of a dose curve and cAMP activity (**FIG. 17B**) of CRTH2-41-51.
- [0042] **Figure 18** depicts a graph of a dose curve of CRTH2-44-59.
- [0043] **Figure 19** depicts a graph FACS analysis of CRTH2-44-59.
- [0044] **Figures 20A-20E** depict FACS analysis plots of cell binding as measured by mean fluorescence intensity (MFI) vs. 8-point titrations with CRTH2R IgG using CRTH2-74, CRTH2-24, CRTH2-28, CRTH2-39, CRTH2-19, CRTH2-9, CRTH2-8, CRTH2-27, CRTH2-45, CRTH2-35, CRTH2-50, CRTH2-66, CRTH2-57, CRTH2-32, CRTH2-15, CRTH2-25, CRTH2-42, CRTH2-55, CRTH2-60, and CRTH2-70.
- [0045] **Figure 21A** depicts an example gated dot plot showing CRTH2-27 binding at 100 nM.
- [0046] **Figure 21B** depicts an example APC histogram showing CRTH2-27 binding at 100 nM.
- [0047] **Figure 22A** depicts binding analysis as in previous figures using comparator antibody gPCR-51.
- [0048] **Figure 22B** depicts binding analysis as in previous figures using comparator antibody gPCR-52.
- [0049] **Figures 23A-23B** depict IgG binding curves with CRTH2-9, CRTH2-27, CRTH2-50, CRTH2-32, and CRTH2-42, which have functional effects in cAMP assays.
- [0050] **Figure 24A** depicts results of CRTH2R cAMP assays across all antibodies tested at 300, 100, and 33 nM.
- [0051] **Figure 24B** depicts results of CRTH2R cAMP assays across all antibodies tested at 33 nM.
- [0052] **Figure 25** indicates the negative allosteric effect seen in five of the CRTH2R IgG (CRTH2-9, CRTH2-27, CRTH2-50, CRTH2-32, and CRTH2-42).
- [0053] **Figures 26A-26C** depict control experiments of allosteric modulators, showing comparator antibody 52 is a positive allosteric modulator.
- [0054] **Figures 27A-27D** depict activity of CRTH2R in β -arrestin recruitment assays.
- [0055] **Figure 28** depicts a schema of design of phage-displayed hyperimmune libraries generated herein.

[0056] **Figures 29A-29F** depict graphs of binding affinity for the CRTH2R immunoglobulins CRTH2-48-03 (**FIG. 29A**), CRTH2-48-21 (**FIG. 29B**), and CRTH2-48-27 (**FIG. 29C**) and cAMP assays for CRTH2-48-03 (**FIG. 29D**), CRTH2-48-21 (**FIG. 29E**), and CRTH2-48-27 (**FIG. 29F**).

[0057] **Figure 30A** depicts a schema of heavy chain IGHV3-23 design.

[0058] **Figure 30B** depicts a schema of heavy chain IGHV1-69 design.

[0059] **Figure 30C** depicts a schema of light chains IGKV 2-28 and IGLV 1-51 design.

[0060] **Figure 30D** depicts a schema of the theoretical diversity and final diversity of a GPCR library.

DETAILED DESCRIPTION

[0061] The present disclosure employs, unless otherwise indicated, conventional molecular biology techniques, which are within the skill of the art. Unless defined otherwise, all technical and scientific terms used herein have the same meaning as is commonly understood by one of ordinary skill in the art.

[0062] Definitions

[0063] Throughout this disclosure, various embodiments are presented in a range format. It should be understood that the description in range format is merely for convenience and brevity and should not be construed as an inflexible limitation on the scope of any embodiments. Accordingly, the description of a range should be considered to have specifically disclosed all the possible subranges as well as individual numerical values within that range to the tenth of the unit of the lower limit unless the context clearly dictates otherwise. For example, description of a range such as from 1 to 6 should be considered to have specifically disclosed subranges such as from 1 to 3, from 1 to 4, from 1 to 5, from 2 to 4, from 2 to 6, from 3 to 6 etc., as well as individual values within that range, for example, 1.1, 2, 2.3, 5, and 5.9. This applies regardless of the breadth of the range. The upper and lower limits of these intervening ranges may independently be included in the smaller ranges, and are also encompassed within the disclosure, subject to any specifically excluded limit in the stated range. Where the stated range includes one or both of the limits, ranges excluding either or both of those included limits are also included in the disclosure, unless the context clearly dictates otherwise.

[0064] The terminology used herein is for the purpose of describing particular embodiments only and is not intended to be limiting of any embodiment. As used herein, the singular forms “a,” “an” and “the” are intended to include the plural forms as well, unless the context clearly indicates otherwise. It will be further understood that the terms “comprises” and/or “comprising,” when used in this specification, specify the presence of stated features, integers, steps, operations, elements, and/or components, but do not preclude the presence or addition of one or more other features,

integers, steps, operations, elements, components, and/or groups thereof. As used herein, the term “and/or” includes any and all combinations of one or more of the associated listed items.

[0065] Unless specifically stated or obvious from context, as used herein, the term “about” in reference to a number or range of numbers is understood to mean the stated number and numbers +/- 10% thereof, or 10% below the lower listed limit and 10% above the higher listed limit for the values listed for a range.

[0066] Unless specifically stated, as used herein, the term “nucleic acid” encompasses double- or triple-stranded nucleic acids, as well as single-stranded molecules. In double- or triple-stranded nucleic acids, the nucleic acid strands need not be coextensive (i.e., a double-stranded nucleic acid need not be double-stranded along the entire length of both strands). Nucleic acid sequences, when provided, are listed in the 5' to 3' direction, unless stated otherwise. Methods described herein provide for the generation of isolated nucleic acids. Methods described herein additionally provide for the generation of isolated and purified nucleic acids. A “nucleic acid” as referred to herein can comprise at least 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, or more bases in length. Moreover, provided herein are methods for the synthesis of any number of polypeptide-segments encoding nucleotide sequences, including sequences encoding non-ribosomal peptides (NRPs), sequences encoding non-ribosomal peptide-synthetase (NRPS) modules and synthetic variants, polypeptide segments of other modular proteins, such as antibodies, polypeptide segments from other protein families, including non-coding DNA or RNA, such as regulatory sequences e.g. promoters, transcription factors, enhancers, siRNA, shRNA, RNAi, miRNA, small nucleolar RNA derived from microRNA, or any functional or structural DNA or RNA unit of interest. The following are non-limiting examples of polynucleotides: coding or non-coding regions of a gene or gene fragment, intergenic DNA, loci (locus) defined from linkage analysis, exons, introns, messenger RNA (mRNA), transfer RNA, ribosomal RNA, short interfering RNA (siRNA), short-hairpin RNA (shRNA), micro-RNA (miRNA), small nucleolar RNA, ribozymes, complementary DNA (cDNA), which is a DNA representation of mRNA, usually obtained by reverse transcription of messenger RNA (mRNA) or by amplification; DNA molecules produced synthetically or by amplification, genomic DNA, recombinant polynucleotides, branched polynucleotides, plasmids, vectors, isolated DNA of any sequence, isolated RNA of any sequence, nucleic acid probes, and primers. cDNA encoding for a gene or gene fragment referred herein may comprise at least one region encoding for exon sequences without an intervening intron sequence in the genomic equivalent sequence.

[0067] As used herein, the term “percent (%) sequence identity” with respect to a sequence is defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in the specific sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as EMBOSS MATCHER, EMBOSS WATER, EMBOSS STRETCHER, EMBOSS NEEDLE, EMBOSS LALIGN, BLAST, BLAST-2, ALIGN or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared.

[0068] CRTH2R Libraries

[0069] Provided herein are methods and compositions relating to G protein-coupled receptor (GPCR) binding libraries for Prostaglandin D2 receptor 2 (DP2 or CRTH2R) comprising nucleic acids encoding for a scaffold comprising a CRTH2R binding domain. Scaffolds as described herein can stably support a CRTH2R binding domain. The CRTH2R binding domain may be designed based on surface interactions of a CRTH2R ligand and CRTH2R. Libraries as described herein may be further variegated to provide for variant libraries comprising nucleic acids each encoding for a predetermined variant of at least one predetermined reference nucleic acid sequence. Further described herein are protein libraries that may be generated when the nucleic acid libraries are translated. In some instances, nucleic acid libraries as described herein are transferred into cells to generate a cell library. Also provided herein are downstream applications for the libraries synthesized using methods described herein. Downstream applications include identification of variant nucleic acids or protein sequences with enhanced biologically relevant functions, e.g., improved stability, affinity, binding, functional activity, and for the treatment or prevention of a disease state associated with CRTH2R signaling.

[0070] Methods, compositions, and systems described herein for the optimization of CRTH2R immunoglobulins or antibodies comprise a ratio-variant approach that mirror the natural diversity of antibody sequences. In some instances, libraries of optimized CRTH2R immunoglobulins or antibodies comprise variant CRTH2R immunoglobulin or antibody sequences. In some instances, the variant CRTH2R immunoglobulin or antibody sequences are designed comprising variant CDR regions. In some instances, the variant CRTH2R immunoglobulin or antibody sequences comprising variant CDR regions are generated by shuffling the natural CDR sequences in a llama, humanized, or chimeric framework. In some instances, such libraries are synthesized, cloned into

expression vectors, and translation products (antibodies) evaluated for activity. In some instances, fragments of sequences are synthesized and subsequently assembled. In some instances, expression vectors are used to display and enrich desired antibodies, such as phage display. In some instances, the phage vector is a Fab phagemid vector. Selection pressures used during enrichment in some instances includes binding affinity, toxicity, immunological tolerance, stability, or other factor. Such expression vectors allow antibodies with specific properties to be selected ("panning"), and subsequent propagation or amplification of such sequences enriches the library with these sequences. Panning rounds can be repeated any number of times, such as 1, 2, 3, 4, 5, 6, 7, or more than 7 rounds. In some instances, each round of panning involves a number of washes. In some instances, each round of panning involves at least or about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or more than 16 washes.

[0071] Described herein are methods and systems of in-silico library design. Libraries as described herein, in some instances, are designed based on a database comprising a variety of antibody sequences. In some instances, the database comprises a plurality of variant antibody sequences against various targets. In some instances, the database comprises at least 100, 500, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, or more than 5000 antibody sequences. An exemplary database is an iCAN database. In some instances, the database comprises naïve and memory B-cell receptor sequences. In some instances, the naïve and memory B-cell receptor sequences are human, mouse, or primate sequences. In some instances, the naïve and memory B-cell receptor sequences are human sequences. In some instances, the database is analyzed for position specific variation. In some instances, antibodies described herein comprise position specific variations in CDR regions. In some instances, the CDR regions comprise multiple sites for variation.

[0072] *Scaffold Libraries*

[0073] Provided herein are libraries comprising nucleic acids encoding for a scaffold, wherein sequences for CRTH2R binding domains are placed in the scaffold. Scaffold described herein allow for improved stability for a range of CRTH2R binding domain encoding sequences when inserted into the scaffold, as compared to an unmodified scaffold. Exemplary scaffolds include, but are not limited to, a protein, a peptide, an immunoglobulin, derivatives thereof, or combinations thereof. In some instances, the scaffold is an immunoglobulin. Scaffolds as described herein comprise improved functional activity, structural stability, expression, specificity, or a combination thereof. In some instances, scaffolds comprise long regions for supporting a CRTH2R binding domain.

[0074] Provided herein are libraries comprising nucleic acids encoding for a scaffold, wherein the scaffold is an immunoglobulin. In some instances, the immunoglobulin is an antibody. As used herein, the term antibody will be understood to include proteins having the characteristic two-armed, Y-shape of a typical antibody molecule as well as one or more fragments of an antibody that retain the ability to specifically bind to an antigen. Exemplary antibodies include, but are not limited to, a monoclonal antibody, a polyclonal antibody, a bi-specific antibody, a multispecific antibody, a grafted antibody, a human antibody, a humanized antibody, a synthetic antibody, a chimeric antibody, a camelized antibody, a single-chain Fvs (scFv) (including fragments in which the VL and VH are joined using recombinant methods by a synthetic or natural linker that enables them to be made as a single protein chain in which the VL and VH regions pair to form monovalent molecules, including single chain Fab and scFab), a single chain antibody, a Fab fragment (including monovalent fragments comprising the VL, VH, CL, and CH1 domains), a F(ab')₂ fragment (including bivalent fragments comprising two Fab fragments linked by a disulfide bridge at the hinge region), a Fd fragment (including fragments comprising the VH and CH1 fragment), a Fv fragment (including fragments comprising the VL and VH domains of a single arm of an antibody), a single-domain antibody (dAb or sdAb) (including fragments comprising a VH domain), an isolated complementarity determining region (CDR), a diabody (including fragments comprising bivalent dimers such as two VL and VH domains bound to each other and recognizing two different antigens), a fragment comprised of only a single monomeric variable domain, disulfide-linked Fvs (sdFv), an intrabody, an anti-idiotypic (anti-Id) antibody, or ab antigen-binding fragments thereof. In some instances, the libraries disclosed herein comprise nucleic acids encoding for a scaffold, wherein the scaffold is a Fv antibody, including Fv antibodies comprised of the minimum antibody fragment which contains a complete antigen-recognition and antigen-binding site. In some embodiments, the Fv antibody consists of a dimer of one heavy chain and one light chain variable domain in tight, non-covalent association, and the three hypervariable regions of each variable domain interact to define an antigen-binding site on the surface of the VH-VL dimer. In some embodiments, the six hypervariable regions confer antigen-binding specificity to the antibody. In some embodiments, a single variable domain (or half of an Fv comprising only three hypervariable regions specific for an antigen, including single domain antibodies isolated from camelid animals comprising one heavy chain variable domain or variable region of a heavy chain such as VHH antibodies or nanobodies) has the ability to recognize and bind antigen. In some instances, the libraries disclosed herein comprise nucleic acids encoding for a scaffold, wherein the scaffold is a single-chain Fv or scFv, including antibody fragments comprising a VH, a VL, or both a VH and VL domain, wherein both domains are present in a single polypeptide chain.

In some embodiments, the Fv polypeptide further comprises a polypeptide linker between the VH and VL domains allowing the scFv to form the desired structure for antigen binding. In some instances, a scFv is linked to the Fc fragment or a VHH is linked to the Fc fragment (including minibodies). In some instances, the antibody comprises immunoglobulin molecules and immunologically active fragments of immunoglobulin molecules, e.g., molecules that contain an antigen binding site. Immunoglobulin molecules are of any type (e.g., IgG, IgE, IgM, IgD, IgA and IgY), class (e.g., IgG 1, IgG 2, IgG 3, IgG 4, IgA 1 and IgA 2) or subclass.

[0075] In some embodiments, libraries comprise immunoglobulins that are adapted to the species of an intended therapeutic target. Generally, these methods include “mammalization” and comprises methods for transferring donor antigen-binding information to a less immunogenic mammal antibody acceptor to generate useful therapeutic treatments. In some instances, the mammal is mouse, rat, equine, sheep, cow, primate (e.g., chimpanzee, baboon, gorilla, orangutan, monkey), dog, cat, pig, donkey, rabbit, and human. In some instances, provided herein are libraries and methods for felinization and caninization of antibodies.

[0076] “Humanized” forms of non-human antibodies can be chimeric antibodies that contain minimal sequence derived from the non-human antibody. A humanized antibody is generally a human antibody (recipient antibody) in which residues from one or more CDRs are replaced by residues from one or more CDRs of a non-human antibody (donor antibody). The donor antibody can be any suitable non-human antibody, such as a mouse, rat, rabbit, chicken, or non-human primate antibody having a desired specificity, affinity, or biological effect. In some instances, selected framework region residues of the recipient antibody are replaced by the corresponding framework region residues from the donor antibody. Humanized antibodies may also comprise residues that are not found in either the recipient antibody or the donor antibody. In some instances, these modifications are made to further refine antibody performance.

[0077] “Caninization” can comprise a method for transferring non-canine antigen-binding information from a donor antibody to a less immunogenic canine antibody acceptor to generate treatments useful as therapeutics in dogs. In some instances, caninized forms of non-canine antibodies provided herein are chimeric antibodies that contain minimal sequence derived from non-canine antibodies. In some instances, caninized antibodies are canine antibody sequences (“acceptor” or “recipient” antibody) in which hypervariable region residues of the recipient are replaced by hypervariable region residues from a non-canine species (“donor” antibody) such as mouse, rat, rabbit, cat, dogs, goat, chicken, bovine, horse, llama, camel, dromedaries, sharks, non-human primates, human, humanized, recombinant sequence, or an engineered sequence having the desired properties. In some instances, framework region (FR) residues of the canine antibody are

replaced by corresponding non-canine FR residues. In some instances, caninized antibodies include residues that are not found in the recipient antibody or in the donor antibody. In some instances, these modifications are made to further refine antibody performance. The caninized antibody may also comprise at least a portion of an immunoglobulin constant region (Fc) of a canine antibody.

[0078] “Felinization” can comprise a method for transferring non-feline antigen-binding information from a donor antibody to a less immunogenic feline antibody acceptor to generate treatments useful as therapeutics in cats. In some instances, felinized forms of non-feline antibodies provided herein are chimeric antibodies that contain minimal sequence derived from non-feline antibodies. In some instances, felinized antibodies are feline antibody sequences (“acceptor” or “recipient” antibody) in which hypervariable region residues of the recipient are replaced by hypervariable region residues from a non-feline species (“donor” antibody) such as mouse, rat, rabbit, cat, dogs, goat, chicken, bovine, horse, llama, camel, dromedaries, sharks, non-human primates, human, humanized, recombinant sequence, or an engineered sequence having the desired properties. In some instances, framework region (FR) residues of the feline antibody are replaced by corresponding non-feline FR residues. In some instances, felinized antibodies include residues that are not found in the recipient antibody or in the donor antibody. In some instances, these modifications are made to further refine antibody performance. The felinized antibody may also comprise at least a portion of an immunoglobulin constant region (Fc) of a felinized antibody.

[0079] Provided herein are libraries comprising nucleic acids encoding for a scaffold, wherein the scaffold is a non-immunoglobulin. In some instances, the scaffold is a non-immunoglobulin binding domain. For example, the scaffold is an antibody mimetic. Exemplary antibody mimetics include, but are not limited to, anticalins, affilins, affibody molecules, affimers, affitins, alphabodies, avimers, atrimers, DARPs, fynomers, Kunitz domain-based proteins, monobodies, anticalins, knottins, armadillo repeat protein-based proteins, and bicyclic peptides.

[0080] Libraries described herein comprising nucleic acids encoding for a scaffold, wherein the scaffold is an immunoglobulin, comprise variations in at least one region of the immunoglobulin. Exemplary regions of the antibody for variation include, but are not limited to, a complementarity-determining region (CDR), a variable domain, or a constant domain. In some instances, the CDR is CDR1, CDR2, or CDR3. In some instances, the CDR is a heavy domain including, but not limited to, CDRH1, CDRH2, and CDRH3. In some instances, the CDR is a light domain including, but not limited to, CDRL1, CDRL2, and CDRL3. In some instances, the variable domain is variable domain, light chain (VL) or variable domain, heavy chain (VH). In some instances, the CDR1, CDR2, or CDR3 is of a variable domain, light chain (VL). CDR1, CDR2, or CDR3 of a variable domain, light chain (VL) can be referred to as CDRL1, CDRL2, or CDRL3, respectively. CDR1,

CDR2, or CDR3 of a variable domain, heavy chain (VH) can be referred to as CDRH1, CDRH2, or CDRH3, respectively. In some instances, the VL domain comprises kappa or lambda chains. In some instances, the constant domain is constant domain, light chain (CL) or constant domain, heavy chain (CH).

[0081] Methods described herein provide for synthesis of libraries comprising nucleic acids encoding for a scaffold, wherein each nucleic acid encodes for a predetermined variant of at least one predetermined reference nucleic acid sequence. In some cases, the predetermined reference sequence is a nucleic acid sequence encoding for a protein, and the variant library comprises sequences encoding for variation of at least a single codon such that a plurality of different variants of a single residue in the subsequent protein encoded by the synthesized nucleic acid are generated by standard translation processes. In some instances, the scaffold library comprises varied nucleic acids collectively encoding variations at multiple positions. In some instances, the variant library comprises sequences encoding for variation of at least a single codon of a CDRH1, CDRH2, CDRH3, CDRL1, CDRL2, CDRL3, VL, or VH domain. In some instances, the variant library comprises sequences encoding for variation of multiple codons of a CDRH1, CDRH2, CDRH3, CDRL1, CDRL2, CDRL3, VL, or VH domain. In some instances, the variant library comprises sequences encoding for variation of multiple codons of framework element 1 (FW1), framework element 2 (FW2), framework element 3 (FW3), or framework element 4 (FW4). Exemplary number of codons for variation include, but are not limited to, at least or about 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 225, 250, 275, 300, or more than 300 codons.

[0082] In some instances, the at least one region of the immunoglobulin for variation is from heavy chain V-gene family, heavy chain D-gene family, heavy chain J-gene family, light chain V-gene family, or light chain J-gene family. In some instances, the light chain V-gene family comprises immunoglobulin kappa (IGK) gene or immunoglobulin lambda (IGL). Exemplary genes include, but are not limited to, IGHV1-18, IGHV1-69, IGHV1-8, IGHV3-21, IGHV3-23, IGHV3-30/33m, IGHV3-28, IGHV1-69, IGHV3-74, IGHV4-39, IGHV4-59/61, IGKV1-39, IGKV1-9, IGKV2-28, IGKV3-11, IGKV3-15, IGKV3-20, IGKV4-1, IGLV1-51, IGLV2-14, IGLV1-40, and IGLV3-1. In some instances, the gene is IGHV1-69, IGHV3-30, IGHV3-23, IGHV3, IGHV1-46, IGHV3-7, IGHV1, or IGHV1-8. In some instances, the gene is IGHV1-69 and IGHV3-30. In some instances, the gene is IGHJ3, IGHJ6, IGHJ, IGHJ4, IGHJ5, IGHJ2, or IGH1. In some instances, the gene is IGHJ3, IGHJ6, IGHJ, or IGHJ4.

[0083] Provided herein are libraries comprising nucleic acids encoding for immunoglobulin scaffolds, wherein the libraries are synthesized with various numbers of fragments. In some

instances, the fragments comprise the CDRH1, CDRH2, CDRH3, CDRL1, CDRL2, CDRL3, VL, or VH domain. In some instances, the fragments comprise framework element 1 (FW1), framework element 2 (FW2), framework element 3 (FW3), or framework element 4 (FW4). In some instances, the scaffold libraries are synthesized with at least or about 2 fragments, 3 fragments, 4 fragments, 5 fragments, or more than 5 fragments. The length of each of the nucleic acid fragments or average length of the nucleic acids synthesized may be at least or about 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, or more than 600 base pairs. In some instances, the length is about 50 to 600, 75 to 575, 100 to 550, 125 to 525, 150 to 500, 175 to 475, 200 to 450, 225 to 425, 250 to 400, 275 to 375, or 300 to 350 base pairs.

[0084] Libraries comprising nucleic acids encoding for immunoglobulin scaffolds as described herein comprise various lengths of amino acids when translated. In some instances, the length of each of the amino acid fragments or average length of the amino acid synthesized may be at least or 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, or more than 150 amino acids. In some instances, the length of the amino acid is about 15 to 150, 20 to 145, 25 to 140, 30 to 135, 35 to 130, 40 to 125, 45 to 120, 50 to 115, 55 to 110, 60 to 110, 65 to 105, 70 to 100, or 75 to 95 amino acids. In some instances, the length of the amino acid is about 22 amino acids to about 75 amino acids. In some instances, the immunoglobulin scaffolds comprise at least or about 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, or more than 5000 amino acids.

[0085] A number of variant sequences for the at least one region of the immunoglobulin for variation are de novo synthesized using methods as described herein. In some instances, a number of variant sequences is de novo synthesized for CDRH1, CDRH2, CDRH3, CDRL1, CDRL2, CDRL3, VL, VH, or combinations thereof. In some instances, a number of variant sequences is de novo synthesized for framework element 1 (FW1), framework element 2 (FW2), framework element 3 (FW3), or framework element 4 (FW4). The number of variant sequences may be at least or about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, or more than 500 sequences. In some instances, the number of variant sequences is at least or about 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, or more than 8000 sequences. In some instances, the number of variant sequences is about 10 to 500, 25 to 475, 50 to 450, 75 to 425, 100 to 400, 125 to 375, 150 to 350, 175 to 325, 200 to 300, 225 to 375, 250 to 350, or 275 to 325 sequences.

[0086] Variant sequences for the at least one region of the immunoglobulin, in some instances, vary in length or sequence. In some instances, the at least one region that is de novo synthesized is

for CDRH1, CDRH2, CDRH3, CDRL1, CDRL2, CDRL3, VL, VH, or combinations thereof. In some instances, the at least one region that is de novo synthesized is for framework element 1 (FW1), framework element 2 (FW2), framework element 3 (FW3), or framework element 4 (FW4). In some instances, the variant sequence comprises at least or about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, or more than 50 variant nucleotides or amino acids as compared to wild-type. In some instances, the variant sequence comprises at least or about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, or 50 additional nucleotides or amino acids as compared to wild-type. In some instances, the variant sequence comprises at least or about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, or 50 less nucleotides or amino acids as compared to wild-type. In some instances, the libraries comprise at least or about 10^1 , 10^2 , 10^3 , 10^4 , 10^5 , 10^6 , 10^7 , 10^8 , 10^9 , 10^{10} , or more than 10^{10} variants.

[0087] Following synthesis of scaffold libraries, scaffold libraries may be used for screening and analysis. For example, scaffold libraries are assayed for library displayability and panning. In some instances, displayability is assayed using a selectable tag. Exemplary tags include, but are not limited to, a radioactive label, a fluorescent label, an enzyme, a chemiluminescent tag, a colorimetric tag, an affinity tag or other labels or tags that are known in the art. In some instances, the tag is histidine, polyhistidine, myc, hemagglutinin (HA), or FLAG. In some instances, scaffold libraries are assayed by sequencing using various methods including, but not limited to, single-molecule real-time (SMRT) sequencing, Polony sequencing, sequencing by ligation, reversible terminator sequencing, proton detection sequencing, ion semiconductor sequencing, nanopore sequencing, electronic sequencing, pyrosequencing, Maxam-Gilbert sequencing, chain termination (e.g., Sanger) sequencing, +S sequencing, or sequencing by synthesis.

[0088] In some instances, the scaffold libraries are assayed for functional activity, structural stability (e.g., thermal stable or pH stable), expression, specificity, or a combination thereof. In some instances, the scaffold libraries are assayed for scaffolds capable of folding. In some instances, a region of the antibody is assayed for functional activity, structural stability, expression, specificity, folding, or a combination thereof. For example, a VH region or VL region is assayed for functional activity, structural stability, expression, specificity, folding, or a combination thereof.

[0089] *CRTH2R Libraries*

[0090] Provided herein are CRTH2R binding libraries comprising nucleic acids encoding for scaffolds comprising sequences for CRTH2R binding domains. In some instances, the scaffolds are immunoglobulins. In some instances, the scaffolds comprising sequences for CRTH2R binding domains are determined by interactions between the CRTH2R binding domains and the CRTH2R.

[0091] Provided herein are libraries comprising nucleic acids encoding scaffolds comprising CRTH2R binding domains, wherein the CRTH2R binding domains are designed based on surface interactions on CRTH2R. In some instances, the CRTH2R binding domain comprises a sequence as defined by SEQ ID NO: 1. In some instances, the CRTH2R binding domains interact with the amino- or carboxy-terminus of the CRTH2R. In some instances, the CRTH2R binding domains interact with at least one transmembrane domain including, but not limited to, transmembrane domain 1 (TM1), transmembrane domain 2 (TM2), transmembrane domain 3 (TM3), transmembrane domain 4 (TM4), transmembrane domain 5 (TM5), transmembrane domain 6 (TM6), and transmembrane domain 7 (TM7). In some instances, the CRTH2R binding domains interact with an intracellular surface of the CRTH2R. For example, the CRTH2R binding domains interact with at least one intracellular loop including, but not limited to, intracellular loop 1 (ICL1), intracellular loop 2 (ICL2), and intracellular loop 3 (ICL3). In some instances, the CRTH2R binding domains interact with an extracellular surface of the CRTH2R. For example, the CRTH2R binding domains interact with at least one extracellular domain (ECD) or extracellular loop (ECL) of the CRTH2R. The extracellular loops include, but are not limited to, extracellular loop 1 (ECL1), extracellular loop 2 (ECL2), and extracellular loop 3 (ECL3).

[0092] Described herein are CRTH2R binding domains, wherein the CRTH2R binding domains are designed based on surface interactions between a CRTH2R ligand and the CRTH2R. In some instances, the ligand is a peptide. In some instances, the ligand is a CRTH2R agonist. In some instances, the ligand is a CRTH2R antagonist. In some instances, the ligand is a CRTH2R allosteric modulator. In some instances, the allosteric modulator is a negative allosteric modulator. In some instances, the allosteric modulator is a positive allosteric modulator.

[0093] Sequences of CRTH2R binding domains based on surface interactions between a CRTH2R ligand and the CRTH2R are analyzed using various methods. For example, multispecies computational analysis is performed. In some instances, a structure analysis is performed. In some instances, a sequence analysis is performed. Sequence analysis can be performed using a database known in the art. Non-limiting examples of databases include, but are not limited to, NCBI BLAST (blast.ncbi.nlm.nih.gov/Blast.cgi), UCSC Genome Browser (genome.ucsc.edu/), UniProt (www.uniprot.org/), and IUPHAR/BPS Guide to PHARMACOLOGY (guidetopharmacology.org/).

[0094] Described herein are CRTH2R binding domains designed based on sequence analysis among various organisms. For example, sequence analysis is performed to identify homologous sequences in different organisms. Exemplary organisms include, but are not limited to, mouse, rat, equine, sheep, cow, primate (*e.g.*, chimpanzee, baboon, gorilla, orangutan, monkey), dog, cat, pig, donkey, rabbit, fish, fly, and human.

[0095] Following identification of CRTH2R binding domains, libraries comprising nucleic acids encoding for the CRTH2R binding domains may be generated. In some instances, libraries of CRTH2R binding domains comprise sequences of CRTH2R binding domains designed based on conformational ligand interactions, peptide ligand interactions, small molecule ligand interactions, extracellular domains of CRTH2R, or antibodies that target CRTH2R. In some instances, libraries of CRTH2R binding domains comprise sequences of CRTH2R binding domains designed based on peptide ligand interactions. Libraries of CRTH2R binding domains may be translated to generate protein libraries. In some instances, libraries of CRTH2R binding domains are translated to generate peptide libraries, immunoglobulin libraries, derivatives thereof, or combinations thereof. In some instances, libraries of CRTH2R binding domains are translated to generate protein libraries that are further modified to generate peptidomimetic libraries. In some instances, libraries of CRTH2R binding domains are translated to generate protein libraries that are used to generate small molecules.

[0096] Methods described herein provide for synthesis of libraries of CRTH2R binding domains comprising nucleic acids each encoding for a predetermined variant of at least one predetermined reference nucleic acid sequence. In some cases, the predetermined reference sequence is a nucleic acid sequence encoding for a protein, and the variant library comprises sequences encoding for variation of at least a single codon such that a plurality of different variants of a single residue in the subsequent protein encoded by the synthesized nucleic acid are generated by standard translation processes. In some instances, the libraries of CRTH2R binding domains comprise varied nucleic acids collectively encoding variations at multiple positions. In some instances, the variant library comprises sequences encoding for variation of at least a single codon in a CRTH2R binding domain. In some instances, the variant library comprises sequences encoding for variation of multiple codons in a CRTH2R binding domain. An exemplary number of codons for variation include, but are not limited to, at least or about 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 225, 250, 275, 300, or more than 300 codons.

[0097] Methods described herein provide for synthesis of libraries comprising nucleic acids encoding for the CRTH2R binding domains, wherein the libraries comprise sequences encoding for variation of length of the CRTH2R binding domains. In some instances, the library comprises sequences encoding for variation of length of at least or about 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 225, 250, 275, 300, or more than 300 codons less as compared to a predetermined reference sequence. In some instances, the library comprises sequences encoding for variation of length of at least or about 1, 5, 10, 15, 20, 25, 30,

35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 225, 250, 275, 300, or more than 300 codons more as compared to a predetermined reference sequence.

[0098] Following identification of CRTH2R binding domains, the CRTH2R binding domains may be placed in scaffolds as described herein. In some instances, the scaffolds are immunoglobulins. In some instances, the CRTH2R binding domains are placed in the CDRH3 region. CRTH2R binding domains that may be placed in scaffolds can also be referred to as a motif. Scaffolds comprising CRTH2R binding domains may be designed based on binding, specificity, stability, expression, folding, or downstream activity. In some instances, the scaffolds comprising CRTH2R binding domains enable contact with the CRTH2R. In some instances, the scaffolds comprising CRTH2R binding domains enables high affinity binding with the CRTH2R. An exemplary amino acid sequence of CRTH2R binding domain is described in **Table 1A**.

Table 1A. CRTH2R binding domain amino acid sequences

SEQ ID NO	GPCR	Amino Acid Sequence
1	CRTH2R	MSANATLKPLCPILEQMSRLQSHSNTSIRYIDHAAVLLHGLASLLGLV ENGVLFFVVGCRMQRQTVVTTWVLHLALSDLLASASLPFFTYFLAVGH SWELGTTFCKLHSSIFFLNMFASGFLLSAISLDRCLQVVRPVWAQNHR TVAAAHKVCLVLWALAVLNTVPYFVFRDTISRLDGRIMCYYNVLLL NPGPDRDATCNSRQAALAVSKFLLAFLVPLAIIASSHAAVSLRLQHRG RRRPGRFVRLVAADVAAAFALCWGPYHVFSLLEARAHANPGLRPLVW RGLPFVTSLAFFNSVANPVLVLTCPDMLRKLRRSLRTVLESVLVDDS ELGGAGSSRRRRTSSTARSASPLALCSRPEEPRGPALLGWLLGSCAA SPQTGPLNRALSSTSS

[0099] Provided herein are scaffolds or immunoglobulins comprising CRTH2R binding domains, wherein the sequences of the CRTH2R binding domains support interaction with CRTH2R. The sequence may be homologous or identical to a sequence of a CRTH2R ligand. In some instances, the CRTH2R binding domain sequence comprises at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 1. In some instances, the CRTH2R binding domain sequence comprises at least or about 95% homology to SEQ ID NO: 1. In some instances, the CRTH2R binding domain sequence comprises at least or about 97% homology to SEQ ID NO: 1. In some instances, the CRTH2R binding domain sequence comprises at least or about 99% homology to SEQ ID NO: 1. In some instances, the CRTH2R binding domain sequence comprises at least or about 100% homology to SEQ ID NO: 1. In some instances, the CRTH2R binding domain sequence comprises at least a portion having at least or about 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160,

170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, or more than 400 amino acids of SEQ ID NO: 1.

[00100] In some embodiments, the CRTH2R antibody or immunoglobulin sequence comprises complementarity determining regions (CDRs) comprising a sequence as set forth in **Table 1B** and **Table 14B**. In some embodiments, the CRTH2R antibody or immunoglobulin sequence comprises complementarity determining regions (CDRs) comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to any one of SEQ ID NOs: 2409-2420 or 2382-2402. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises complementarity determining regions (CDRs) comprising at least or about 95% homology to any one of SEQ ID NOs: 2409-2420 or 2382-2402. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises complementarity determining regions (CDRs) comprising at least or about 97% homology to any one of SEQ ID NOs: 2409-2420 or 2382-2402. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises complementarity determining regions (CDRs) comprising at least or about 99% homology to any one of SEQ ID NOs: 2409-2420 or 2382-2402. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises complementarity determining regions (CDRs) comprising at least or about 100% homology to any one of SEQ ID NOs: 2409-2420 or 2382-2402. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises complementarity determining regions (CDRs) comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of any one of SEQ ID NOs: 2409-2420 or 2382-2402.

Table 1B.

Construct Description	Amino Acid Sequence	SEQ ID NO
IGHV1-69 CDR1	GGTFSSYA	2409
IGHV1-69 CDR2	IPIFGTA	2410
IGHV1-69 CDR3	CARNNNNNNNNNFDYW	2411
IGHV3-23 CDR1	GFTFSSYA	2412
IGHV3-23 CDR2	ISGSGGST	2413
IGHV3-23 CDR3	CAKNNNNNNNNNNFDYW	2414
IGKV1-39 CDR1	QSISSY	2415
IGKV1-39 CDR2	AAS	2416
IGKV1-39 CDR3	CQSYSTPNTF	2417
IGKV3-20 CDR1	QSVSSSY	2418
IGKV3-20 CDR2	GAS	2419
IGKV3-20 CDR3	CQQYGSSPNTF	2420

[00101] In some embodiments, the CRTH2R antibody or immunoglobulin sequence comprises a CDR1 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to any one of SEQ ID NOs: 2409, 2412, 2415, or 2418. In

some instances, the CRTH2R antibody or immunoglobulin sequence comprises CDR1 comprising at least or about 95% homology to any one of SEQ ID NO: 2409, 2412, 2415, or 2418. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises CDR1 comprising at least or about 97% homology to any one of SEQ ID NO: 2409, 2412, 2415, or 2418. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises CDR1 comprising at least or about 99% homology to any one of SEQ ID NO: 2409, 2412, 2415, or 2418. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises CDR1 comprising 100% homology to any one of SEQ ID NO: 2409, 2412, 2415, or 2418. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises CDR1 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of any one of SEQ ID NO: 2409, 2412, 2415, or 2418.

[00102] In some embodiments, the CRTH2R antibody or immunoglobulin sequence comprises a CDR2 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to any one of SEQ ID NOs: 2410, 2413, 2416, or 2419. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises CDR2 comprising at least or about 95% homology to any one of SEQ ID NO: 2410, 2413, 2416, or 2419. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises CDR2 comprising at least or about 97% homology to any one of SEQ ID NO: 2410, 2413, 2416, or 2419. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises CDR2 comprising at least or about 99% homology to any one of SEQ ID NO: 2410, 2413, 2416, or 2419. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises CDR2 comprising at 100% homology to any one of SEQ ID NO: 2410, 2413, 2416, or 2419. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises CDR2 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of any one of SEQ ID NO: 2410, 2413, 2416, or 2419.

[00103] In some embodiments, the CRTH2R antibody or immunoglobulin sequence comprises a CDR3 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to any one of SEQ ID NOs: 2411, 2414, 2417, 2420, or 2382-2402. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises CDR3 comprising at least or about 95% homology to any one of SEQ ID NO: 2411, 2414, 2417, 2420, or 2382-2402. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises CDR3 comprising at least or about 97% homology to any one of SEQ ID NO: 2411, 2414, 2417, 2420, or 2382-2402. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises CDR3 comprising at least or about 99% homology to any one of SEQ ID NO:

2411, 2414, 2417, 2420, or 2382-2402. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises CDR3 comprising 100% homology to any one of SEQ ID NO: 2411, 2414, 2417, 2420, or 2382-2402. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises CDR3 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of any one of SEQ ID NO: 2411, 2414, 2417, 2420, or 2382-2402.

[00104] In some embodiments, the CRTH2R antibody or immunoglobulin sequence comprises a CDRH1 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2409; a CDRH2 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2410; and a CDRH3 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2411. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises CDRH1 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2409; a CDRH2 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2410; and a CDRH3 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2411. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises CDRH1 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2409; a CDRH2 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2410; and a CDRH3 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2411.

[00105] In some embodiments, the CRTH2R antibody or immunoglobulin sequence comprises a CDRH1 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2412; a CDRH2 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2413; and a CDRH3 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2414. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises CDRH1 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2412; a CDRH2 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2413; and a CDRH3 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2414. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises CDRH1 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than

16 amino acids of SEQ ID NO: 2412; a CDRH2 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2413; and a CDRH3 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2414.

[00106] In some embodiments, the CRTH2R antibody or immunoglobulin sequence comprises a CDRL1 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2415; a CDRL2 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2416; and a CDRL3 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2417. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises CDRL1 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2415; a CDRL2 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2416; and a CDRL3 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2417. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises CDRL1 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2415; a CDRL2 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2416; and a CDRL3 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2417.

[00107] In some embodiments, the CRTH2R antibody or immunoglobulin sequence comprises a CDRL1 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2418; a CDRL2 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2419; and a CDRL3 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2420. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises CDRL1 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2418; a CDRL2 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2419; and a CDRL3 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2420. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises CDRL1 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2418; a CDRL2 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2419; and a

CDRL3 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2420.

[00108] In some embodiments, the CRTH2R antibody or immunoglobulin sequence comprises a CDRH1 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2409; a CDRH2 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2410; a CDRH3 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2411; a CDRL1 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2415; a CDRL2 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2416; and a CDRL3 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2417. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises CDRH1 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2409; a CDRH2 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2410; a CDRH3 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2411; a CDRL1 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2415; a CDRL2 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2416; and a CDRL3 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2417. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises CDRH1 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2409; a CDRH2 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2410; a CDRH3 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2411; a CDRL1 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2415; a CDRL2 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2416; and a CDRL3 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2417.

[00109] In some embodiments, the CRTH2R antibody or immunoglobulin sequence comprises a CDRH1 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2409; a CDRH2 comprising at least or

about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2410; a CDRH3 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2411, a CDRL1 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2418; a CDRL2 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2419; and a CDRL3 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2420. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises CDRH1 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2409; a CDRH2 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2410; a CDRH3 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2411; a CDRL1 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2418; a CDRL2 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2419; and a CDRL3 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2420. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises CDRH1 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2409; a CDRH2 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2410; a CDRH3 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2411; a CDRL1 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2418; a CDRL2 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2419; and a CDRL3 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2420.

[00110] In some embodiments, the CRTH2R antibody or immunoglobulin sequence comprises a CDRH1 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2412; a CDRH2 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2413; a CDRH3 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2414, a CDRL1 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2415; a CDRL2

comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2416; and a CDRL3 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2417. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises CDRH1 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2412; a CDRH2 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2413; a CDRH3 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2414; a CDRL1 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2415; a CDRL2 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2416; and a CDRL3 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2417. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises CDRH1 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2412; a CDRH2 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2413; a CDRH3 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2414; a CDRL1 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2415; a CDRL2 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2416; and a CDRL3 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2417.

[00111] In some embodiments, the CRTH2R antibody or immunoglobulin sequence comprises a CDRH1 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2412; a CDRH2 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2413; a CDRH3 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2414; a CDRL1 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2418; a CDRL2 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2419; and a CDRL3 comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to SEQ ID NO: 2420. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises CDRH1 comprising at least or about 95%, 97%, 99%, or 100% homology to

SEQ ID NO: 2412; a CDRH2 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2413; a CDRH3 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2414; a CDRL1 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2418; a CDRL2 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2419; and a CDRL3 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 2420. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises CDRH1 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2412; a CDRH2 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2413; a CDRH3 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2414; a CDRL1 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2418; a CDRL2 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2419; and a CDRL3 comprising at least a portion having at least or about 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, or more than 16 amino acids of SEQ ID NO: 2420.

[00112] Described herein, in some embodiments, are antibodies or immunoglobulins that bind to the CRTH2R. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises a heavy chain variable domain comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to any one of SEQ ID NO: 2338-2360 or 2403-2405. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises a heavy chain variable domain comprising at least or about 95% sequence identity to any one of SEQ ID NO: 2338-2360 or 2403-2405. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises a heavy chain variable domain comprising at least or about 97% sequence identity to any one of SEQ ID NO: 2338-2360 or 2403-2405. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises a heavy chain variable domain comprising at least or about 99% sequence identity to any one of SEQ ID NO: 2338-2360 or 2403-2405. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises a heavy chain variable domain comprising at least or about 100% sequence identity to any one of SEQ ID NO: 2338-2360 or 2403-2405. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises a heavy chain variable domain comprising at least a portion having at least or about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330,

340, 350, 360, 370, 380, 390, 400, or more than 400 amino acids of any one of SEQ ID NO: 2338-2360 and 2403-2405.

[00113] In some instances, the CRTH2R antibody or immunoglobulin sequence comprises a light chain variable domain comprising at least or about 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% sequence identity to any one of SEQ ID NO: 2361-2381 or 2406-2408. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises a light chain variable domain comprising at least or about 95% sequence identity to any one of SEQ ID NO: 2361-2381 and 2406-2408. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises a light chain variable domain comprising at least or about 97% sequence identity to any one of SEQ ID NO: 2361-2381 or 2406-2408. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises a light chain variable domain comprising at least or about 99% sequence identity to any one of SEQ ID NO: 2361-2381 or 2406-2408. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises a light chain variable domain comprising at least or about 100% sequence identity to any one of SEQ ID NO: 2361-2381 or 2406-2408. In some instances, the CRTH2R antibody or immunoglobulin sequence comprises a light chain variable domain comprising at least a portion having at least or about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, or more than 400 amino acids of any one of SEQ ID NO: 2361-2381 or 2406-2408.

[00114] Provided herein are CRTH2R binding libraries comprising nucleic acids encoding for scaffolds or immunoglobulins comprising CRTH2R binding domains comprise variation in domain type, domain length, or residue variation. In some instances, the domain is a region in the scaffold comprising the CRTH2R binding domains. For example, the region is the VH, CDRH3, or VL domain. In some instances, the domain is the CRTH2R binding domain.

[00115] Methods described herein provide for synthesis of a CRTH2R binding library of nucleic acids each encoding for a predetermined variant of at least one predetermined reference nucleic acid sequence. In some cases, the predetermined reference sequence is a nucleic acid sequence encoding for a protein, and the variant library comprises sequences encoding for variation of at least a single codon such that a plurality of different variants of a single residue in the subsequent protein encoded by the synthesized nucleic acid are generated by standard translation processes. In some instances, the CRTH2R binding library comprises varied nucleic acids collectively encoding variations at multiple positions. In some instances, the variant library comprises sequences encoding for variation of at least a single codon of a VH, CDRH3, or VL domain. In some instances, the variant library comprises sequences encoding for variation of at least a single codon

in a CRTH2R binding domain. For example, at least one single codon of a CRTH2R binding domain as listed in **Table 1A** is varied. In some instances, the variant library comprises sequences encoding for variation of multiple codons of a VH, CDRH3, or VL domain. In some instances, the variant library comprises sequences encoding for variation of multiple codons in a CRTH2R binding domain. An exemplary number of codons for variation include, but are not limited to, at least or about 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 225, 250, 275, 300, or more than 300 codons.

[00116] Methods described herein provide for synthesis of a CRTH2R binding library of nucleic acids each encoding for a predetermined variant of at least one predetermined reference nucleic acid sequence, wherein the CRTH2R binding library comprises sequences encoding for variation of length of a domain. In some instances, the domain is VH, CDRH3, or VL domain. In some instances, the domain is the CRTH2R binding domain. In some instances, the library comprises sequences encoding for variation of length of at least or about 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 225, 250, 275, 300, or more than 300 codons less as compared to a predetermined reference sequence. In some instances, the library comprises sequences encoding for variation of length of at least or about 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 225, 250, 275, 300, or more than 300 codons more as compared to a predetermined reference sequence.

[00117] Provided herein are CRTH2R binding libraries comprising nucleic acids encoding for scaffolds comprising CRTH2R binding domains, wherein the CRTH2R binding libraries are synthesized with various numbers of fragments. In some instances, the fragments comprise the VH, CDRH3, or VL domain. In some instances, the CRTH2R binding libraries are synthesized with at least or about 2 fragments, 3 fragments, 4 fragments, 5 fragments, or more than 5 fragments. The length of each of the nucleic acid fragments or average length of the nucleic acids synthesized may be at least or about 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, or more than 600 base pairs. In some instances, the length is about 50 to 600, 75 to 575, 100 to 550, 125 to 525, 150 to 500, 175 to 475, 200 to 450, 225 to 425, 250 to 400, 275 to 375, or 300 to 350 base pairs.

[00118] CRTH2R binding libraries comprising nucleic acids encoding for scaffolds comprising CRTH2R binding domains as described herein comprise various lengths of amino acids when translated. In some instances, the length of each of the amino acid fragments or average length of the amino acid synthesized may be at least or 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, or more than 150 amino acids. In some instances, the length of the amino acid is about 15 to 150, 20 to 145, 25 to 140, 30

to 135, 35 to 130, 40 to 125, 45 to 120, 50 to 115, 55 to 110, 60 to 110, 65 to 105, 70 to 100, or 75 to 95 amino acids. In some instances, the length of the amino acid is about 22 to about 75 amino acids.

[00119] CRTH2R binding libraries comprising de novo synthesized variant sequences encoding for scaffolds comprising CRTH2R binding domains comprise a number of variant sequences. In some instances, a number of variant sequences is de novo synthesized for a CDRH1, CDRH2, CDRH3, CDRL1, CDRL2, CDRL3, VL, VH, or a combination thereof. In some instances, a number of variant sequences is de novo synthesized for framework element 1 (FW1), framework element 2 (FW2), framework element 3 (FW3), or framework element 4 (FW4). In some instances, a number of variant sequences is de novo synthesized for a CRTH2R binding domain. For example, the number of variant sequences is about 1 to about 10 sequences for the VH domain, about 10^8 sequences for the CRTH2R binding domain, and about 1 to about 44 sequences for the VK domain. See **FIGS. 2A-2B**. The number of variant sequences may be at least or about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, or more than 500 sequences. In some instances, the number of variant sequences is about 10 to 300, 25 to 275, 50 to 250, 75 to 225, 100 to 200, or 125 to 150 sequences.

[00120] CRTH2R binding libraries comprising de novo synthesized variant sequences encoding for scaffolds comprising CRTH2R binding domains comprise improved diversity. For example, variants are generated by placing CRTH2R binding domain variants in immunoglobulin scaffold variants comprising N-terminal CDRH3 variations and C-terminal CDRH3 variations. In some instances, variants include affinity maturation variants. Alternatively or in combination, variants include variants in other regions of the immunoglobulin including, but not limited to, CDRH1, CDRH2, CDRL1, CDRL2, and CDRL3. In some instances, the number of variants of the CRTH2R binding libraries is least or about 10^4 , 10^5 , 10^6 , 10^7 , 10^8 , 10^9 , 10^{10} , 10^{11} , 10^{12} , 10^{13} , 10^{14} , 10^{15} , 10^{16} , 10^{17} , 10^{18} , 10^{19} , 10^{20} , or more than 10^{20} non-identical sequences non-identical sequences. For example, a library comprising about 10 variant sequences for a VH region, about 237 variant sequences for a CDRH3 region, and about 43 variant sequences for a VL and CDRL3 region comprises 10^5 non-identical sequences ($10 \times 237 \times 43$).

[00121] Provided herein are libraries comprising nucleic acids encoding for a CRTH2R antibody comprising variation in at least one region of the antibody, wherein the region is the CDR region. In some instances, the CRTH2R antibody is a single domain antibody comprising one heavy chain variable domain such as a VHH antibody. In some instances, the VHH antibody comprises variation in one or more CDR regions. In some instances, libraries described herein comprise at

least or about 1, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1200, 1400, 1600, 1800, 2000, 2400, 2600, 2800, 3000, or more than 3000 sequences of a CDR1, CDR2, or CDR3. In some instances, libraries described herein comprise at least or about 10^4 , 10^5 , 10^6 , 10^7 , 10^8 , 10^9 , 10^{10} , 10^{11} , 10^{12} , 10^{13} , 10^{14} , 10^{15} , 10^{16} , 10^{17} , 10^{18} , 10^{19} , 10^{20} , or more than 10^{20} sequences of a CDR1, CDR2, or CDR3. For example, the libraries comprise at least 2000 sequences of a CDR1, at least 1200 sequences for CDR2, and at least 1600 sequences for CDR3. In some instances, each sequence is non-identical.

[00122] In some instances, the CDR1, CDR2, or CDR3 is of a variable domain, light chain (VL). CDR1, CDR2, or CDR3 of a variable domain, light chain (VL) can be referred to as CDRL1, CDRL2, or CDRL3, respectively. In some instances, libraries described herein comprise at least or about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1200, 1400, 1600, 1800, 2000, 2400, 2600, 2800, 3000, or more than 3000 sequences of a CDR1, CDR2, or CDR3 of the VL. In some instances, libraries described herein comprise at least or about 10^4 , 10^5 , 10^6 , 10^7 , 10^8 , 10^9 , 10^{10} , 10^{11} , 10^{12} , 10^{13} , 10^{14} , 10^{15} , 10^{16} , 10^{17} , 10^{18} , 10^{19} , 10^{20} , or more than 10^{20} sequences of a CDR1, CDR2, or CDR3 of the VL. For example, the libraries comprise at least 20 sequences of a CDR1 of the VL, at least 4 sequences of a CDR2 of the VL, and at least 140 sequences of a CDR3 of the VL. In some instances, the libraries comprise at least 2 sequences of a CDR1 of the VL, at least 1 sequence of CDR2 of the VL, and at least 3000 sequences of a CDR3 of the VL. In some instances, the VL is IGKV1-39, IGKV1-9, IGKV2-28, IGKV3-11, IGKV3-15, IGKV3-20, IGKV4-1, IGLV1-51, IGLV2-14, IGLV1-40, or IGLV3-1. In some instances, the VL is IGKV2-28. In some instances, the VL is IGLV1-51.

[00123] In some instances, the CDR1, CDR2, or CDR3 is of a variable domain, heavy chain (VH). CDR1, CDR2, or CDR3 of a variable domain, heavy chain (VH) can be referred to as CDRH1, CDRH2, or CDRH3, respectively. In some instances, libraries described herein comprise at least or about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1200, 1400, 1600, 1800, 2000, 2400, 2600, 2800, 3000, or more than 3000 sequences of a CDR1, CDR2, or CDR3 of the VH. In some instances, libraries described herein comprise at least or about 10^4 , 10^5 , 10^6 , 10^7 , 10^8 , 10^9 , 10^{10} , 10^{11} , 10^{12} , 10^{13} , 10^{14} , 10^{15} , 10^{16} , 10^{17} , 10^{18} , 10^{19} , 10^{20} , or more than 10^{20} sequences of a CDR1, CDR2, or CDR3 of the VH. For example, the libraries comprise at least 30 sequences of a CDR1 of the VH, at least 570 sequences of a CDR2 of the VH, and at least 10^8 sequences of a CDR3 of the VH. In some instances, the 20 or IGHV4-59/61. In some instances, the VH is IGHV1-69, IGHV3-30, IGHV3-23, IGHV3, IGHV1-46, IGHV3-7, IGHV1, or IGHV1-8. In some instances, the VH is IGHV1-69 and IGHV3-30. In some instances, the VH is IGHV3-23.

[00124] Libraries as described herein, in some embodiments, comprise varying lengths of a CDRL1, CDRL2, CDRL3, CDRH1, CDRH2, or CDRH3. In some instances, the length of the CDRL1, CDRL2, CDRL3, CDRH1, CDRH2, or CDRH3 comprises at least or about 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, 40, 50, 60, 70, 80, 90, or more than 90 amino acids in length. For example, the CDRH3 comprises at least or about 12, 15, 16, 17, 20, 21, or 23 amino acids in length. In some instances, the CDRL1, CDRL2, CDRL3, CDRH1, CDRH2, or CDRH3 comprises a range of about 1 to about 10, about 5 to about 15, about 10 to about 20, or about 15 to about 30 amino acids in length.

[00125] Libraries comprising nucleic acids encoding for antibodies having variant CDR sequences as described herein comprise various lengths of amino acids when translated. In some instances, the length of each of the amino acid fragments or average length of the amino acid synthesized may be at least or 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, or more than 150 amino acids. In some instances, the length of the amino acid is about 15 to 150, 20 to 145, 25 to 140, 30 to 135, 35 to 130, 40 to 125, 45 to 120, 50 to 115, 55 to 110, 60 to 110, 65 to 105, 70 to 100, or 75 to 95 amino acids. In some instances, the length of the amino acid is about 22 amino acids to about 75 amino acids. In some instances, the antibodies comprise at least or about 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, or more than 5000 amino acids.

[00126] Ratios of the lengths of a CDRL1, CDRL2, CDRL3, CDRH1, CDRH2, or CDRH3 may vary in libraries described herein. In some instances, a CDRL1, CDRL2, CDRL3, CDRH1, CDRH2, or CDRH3 comprising at least or about 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, 40, 50, 60, 70, 80, 90, or more than 90 amino acids in length comprises about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or more than 90% of the library. For example, a CDRH3 comprising about 23 amino acids in length is present in the library at 40%, a CDRH3 comprising about 21 amino acids in length is present in the library at 30%, a CDRH3 comprising about 17 amino acids in length is present in the library at 20%, and a CDRH3 comprising about 12 amino acids in length is present in the library at 10%. In some instances, a CDRH3 comprising about 20 amino acids in length is present in the library at 40%, a CDRH3 comprising about 16 amino acids in length is present in the library at 30%, a CDRH3 comprising about 15 amino acids in length is present in the library at 20%, and a CDRH3 comprising about 12 amino acids in length is present in the library at 10%.

[00127] Libraries as described herein encoding for a VHH antibody comprise variant CDR sequences that are shuffled to generate a library with a theoretical diversity of at least or about 10^7 , 10^8 , 10^9 , 10^{10} , 10^{11} , 10^{12} , 10^{13} , 10^{14} , 10^{15} , 10^{16} , 10^{17} , 10^{18} , 10^{19} , 10^{20} , or more than 10^{20} sequences.

In some instances, the library has a final library diversity of at least or about 10^7 , 10^8 , 10^9 , 10^{10} , 10^{11} , 10^{12} , 10^{13} , 10^{14} , 10^{15} , 10^{16} , 10^{17} , 10^{18} , 10^{19} , 10^{20} , or more than 10^{20} sequences.

[00128] Provided herein are CRTH2R binding libraries encoding for an immunoglobulin. In some instances, the CRTH2R immunoglobulin is an antibody. In some instances, the CRTH2R immunoglobulin is a VHH antibody. In some instances, the CRTH2R immunoglobulin comprises a binding affinity (e.g., K_D) to CRTH2R of less than 1 nM, less than 1.2 nM, less than 2 nM, less than 5 nM, less than 10 nM, less than 11 nM, less than 13.5 nM, less than 15 nM, less than 20 nM, less than 25 nM, or less than 30 nM. In some instances, the CRTH2R immunoglobulin comprises a K_D of less than 1 nM. In some instances, the CRTH2R immunoglobulin comprises a K_D of less than 1.2 nM. In some instances, the CRTH2R immunoglobulin comprises a K_D of less than 2 nM. In some instances, the CRTH2R immunoglobulin comprises a K_D of less than 5 nM. In some instances, the CRTH2R immunoglobulin comprises a K_D of less than 10 nM. In some instances, the CRTH2R immunoglobulin comprises a K_D of less than 13.5 nM. In some instances, the CRTH2R immunoglobulin comprises a K_D of less than 15 nM. In some instances, the CRTH2R immunoglobulin comprises a K_D of less than 20 nM. In some instances, the CRTH2R immunoglobulin comprises a K_D of less than 25 nM. In some instances, the CRTH2R immunoglobulin comprises a K_D of less than 30 nM.

[00129] In some instances, the CRTH2R immunoglobulin is a CRTH2R agonist. In some instances, the CRTH2R immunoglobulin is a CRTH2R antagonist. In some instances, the CRTH2R immunoglobulin is a CRTH2R allosteric modulator. In some instances, the allosteric modulator is a negative allosteric modulator. In some instances, the allosteric modulator is a positive allosteric modulator. In some instances, the CRTH2R immunoglobulin results in agonistic, antagonistic, or allosteric effects at a concentration of at least or about 1 nM, 2 nM, 4 nM, 6 nM, 8 nM, 10 nM, 20 nM, 30 nM, 40 nM, 50 nM, 60 nM, 70 nM, 80 nM, 90 nM, 100 nM, 120 nM, 140 nM, 160 nM, 180 nM, 200 nM, 300 nM, 400 nM, 500 nM, 600 nM, 700 nM, 800 nM, 900 nM, 1000 nM, or more than 1000 nM. In some instances, the CRTH2R immunoglobulin is a negative allosteric modulator. In some instances, the CRTH2R immunoglobulin is a negative allosteric modulator at a concentration of at least or about 0.001, 0.005, 0.01, 0.05, 0.1, 0.5, 1 nM, 2 nM, 4 nM, 6 nM, 8 nM, 10 nM, 20 nM, 30 nM, 40 nM, 50 nM, 60 nM, 70 nM, 80 nM, 90 nM, 100 nM, or more than 100 nM. In some instances, the CRTH2R immunoglobulin is a negative allosteric modulator at a concentration in a range of about 0.001 to about 100, 0.01 to about 90, about 0.1 to about 80, 1 to about 50, about 10 to about 40 nM, or about 1 to about 10 nM. In some instances, the CRTH2R immunoglobulin comprises an EC_{50} or IC_{50} of at least or about 0.001, 0.0025, 0.005, 0.01, 0.025, 0.05, 0.06, 0.07, 0.08, 0.9, 0.1, 0.5, 1, 2, 3, 4, 5, 6, or more than 6 nM.

In some instances, the CRTH2R immunoglobulin comprises an EC₅₀ or IC₅₀ of at least or about 1 nM, 2 nM, 4 nM, 6 nM, 8 nM, 10 nM, 20 nM, 30 nM, 40 nM, 50 nM, 60 nM, 70 nM, 80 nM, 90 nM, 100 nM, or more than 100 nM.

[00130] CRTH2R immunoglobulins as described herein may comprise improved properties. In some instances, the CRTH2R immunoglobulins are monomeric. In some instances, the CRTH2R immunoglobulins are not prone to aggregation. In some instances, at least or about 70%, 75%, 80%, 85%, 90%, 95%, or 99% of the CRTH2R immunoglobulins are monomeric. In some instances, the CRTH2R immunoglobulins are thermostable. In some instances, the CRTH2R immunoglobulins result in reduced non-specific binding.

[00131] Following synthesis of CRTH2R binding libraries comprising nucleic acids encoding scaffolds comprising CRTH2R binding domains, libraries may be used for screening and analysis. For example, libraries are assayed for library displayability and panning. In some instances, displayability is assayed using a selectable tag. Exemplary tags include, but are not limited to, a radioactive label, a fluorescent label, an enzyme, a chemiluminescent tag, a colorimetric tag, an affinity tag or other labels or tags that are known in the art. In some instances, the tag is histidine, polyhistidine, myc, hemagglutinin (HA), or FLAG. For example as seen in **FIG. 3**, the CRTH2R binding libraries comprises nucleic acids encoding scaffolds comprising CRTH2R binding domains with multiple tags such as GFP, FLAG, and Lucy as well as a DNA barcode. In some instances, libraries are assayed by sequencing using various methods including, but not limited to, single-molecule real-time (SMRT) sequencing, Polony sequencing, sequencing by ligation, reversible terminator sequencing, proton detection sequencing, ion semiconductor sequencing, nanopore sequencing, electronic sequencing, pyrosequencing, Maxam-Gilbert sequencing, chain termination (e.g., Sanger) sequencing, +S sequencing, or sequencing by synthesis.

[00132] *Expression Systems*

[00133] Provided herein are libraries comprising nucleic acids encoding for scaffolds comprising CRTH2R binding domains, wherein the libraries have improved specificity, stability, expression, folding, or downstream activity. In some instances, libraries described herein are used for screening and analysis.

[00134] Provided herein are libraries comprising nucleic acids encoding for scaffolds comprising CRTH2R binding domains, wherein the nucleic acid libraries are used for screening and analysis. In some instances, screening and analysis comprises *in vitro*, *in vivo*, or *ex vivo* assays. Cells for screening include primary cells taken from living subjects or cell lines. Cells may be from prokaryotes (e.g., bacteria and fungi) or eukaryotes (e.g., animals and plants). Exemplary animal cells include, without limitation, those from a mouse, rabbit, primate, and insect. In some

instances, cells for screening include a cell line including, but not limited to, Chinese Hamster Ovary (CHO) cell line, human embryonic kidney (HEK) cell line, or baby hamster kidney (BHK) cell line. In some instances, nucleic acid libraries described herein may also be delivered to a multicellular organism. Exemplary multicellular organisms include, without limitation, a plant, a mouse, rabbit, primate, and insect.

[00135] Nucleic acid libraries or protein libraries encoded thereof described herein may be screened for various pharmacological or pharmacokinetic properties. In some instances, the libraries are screened using *in vitro* assays, *in vivo* assays, or *ex vivo* assays. For example, *in vitro* pharmacological or pharmacokinetic properties that are screened include, but are not limited to, binding affinity, binding specificity, and binding avidity. Exemplary *in vivo* pharmacological or pharmacokinetic properties of libraries described herein that are screened include, but are not limited to, therapeutic efficacy, activity, preclinical toxicity properties, clinical efficacy properties, clinical toxicity properties, immunogenicity, potency, and clinical safety properties.

[00136] Pharmacological or pharmacokinetic properties that may be screened include, but are not limited to, cell binding affinity and cell activity. For example, cell binding affinity assays or cell activity assays are performed to determine agonistic, antagonistic, or allosteric effects of libraries described herein. In some instances, the cell activity assay is a cAMP assay. In some instances, libraries as described herein are compared to cell binding or cell activity of ligands of CRTH2R.

[00137] Libraries as described herein may be screened in cell-based assays or in non-cell-based assays. Examples of non-cell-based assays include, but are not limited to, using viral particles, using *in vitro* translation proteins, and using proteoliposomes with CRTH2R.

[00138] Nucleic acid libraries as described herein may be screened by sequencing. In some instances, next generation sequence is used to determine sequence enrichment of CRTH2R binding variants. In some instances, V gene distribution, J gene distribution, V gene family, CDR3 counts per length, or a combination thereof is determined. In some instances, clonal frequency, clonal accumulation, lineage accumulation, or a combination thereof is determined. In some instances, number of sequences, sequences with VH clones, clones, clones greater than 1, clonotypes, clonotypes greater than 1, lineages, simpsons, or a combination thereof is determined. In some instances, a percentage of non-identical CDR3s is determined. For example, the percentage of non-identical CDR3s is calculated as the number of non-identical CDR3s in a sample divided by the total number of sequences that had a CDR3 in the sample.

[00139] Provided herein are nucleic acid libraries, wherein the nucleic acid libraries may be expressed in a vector. Expression vectors for inserting nucleic acid libraries disclosed herein may

comprise eukaryotic or prokaryotic expression vectors. Exemplary expression vectors include, without limitation, mammalian expression vectors: pSF-CMV-NEO-NH2-PPT-3XFLAG, pSF-CMV-NEO-COOH-3XFLAG, pSF-CMV-PURO-NH2-GST-TEV, pSF-OXB20-COOH-TEV-FLAG(R)-6His, pCEP4 pDEST27, pSF-CMV-Ub-KrYFP, pSF-CMV-FMDV-daGFP, pEF1a-mCherry-N1 Vector, pEF1a-tdTomato Vector, pSF-CMV-FMDV-Hygro, pSF-CMV-PGK-Puro, pMCP-tag(m), and pSF-CMV-PURO-NH2-CMYC; bacterial expression vectors: pSF-OXB20-BetaGal, pSF-OXB20-Fluc, pSF-OXB20, and pSF-Tac; plant expression vectors: pRI 101-AN DNA and pCambia2301; and yeast expression vectors: pTYB21 and pKLAC2, and insect vectors: pAc5.1/V5-His A and pDEST8. In some instances, the vector is pcDNA3 or pcDNA3.1.

[00140] Described herein are nucleic acid libraries that are expressed in a vector to generate a construct comprising a scaffold comprising sequences of CRTH2R binding domains. In some instances, a size of the construct varies. In some instances, the construct comprises at least or about 500, 600, 700, 800, 900, 1000, 1100, 1300, 1400, 1500, 1600, 1700, 1800, 2000, 2400, 2600, 2800, 3000, 3200, 3400, 3600, 3800, 4000, 4200, 4400, 4600, 4800, 5000, 6000, 7000, 8000, 9000, 10000, or more than 10000 bases. In some instances, a the construct comprises a range of about 300 to 1,000, 300 to 2,000, 300 to 3,000, 300 to 4,000, 300 to 5,000, 300 to 6,000, 300 to 7,000, 300 to 8,000, 300 to 9,000, 300 to 10,000, 1,000 to 2,000, 1,000 to 3,000, 1,000 to 4,000, 1,000 to 5,000, 1,000 to 6,000, 1,000 to 7,000, 1,000 to 8,000, 1,000 to 9,000, 1,000 to 10,000, 2,000 to 3,000, 2,000 to 4,000, 2,000 to 5,000, 2,000 to 6,000, 2,000 to 7,000, 2,000 to 8,000, 2,000 to 9,000, 2,000 to 10,000, 3,000 to 4,000, 3,000 to 5,000, 3,000 to 6,000, 3,000 to 7,000, 3,000 to 8,000, 3,000 to 9,000, 3,000 to 10,000, 4,000 to 5,000, 4,000 to 6,000, 4,000 to 7,000, 4,000 to 8,000, 4,000 to 9,000, 4,000 to 10,000, 5,000 to 6,000, 5,000 to 7,000, 5,000 to 8,000, 5,000 to 9,000, 5,000 to 10,000, 6,000 to 7,000, 6,000 to 8,000, 6,000 to 9,000, 6,000 to 10,000, 7,000 to 8,000, 7,000 to 9,000, 7,000 to 10,000, 8,000 to 9,000, 8,000 to 10,000, or 9,000 to 10,000 bases.

[00141] Provided herein are libraries comprising nucleic acids encoding for scaffolds comprising CRTH2R binding domains, wherein the nucleic acid libraries are expressed in a cell. In some instances, the libraries are synthesized to express a reporter gene. Exemplary reporter genes include, but are not limited to, acetohydroxyacid synthase (AHAS), alkaline phosphatase (AP), beta galactosidase (LacZ), beta glucuronidase (GUS), chloramphenicol acetyltransferase (CAT), green fluorescent protein (GFP), red fluorescent protein (RFP), yellow fluorescent protein (YFP), cyan fluorescent protein (CFP), cerulean fluorescent protein, citrine fluorescent protein, orange fluorescent protein, cherry fluorescent protein, turquoise fluorescent protein, blue fluorescent protein, horseradish peroxidase (HRP), luciferase (Luc), nopaline synthase (NOS), octopine synthase (OCS), luciferase, and derivatives thereof. Methods to determine modulation of a reporter

gene are well known in the art, and include, but are not limited to, fluorometric methods (*e.g.* fluorescence spectroscopy, Fluorescence Activated Cell Sorting (FACS), fluorescence microscopy), and antibiotic resistance determination.

[00142] Diseases and Disorders

[00143] Provided herein are CRTH2R binding libraries comprising nucleic acids encoding for scaffolds comprising CRTH2R binding domains that may have therapeutic effects. In some instances, the CRTH2R binding libraries result in protein when translated that is used to treat a disease or disorder. In some instances, the protein is an immunoglobulin. In some instances, the protein is a peptidomimetic. Exemplary diseases include, but are not limited to, cancer, inflammatory diseases or disorders, a metabolic disease or disorder, a cardiovascular disease or disorder, a respiratory disease or disorder, pain, a digestive disease or disorder, a reproductive disease or disorder, an endocrine disease or disorder, or a neurological disease or disorder. In some instances, the cancer is a solid cancer or a hematologic cancer. In some instances, an inhibitor of prostaglandin D2 receptor 2 (DP2 or CRTH2R) as described herein is used for treatment of a disease or disorder of the central nervous system, kidney, intestine, lung, hair, skin, bone, or cartilage. In some instances, an inhibitor or antagonist of CRTH2R as described herein is used for treatment of a disease or disorder characterized by an inflammatory response. In some instances, an inhibitor or antagonist of CRTH2R as described herein is used for treatment of an allergic reaction. In some instances, the allergic reaction is chronic idiopathic urticaria. In some instances, the allergic reaction is allergic rhinitis. In some instances, an inhibitor or antagonist of CRTH2R as described herein is used for treatment of asthma. In some instances, an inhibitor or antagonist of CRTH2R as described herein is used for treatment of alopecia or baldness. In some instances, the subject is a mammal. In some instances, the subject is a mouse, rabbit, dog, or human. Subjects treated by methods described herein may be infants, adults, or children. Pharmaceutical compositions comprising antibodies or antibody fragments as described herein may be administered intravenously or subcutaneously.

[00144] Variant Libraries

[00145] *Codon variation*

[00146] Variant nucleic acid libraries described herein may comprise a plurality of nucleic acids, wherein each nucleic acid encodes for a variant codon sequence compared to a reference nucleic acid sequence. In some instances, each nucleic acid of a first nucleic acid population contains a variant at a single variant site. In some instances, the first nucleic acid population contains a plurality of variants at a single variant site such that the first nucleic acid population contains more than one variant at the same variant site. The first nucleic acid population may comprise nucleic

acids collectively encoding multiple codon variants at the same variant site. The first nucleic acid population may comprise nucleic acids collectively encoding up to 19 or more codons at the same position. The first nucleic acid population may comprise nucleic acids collectively encoding up to 60 variant triplets at the same position, or the first nucleic acid population may comprise nucleic acids collectively encoding up to 61 different triplets of codons at the same position. Each variant may encode for a codon that results in a different amino acid during translation. **Table 3** provides a listing of each codon possible (and the representative amino acid) for a variant site.

Table 2. List of codons and amino acids

Amino Acids	One letter code	Three letter code	Codons					
Alanine	A	Ala	GCA	GCC	GCG	GCT		
Cysteine	C	Cys	TGC	TGT				
Aspartic acid	D	Asp	GAC	GAT				
Glutamic acid	E	Glu	GAA	GAG				
Phenylalanine	F	Phe	TTC	TTT				
Glycine	G	Gly	GGA	GGC	GGG	GGT		
Histidine	H	His	CAC	CAT				
Isoleucine	I	Iso	ATA	ATC	ATT			
Lysine	K	Lys	AAA	AAG				
Leucine	L	Leu	TTA	TTG	CTA	CTC	CTG	CTT
Methionine	M	Met	ATG					
Asparagine	N	Asn	AAC	AAT				
Proline	P	Pro	CCA	CCC	CCG	CCT		
Glutamine	Q	Gln	CAA	CAG				

Arginine	R	Arg	AGA	AGG	CGA	CGC	CGG	CGT
Serine	S	Ser	AGC	AGT	TCA	TCC	TCG	TCT
Threonine	T	Thr	ACA	ACC	ACG	ACT		
Valine	V	Val	GTA	GTC	GTG	GTT		
Tryptophan	W	Trp	TGG					
Tyrosine	Y	Tyr	TAC	TAT				

[00147] A nucleic acid population may comprise varied nucleic acids collectively encoding up to 20 codon variations at multiple positions. In such cases, each nucleic acid in the population comprises variation for codons at more than one position in the same nucleic acid. In some instances, each nucleic acid in the population comprises variation for codons at 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more codons in a single nucleic acid. In some instances, each variant long nucleic acid comprises variation for codons at 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 or more codons in a single long nucleic acid. In some instances, the variant nucleic acid population comprises variation for codons at 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 or more codons in a single nucleic acid. In some instances, the variant nucleic acid population comprises variation for codons in at least about 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 or more codons in a single long nucleic acid.

[00148] Highly Parallel Nucleic Acid Synthesis

[00149] Provided herein is a platform approach utilizing miniaturization, parallelization, and vertical integration of the end-to-end process from polynucleotide synthesis to gene assembly within nanowells on silicon to create a revolutionary synthesis platform. Devices described herein provide, with the same footprint as a 96-well plate, a silicon synthesis platform is capable of increasing throughput by a factor of up to 1,000 or more compared to traditional synthesis methods, with production of up to approximately 1,000,000 or more polynucleotides, or 10,000 or more genes in a single highly-parallelized run.

[00150] With the advent of next-generation sequencing, high resolution genomic data has become an important factor for studies that delve into the biological roles of various genes in both normal biology and disease pathogenesis. At the core of this research is the central dogma of molecular biology and the concept of “residue-by-residue transfer of sequential information.”

Genomic information encoded in the DNA is transcribed into a message that is then translated into the protein that is the active product within a given biological pathway.

[00151] Another exciting area of study is on the discovery, development and manufacturing of therapeutic molecules focused on a highly-specific cellular target. High diversity DNA sequence libraries are at the core of development pipelines for targeted therapeutics. Gene mutants are used to express proteins in a design, build, and test protein engineering cycle that ideally culminates in an optimized gene for high expression of a protein with high affinity for its therapeutic target. As an example, consider the binding pocket of a receptor. The ability to test all sequence permutations of all residues within the binding pocket simultaneously will allow for a thorough exploration, increasing chances of success. Saturation mutagenesis, in which a researcher attempts to generate all possible mutations at a specific site within the receptor, represents one approach to this development challenge. Though costly and time and labor-intensive, it enables each variant to be introduced into each position. In contrast, combinatorial mutagenesis, where a few selected positions or short stretch of DNA may be modified extensively, generates an incomplete repertoire of variants with biased representation.

[00152] To accelerate the drug development pipeline, a library with the desired variants available at the intended frequency in the right position available for testing—in other words, a precision library, enables reduced costs as well as turnaround time for screening. Provided herein are methods for synthesizing nucleic acid synthetic variant libraries which provide for precise introduction of each intended variant at the desired frequency. To the end user, this translates to the ability to not only thoroughly sample sequence space but also be able to query these hypotheses in an efficient manner, reducing cost and screening time. Genome-wide editing can elucidate important pathways, libraries where each variant and sequence permutation can be tested for optimal functionality, and thousands of genes can be used to reconstruct entire pathways and genomes to re-engineer biological systems for drug discovery.

[00153] In a first example, a drug itself can be optimized using methods described herein. For example, to improve a specified function of an antibody, a variant polynucleotide library encoding for a portion of the antibody is designed and synthesized. A variant nucleic acid library for the antibody can then be generated by processes described herein (e.g., PCR mutagenesis followed by insertion into a vector). The antibody is then expressed in a production cell line and screened for enhanced activity. Example screens include examining modulation in binding affinity to an antigen, stability, or effector function (e.g., ADCC, complement, or apoptosis). Exemplary regions to optimize the antibody include, without limitation, the Fc region, Fab region, variable region of

the Fab region, constant region of the Fab region, variable domain of the heavy chain or light chain (V_H or V_L), and specific complementarity-determining regions (CDRs) of V_H or V_L .

[00154] Nucleic acid libraries synthesized by methods described herein may be expressed in various cells associated with a disease state. Cells associated with a disease state include cell lines, tissue samples, primary cells from a subject, cultured cells expanded from a subject, or cells in a model system. Exemplary model systems include, without limitation, plant and animal models of a disease state.

[00155] To identify a variant molecule associated with prevention, reduction or treatment of a disease state, a variant nucleic acid library described herein is expressed in a cell associated with a disease state, or one in which a cell a disease state can be induced. In some instances, an agent is used to induce a disease state in cells. Exemplary tools for disease state induction include, without limitation, a Cre/Lox recombination system, LPS inflammation induction, and streptozotocin to induce hypoglycemia. The cells associated with a disease state may be cells from a model system or cultured cells, as well as cells from a subject having a particular disease condition. Exemplary disease conditions include a bacterial, fungal, viral, autoimmune, or proliferative disorder (e.g., cancer). In some instances, the variant nucleic acid library is expressed in the model system, cell line, or primary cells derived from a subject, and screened for changes in at least one cellular activity. Exemplary cellular activities include, without limitation, proliferation, cycle progression, cell death, adhesion, migration, reproduction, cell signaling, energy production, oxygen utilization, metabolic activity, and aging, response to free radical damage, or any combination thereof.

[00156] Substrates

[00157] Devices used as a surface for polynucleotide synthesis may be in the form of substrates which include, without limitation, homogenous array surfaces, patterned array surfaces, channels, beads, gels, and the like. Provided herein are substrates comprising a plurality of clusters, wherein each cluster comprises a plurality of loci that support the attachment and synthesis of polynucleotides. In some instances, substrates comprise a homogenous array surface. For example, the homogenous array surface is a homogenous plate. The term “locus” as used herein refers to a discrete region on a structure which provides support for polynucleotides encoding for a single predetermined sequence to extend from the surface. In some instances, a locus is on a two dimensional surface, e.g., a substantially planar surface. In some instances, a locus is on a three-dimensional surface, e.g., a well, microwell, channel, or post. In some instances, a surface of a locus comprises a material that is actively functionalized to attach to at least one nucleotide for polynucleotide synthesis, or preferably, a population of identical nucleotides for synthesis of a population of polynucleotides. In some instances, polynucleotide refers to a population of

polynucleotides encoding for the same nucleic acid sequence. In some cases, a surface of a substrate is inclusive of one or a plurality of surfaces of a substrate. The average error rates for polynucleotides synthesized within a library described here using the systems and methods provided are often less than 1 in 1000, less than about 1 in 2000, less than about 1 in 3000 or less often without error correction.

[00158] Provided herein are surfaces that support the parallel synthesis of a plurality of polynucleotides having different predetermined sequences at addressable locations on a common support. In some instances, a substrate provides support for the synthesis of more than 50, 100, 200, 400, 600, 800, 1000, 1200, 1400, 1600, 1800, 2,000; 5,000; 10,000; 20,000; 50,000; 100,000; 200,000; 300,000; 400,000; 500,000; 600,000; 700,000; 800,000; 900,000; 1,000,000; 1,200,000; 1,400,000; 1,600,000; 1,800,000; 2,000,000; 2,500,000; 3,000,000; 3,500,000; 4,000,000; 4,500,000; 5,000,000; 10,000,000 or more non-identical polynucleotides. In some cases, the surfaces provide support for the synthesis of more than 50, 100, 200, 400, 600, 800, 1000, 1200, 1400, 1600, 1800, 2,000; 5,000; 10,000; 20,000; 50,000; 100,000; 200,000; 300,000; 400,000; 500,000; 600,000; 700,000; 800,000; 900,000; 1,000,000; 1,200,000; 1,400,000; 1,600,000; 1,800,000; 2,000,000; 2,500,000; 3,000,000; 3,500,000; 4,000,000; 4,500,000; 5,000,000; 10,000,000 or more polynucleotides encoding for distinct sequences. In some instances, at least a portion of the polynucleotides have an identical sequence or are configured to be synthesized with an identical sequence. In some instances, the substrate provides a surface environment for the growth of polynucleotides having at least 80, 90, 100, 120, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500 bases or more.

[00159] Provided herein are methods for polynucleotide synthesis on distinct loci of a substrate, wherein each locus supports the synthesis of a population of polynucleotides. In some cases, each locus supports the synthesis of a population of polynucleotides having a different sequence than a population of polynucleotides grown on another locus. In some instances, each polynucleotide sequence is synthesized with 1, 2, 3, 4, 5, 6, 7, 8, 9 or more redundancy across different loci within the same cluster of loci on a surface for polynucleotide synthesis. In some instances, the loci of a substrate are located within a plurality of clusters. In some instances, a substrate comprises at least 10, 500, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 11000, 12000, 13000, 14000, 15000, 20000, 30000, 40000, 50000 or more clusters. In some instances, a substrate comprises more than 2,000; 5,000; 10,000; 100,000; 200,000; 300,000; 400,000; 500,000; 600,000; 700,000; 800,000; 900,000; 1,000,000; 1,100,000; 1,200,000; 1,300,000; 1,400,000; 1,500,000; 1,600,000; 1,700,000; 1,800,000; 1,900,000; 2,000,000; 300,000; 400,000; 500,000; 600,000; 700,000; 800,000; 900,000; 1,000,000; 1,200,000; 1,400,000; 1,600,000; 1,800,000; 2,000,000;

2,500,000; 3,000,000; 3,500,000; 4,000,000; 4,500,000; 5,000,000; or 10,000,000 or more distinct loci. In some instances, a substrate comprises about 10,000 distinct loci. The amount of loci within a single cluster is varied in different instances. In some cases, each cluster includes 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 120, 130, 150, 200, 300, 400, 500 or more loci. In some instances, each cluster includes about 50-500 loci. In some instances, each cluster includes about 100-200 loci. In some instances, each cluster includes about 100-150 loci. In some instances, each cluster includes about 109, 121, 130 or 137 loci. In some instances, each cluster includes about 19, 20, 61, 64 or more loci. Alternatively or in combination, polynucleotide synthesis occurs on a homogenous array surface.

[00160] In some instances, the number of distinct polynucleotides synthesized on a substrate is dependent on the number of distinct loci available in the substrate. In some instances, the density of loci within a cluster or surface of a substrate is at least or about 1, 10, 25, 50, 65, 75, 100, 130, 150, 175, 200, 300, 400, 500, 1,000 or more loci per mm^2 . In some cases, a substrate comprises 10-500, 25-400, 50-500, 100-500, 150-500, 10-250, 50-250, 10-200, or 50-200 mm^2 . In some instances, the distance between the centers of two adjacent loci within a cluster or surface is from about 10-500, from about 10-200, or from about 10-100 μm . In some instances, the distance between two centers of adjacent loci is greater than about 10, 20, 30, 40, 50, 60, 70, 80, 90 or 100 μm . In some instances, the distance between the centers of two adjacent loci is less than about 200, 150, 100, 80, 70, 60, 50, 40, 30, 20 or 10 μm . In some instances, each locus has a width of about 0.5, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90 or 100 μm . In some cases, each locus has a width of about 0.5-100, 0.5-50, 10-75, or 0.5-50 μm .

[00161] In some instances, the density of clusters within a substrate is at least or about 1 cluster per 100 mm^2 , 1 cluster per 10 mm^2 , 1 cluster per 5 mm^2 , 1 cluster per 4 mm^2 , 1 cluster per 3 mm^2 , 1 cluster per 2 mm^2 , 1 cluster per 1 mm^2 , 2 clusters per 1 mm^2 , 3 clusters per 1 mm^2 , 4 clusters per 1 mm^2 , 5 clusters per 1 mm^2 , 10 clusters per 1 mm^2 , 50 clusters per 1 mm^2 or more. In some instances, a substrate comprises from about 1 cluster per 10 mm^2 to about 10 clusters per 1 mm^2 . In some instances, the distance between the centers of two adjacent clusters is at least or about 50, 100, 200, 500, 1000, 2000, or 5000 μm . In some cases, the distance between the centers of two adjacent clusters is between about 50-100, 50-200, 50-300, 50-500, and 100-2000 μm . In some cases, the distance between the centers of two adjacent clusters is between about 0.05-50, 0.05-10, 0.05-5, 0.05-4, 0.05-3, 0.05-2, 0.1-10, 0.2-10, 0.3-10, 0.4-10, 0.5-10, 0.5-5, or 0.5-2 mm. In some cases, each cluster has a cross section of about 0.5 to about 2, about 0.5 to about 1, or about 1 to about 2 mm. In some cases, each cluster has a cross section of about 0.5, 0.6, 0.7, 0.8, 0.9, 1, 1.1,

1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9 or 2 mm. In some cases, each cluster has an interior cross section of about 0.5, 0.6, 0.7, 0.8, 0.9, 1, 1.1, 1.15, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9 or 2 mm.

[00162] In some instances, a substrate is about the size of a standard 96 well plate, for example between about 100 and about 200 mm by between about 50 and about 150 mm. In some instances, a substrate has a diameter less than or equal to about 1000, 500, 450, 400, 300, 250, 200, 150, 100 or 50 mm. In some instances, the diameter of a substrate is between about 25-1000, 25-800, 25-600, 25-500, 25-400, 25-300, or 25-200 mm. In some instances, a substrate has a planar surface area of at least about 100; 200; 500; 1,000; 2,000; 5,000; 10,000; 12,000; 15,000; 20,000; 30,000; 40,000; 50,000 mm² or more. In some instances, the thickness of a substrate is between about 50-2000, 50-1000, 100-1000, 200-1000, or 250-1000 mm.

[00163] *Surface materials*

[00164] Substrates, devices, and reactors provided herein are fabricated from any variety of materials suitable for the methods, compositions, and systems described herein. In certain instances, substrate materials are fabricated to exhibit a low level of nucleotide binding. In some instances, substrate materials are modified to generate distinct surfaces that exhibit a high level of nucleotide binding. In some instances, substrate materials are transparent to visible and/or UV light. In some instances, substrate materials are sufficiently conductive, *e.g.*, are able to form uniform electric fields across all or a portion of a substrate. In some instances, conductive materials are connected to an electric ground. In some instances, the substrate is heat conductive or insulated. In some instances, the materials are chemical resistant and heat resistant to support chemical or biochemical reactions, for example polynucleotide synthesis reaction processes. In some instances, a substrate comprises flexible materials. For flexible materials, materials can include, without limitation: nylon, both modified and unmodified, nitrocellulose, polypropylene, and the like. In some instances, a substrate comprises rigid materials. For rigid materials, materials can include, without limitation: glass; fused silica; silicon, plastics (for example polytetrafluoroethylene, polypropylene, polystyrene, polycarbonate, and blends thereof, and the like); metals (for example, gold, platinum, and the like). The substrate, solid support or reactors can be fabricated from a material selected from the group consisting of silicon, polystyrene, agarose, dextran, cellulosic polymers, polyacrylamides, polydimethylsiloxane (PDMS), and glass. The substrates/solid supports or the microstructures, reactors therein may be manufactured with a combination of materials listed herein or any other suitable material known in the art.

[00165] **Surface Architecture**

[00166] Provided herein are substrates for the methods, compositions, and systems described herein, wherein the substrates have a surface architecture suitable for the methods, compositions,

and systems described herein. In some instances, a substrate comprises raised and/or lowered features. One benefit of having such features is an increase in surface area to support polynucleotide synthesis. In some instances, a substrate having raised and/or lowered features is referred to as a three-dimensional substrate. In some cases, a three-dimensional substrate comprises one or more channels. In some cases, one or more loci comprise a channel. In some cases, the channels are accessible to reagent deposition via a deposition device such as a material deposition device. In some cases, reagents and/or fluids collect in a larger well in fluid communication one or more channels. For example, a substrate comprises a plurality of channels corresponding to a plurality of loci with a cluster, and the plurality of channels are in fluid communication with one well of the cluster. In some methods, a library of polynucleotides is synthesized in a plurality of loci of a cluster.

[00167] Provided herein are substrates for the methods, compositions, and systems described herein, wherein the substrates are configured for polynucleotide synthesis. In some instances, the structure is configured to allow for controlled flow and mass transfer paths for polynucleotide synthesis on a surface. In some instances, the configuration of a substrate allows for the controlled and even distribution of mass transfer paths, chemical exposure times, and/or wash efficacy during polynucleotide synthesis. In some instances, the configuration of a substrate allows for increased sweep efficiency, for example by providing sufficient volume for a growing polynucleotide such that the excluded volume by the growing polynucleotide does not take up more than 50, 45, 40, 35, 30, 25, 20, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1%, or less of the initially available volume that is available or suitable for growing the polynucleotide. In some instances, a three-dimensional structure allows for managed flow of fluid to allow for the rapid exchange of chemical exposure.

[00168] Provided herein are substrates for the methods, compositions, and systems described herein, wherein the substrates comprise structures suitable for the methods, compositions, and systems described herein. In some instances, segregation is achieved by physical structure. In some instances, segregation is achieved by differential functionalization of the surface generating active and passive regions for polynucleotide synthesis. In some instances, differential functionalization is achieved by alternating the hydrophobicity across the substrate surface, thereby creating water contact angle effects that cause beading or wetting of the deposited reagents. Employing larger structures can decrease splashing and cross-contamination of distinct polynucleotide synthesis locations with reagents of the neighboring spots. In some cases, a device, such as a material deposition device, is used to deposit reagents to distinct polynucleotide synthesis locations. Substrates having three-dimensional features are configured in a manner that allows for the synthesis of a large number of polynucleotides (*e.g.*, more than about 10,000) with a low error

rate (*e.g.*, less than about 1:500, 1:1000, 1:1500, 1:2,000, 1:3,000, 1:5,000, or 1:10,000). In some cases, a substrate comprises features with a density of about or greater than about 1, 5, 10, 20, 30, 40, 50, 60, 70, 80, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 300, 400 or 500 features per mm².

[00169] A well of a substrate may have the same or different width, height, and/or volume as another well of the substrate. A channel of a substrate may have the same or different width, height, and/or volume as another channel of the substrate. In some instances, the diameter of a cluster or the diameter of a well comprising a cluster, or both, is between about 0.05-50, 0.05-10, 0.05-5, 0.05-4, 0.05-3, 0.05-2, 0.05-1, 0.05-0.5, 0.05-0.1, 0.1-10, 0.2-10, 0.3-10, 0.4-10, 0.5-10, 0.5-5, or 0.5-2 mm. In some instances, the diameter of a cluster or well or both is less than or about 5, 4, 3, 2, 1, 0.5, 0.1, 0.09, 0.08, 0.07, 0.06, or 0.05 mm. In some instances, the diameter of a cluster or well or both is between about 1.0 and 1.3 mm. In some instances, the diameter of a cluster or well, or both is about 1.150 mm. In some instances, the diameter of a cluster or well, or both is about 0.08 mm. The diameter of a cluster refers to clusters within a two-dimensional or three-dimensional substrate.

[00170] In some instances, the height of a well is from about 20-1000, 50-1000, 100-1000, 200-1000, 300-1000, 400-1000, or 500-1000 μm . In some cases, the height of a well is less than about 1000, 900, 800, 700, or 600 μm .

[00171] In some instances, a substrate comprises a plurality of channels corresponding to a plurality of loci within a cluster, wherein the height or depth of a channel is 5-500, 5-400, 5-300, 5-200, 5-100, 5-50, or 10-50 μm . In some cases, the height of a channel is less than 100, 80, 60, 40, or 20 μm .

[00172] In some instances, the diameter of a channel, locus (*e.g.*, in a substantially planar substrate) or both channel and locus (*e.g.*, in a three-dimensional substrate wherein a locus corresponds to a channel) is from about 1-1000, 1-500, 1-200, 1-100, 5-100, or 10-100 μm , for example, about 90, 80, 70, 60, 50, 40, 30, 20 or 10 μm . In some instances, the diameter of a channel, locus, or both channel and locus is less than about 100, 90, 80, 70, 60, 50, 40, 30, 20 or 10 μm . In some instances, the distance between the center of two adjacent channels, loci, or channels and loci is from about 1-500, 1-200, 1-100, 5-200, 5-100, 5-50, or 5-30, for example, about 20 μm .

[00173] Surface Modifications

[00174] Provided herein are methods for polynucleotide synthesis on a surface, wherein the surface comprises various surface modifications. In some instances, the surface modifications are employed for the chemical and/or physical alteration of a surface by an additive or subtractive process to change one or more chemical and/or physical properties of a substrate surface or a

selected site or region of a substrate surface. For example, surface modifications include, without limitation, (1) changing the wetting properties of a surface, (2) functionalizing a surface, i.e., providing, modifying or substituting surface functional groups, (3) defunctionalizing a surface, i.e., removing surface functional groups, (4) otherwise altering the chemical composition of a surface, *e.g.*, through etching, (5) increasing or decreasing surface roughness, (6) providing a coating on a surface, *e.g.*, a coating that exhibits wetting properties that are different from the wetting properties of the surface, and/or (7) depositing particulates on a surface.

[00175] In some cases, the addition of a chemical layer on top of a surface (referred to as adhesion promoter) facilitates structured patterning of loci on a surface of a substrate. Exemplary surfaces for application of adhesion promotion include, without limitation, glass, silicon, silicon dioxide and silicon nitride. In some cases, the adhesion promoter is a chemical with a high surface energy. In some instances, a second chemical layer is deposited on a surface of a substrate. In some cases, the second chemical layer has a low surface energy. In some cases, surface energy of a chemical layer coated on a surface supports localization of droplets on the surface. Depending on the patterning arrangement selected, the proximity of loci and/or area of fluid contact at the loci are alterable.

[00176] In some instances, a substrate surface, or resolved loci, onto which nucleic acids or other moieties are deposited, *e.g.*, for polynucleotide synthesis, are smooth or substantially planar (*e.g.*, two-dimensional) or have irregularities, such as raised or lowered features (*e.g.*, three-dimensional features). In some instances, a substrate surface is modified with one or more different layers of compounds. Such modification layers of interest include, without limitation, inorganic and organic layers such as metals, metal oxides, polymers, small organic molecules and the like.

[00177] In some instances, resolved loci of a substrate are functionalized with one or more moieties that increase and/or decrease surface energy. In some cases, a moiety is chemically inert. In some cases, a moiety is configured to support a desired chemical reaction, for example, one or more processes in a polynucleotide synthesis reaction. The surface energy, or hydrophobicity, of a surface is a factor for determining the affinity of a nucleotide to attach onto the surface. In some instances, a method for substrate functionalization comprises: (a) providing a substrate having a surface that comprises silicon dioxide; and (b) silanizing the surface using, a suitable silanizing agent described herein or otherwise known in the art, for example, an organofunctional alkoxysilane molecule. Methods and functionalizing agents are described in U.S. Patent No. 5474796, which is herein incorporated by reference in its entirety.

[00178] In some instances, a substrate surface is functionalized by contact with a derivatizing composition that contains a mixture of silanes, under reaction conditions effective to couple the

silanes to the substrate surface, typically via reactive hydrophilic moieties present on the substrate surface. Silanization generally covers a surface through self-assembly with organofunctional alkoxy silane molecules. A variety of siloxane functionalizing reagents can further be used as currently known in the art, *e.g.*, for lowering or increasing surface energy. The organofunctional alkoxy silanes are classified according to their organic functions.

[00179] Polynucleotide Synthesis

[00180] Methods of the current disclosure for polynucleotide synthesis may include processes involving phosphoramidite chemistry. In some instances, polynucleotide synthesis comprises coupling a base with phosphoramidite. Polynucleotide synthesis may comprise coupling a base by deposition of phosphoramidite under coupling conditions, wherein the same base is optionally deposited with phosphoramidite more than once, *i.e.*, double coupling. Polynucleotide synthesis may comprise capping of unreacted sites. In some instances, capping is optional. Polynucleotide synthesis may also comprise oxidation or an oxidation step or oxidation steps. Polynucleotide synthesis may comprise deblocking, detritylation, and sulfurization. In some instances, polynucleotide synthesis comprises either oxidation or sulfurization. In some instances, between one or each step during a polynucleotide synthesis reaction, the device is washed, for example, using tetrazole or acetonitrile. Time frames for any one step in a phosphoramidite synthesis method may be less than about 2 min, 1 min, 50 sec, 40 sec, 30 sec, 20 sec and 10 sec.

[00181] Polynucleotide synthesis using a phosphoramidite method may comprise a subsequent addition of a phosphoramidite building block (*e.g.*, nucleoside phosphoramidite) to a growing polynucleotide chain for the formation of a phosphite triester linkage. Phosphoramidite polynucleotide synthesis proceeds in the 3' to 5' direction. Phosphoramidite polynucleotide synthesis allows for the controlled addition of one nucleotide to a growing nucleic acid chain per synthesis cycle. In some instances, each synthesis cycle comprises a coupling step. Phosphoramidite coupling involves the formation of a phosphite triester linkage between an activated nucleoside phosphoramidite and a nucleoside bound to the substrate, for example, via a linker. In some instances, the nucleoside phosphoramidite is provided to the device activated. In some instances, the nucleoside phosphoramidite is provided to the device with an activator. In some instances, nucleoside phosphoramidites are provided to the device in a 1.5, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100-fold excess or more over the substrate-bound nucleosides. In some instances, the addition of nucleoside phosphoramidite is performed in an anhydrous environment, for example, in anhydrous acetonitrile. Following addition of a nucleoside phosphoramidite, the device is optionally washed. In some instances, the coupling step is repeated one or more additional times, optionally with a wash step

between nucleoside phosphoramidite additions to the substrate. In some instances, a polynucleotide synthesis method used herein comprises 1, 2, 3 or more sequential coupling steps. Prior to coupling, in many cases, the nucleoside bound to the device is de-protected by removal of a protecting group, where the protecting group functions to prevent polymerization. A common protecting group is 4,4'-dimethoxytrityl (DMT).

[00182] Following coupling, phosphoramidite polynucleotide synthesis methods optionally comprise a capping step. In a capping step, the growing polynucleotide is treated with a capping agent. A capping step is useful to block unreacted substrate-bound 5'-OH groups after coupling from further chain elongation, preventing the formation of polynucleotides with internal base deletions. Further, phosphoramidites activated with 1H-tetrazole may react, to a small extent, with the O6 position of guanosine. Without being bound by theory, upon oxidation with I₂/water, this side product, possibly via O6-N7 migration, may undergo depurination. The apurinic sites may end up being cleaved in the course of the final deprotection of the polynucleotide thus reducing the yield of the full-length product. The O6 modifications may be removed by treatment with the capping reagent prior to oxidation with I₂/water. In some instances, inclusion of a capping step during polynucleotide synthesis decreases the error rate as compared to synthesis without capping. As an example, the capping step comprises treating the substrate-bound polynucleotide with a mixture of acetic anhydride and 1-methylimidazole. Following a capping step, the device is optionally washed.

[00183] In some instances, following addition of a nucleoside phosphoramidite, and optionally after capping and one or more wash steps, the device bound growing nucleic acid is oxidized. The oxidation step comprises the phosphite triester is oxidized into a tetracoordinated phosphate triester, a protected precursor of the naturally occurring phosphate diester internucleoside linkage. In some instances, oxidation of the growing polynucleotide is achieved by treatment with iodine and water, optionally in the presence of a weak base (*e.g.*, pyridine, lutidine, collidine). Oxidation may be carried out under anhydrous conditions using, *e.g.* tert-Butyl hydroperoxide or (1S)-(+)-(10-camphorsulfonyl)-oxaziridine (CSO). In some methods, a capping step is performed following oxidation. A second capping step allows for device drying, as residual water from oxidation that may persist can inhibit subsequent coupling. Following oxidation, the device and growing polynucleotide is optionally washed. In some instances, the step of oxidation is substituted with a sulfurization step to obtain polynucleotide phosphorothioates, wherein any capping steps can be performed after the sulfurization. Many reagents are capable of the efficient sulfur transfer, including but not limited to 3-(Dimethylaminomethylidene)amino)-3H-1,2,4-dithiazole-3-thione,

DDTT, 3H-1,2-benzodithiol-3-one 1,1-dioxide, also known as Beaucage reagent, and N,N,N',N'-Tetraethylthiuram disulfide (TETD).

[00184] In order for a subsequent cycle of nucleoside incorporation to occur through coupling, the protected 5' end of the device bound growing polynucleotide is removed so that the primary hydroxyl group is reactive with a next nucleoside phosphoramidite. In some instances, the protecting group is DMT and deblocking occurs with trichloroacetic acid in dichloromethane. Conducting detritylation for an extended time or with stronger than recommended solutions of acids may lead to increased depurination of solid support-bound polynucleotide and thus reduces the yield of the desired full-length product. Methods and compositions of the disclosure described herein provide for controlled deblocking conditions limiting undesired depurination reactions. In some instances, the device bound polynucleotide is washed after deblocking. In some instances, efficient washing after deblocking contributes to synthesized polynucleotides having a low error rate.

[00185] Methods for the synthesis of polynucleotides typically involve an iterating sequence of the following steps: application of a protected monomer to an actively functionalized surface (*e.g.*, locus) to link with either the activated surface, a linker or with a previously deprotected monomer; deprotection of the applied monomer so that it is reactive with a subsequently applied protected monomer; and application of another protected monomer for linking. One or more intermediate steps include oxidation or sulfurization. In some instances, one or more wash steps precede or follow one or all of the steps.

[00186] Methods for phosphoramidite-based polynucleotide synthesis comprise a series of chemical steps. In some instances, one or more steps of a synthesis method involve reagent cycling, where one or more steps of the method comprise application to the device of a reagent useful for the step. For example, reagents are cycled by a series of liquid deposition and vacuum drying steps. For substrates comprising three-dimensional features such as wells, microwells, channels and the like, reagents are optionally passed through one or more regions of the device via the wells and/or channels.

[00187] Methods and systems described herein relate to polynucleotide synthesis devices for the synthesis of polynucleotides. The synthesis may be in parallel. For example, at least or about at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 1000, 10000, 50000, 75000, 100000 or more polynucleotides can be synthesized in parallel. The total number polynucleotides that may be synthesized in parallel may be from 2-100000, 3-50000, 4-10000, 5-1000, 6-900, 7-850, 8-800, 9-750, 10-700, 11-650, 12-600, 13-550, 14-500, 15-450, 16-

400, 17-350, 18-300, 19-250, 20-200, 21-150, 22-100, 23-50, 24-45, 25-40, 30-35. Those of skill in the art appreciate that the total number of polynucleotides synthesized in parallel may fall within any range bound by any of these values, for example 25-100. The total number of polynucleotides synthesized in parallel may fall within any range defined by any of the values serving as endpoints of the range. Total molar mass of polynucleotides synthesized within the device or the molar mass of each of the polynucleotides may be at least or at least about 10, 20, 30, 40, 50, 100, 250, 500, 750, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 25000, 50000, 75000, 100000 picomoles, or more. The length of each of the polynucleotides or average length of the polynucleotides within the device may be at least or about at least 10, 15, 20, 25, 30, 35, 40, 45, 50, 100, 150, 200, 300, 400, 500 nucleotides, or more. The length of each of the polynucleotides or average length of the polynucleotides within the device may be at most or about at most 500, 400, 300, 200, 150, 100, 50, 45, 35, 30, 25, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10 nucleotides, or less. The length of each of the polynucleotides or average length of the polynucleotides within the device may fall from 10-500, 9-400, 11-300, 12-200, 13-150, 14-100, 15-50, 16-45, 17-40, 18-35, 19-25. Those of skill in the art appreciate that the length of each of the polynucleotides or average length of the polynucleotides within the device may fall within any range bound by any of these values, for example 100-300. The length of each of the polynucleotides or average length of the polynucleotides within the device may fall within any range defined by any of the values serving as endpoints of the range.

[00188] Methods for polynucleotide synthesis on a surface provided herein allow for synthesis at a fast rate. As an example, at least 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, 35, 40, 45, 50, 55, 60, 70, 80, 90, 100, 125, 150, 175, 200 nucleotides per hour, or more are synthesized. Nucleotides include adenine, guanine, thymine, cytosine, uridine building blocks, or analogs/modified versions thereof. In some instances, libraries of polynucleotides are synthesized in parallel on substrate. For example, a device comprising about or at least about 100; 1,000; 10,000; 30,000; 75,000; 100,000; 1,000,000; 2,000,000; 3,000,000; 4,000,000; or 5,000,000 resolved loci is able to support the synthesis of at least the same number of distinct polynucleotides, wherein polynucleotide encoding a distinct sequence is synthesized on a resolved locus. In some instances, a library of polynucleotides is synthesized on a device with low error rates described herein in less than about three months, two months, one month, three weeks, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2 days, 24 hours or less. In some instances, larger nucleic acids assembled from a polynucleotide library synthesized with low error rate using the substrates and methods described herein are prepared in less than about three months, two months, one month, three weeks, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2 days, 24 hours or less.

[00189] In some instances, methods described herein provide for generation of a library of nucleic acids comprising variant nucleic acids differing at a plurality of codon sites. In some instances, a nucleic acid may have 1 site, 2 sites, 3 sites, 4 sites, 5 sites, 6 sites, 7 sites, 8 sites, 9 sites, 10 sites, 11 sites, 12 sites, 13 sites, 14 sites, 15 sites, 16 sites, 17 sites, 18 sites, 19 sites, 20 sites, 30 sites, 40 sites, 50 sites, or more of variant codon sites.

[00190] In some instances, the one or more sites of variant codon sites may be adjacent. In some instances, the one or more sites of variant codon sites may not be adjacent and separated by 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more codons.

[00191] In some instances, a nucleic acid may comprise multiple sites of variant codon sites, wherein all the variant codon sites are adjacent to one another, forming a stretch of variant codon sites. In some instances, a nucleic acid may comprise multiple sites of variant codon sites, wherein none the variant codon sites are adjacent to one another. In some instances, a nucleic acid may comprise multiple sites of variant codon sites, wherein some the variant codon sites are adjacent to one another, forming a stretch of variant codon sites, and some of the variant codon sites are not adjacent to one another.

[00192] Referring to the Figures, **FIG. 7** illustrates an exemplary process workflow for synthesis of nucleic acids (e.g., genes) from shorter nucleic acids. The workflow is divided generally into phases: (1) de novo synthesis of a single stranded nucleic acid library, (2) joining nucleic acids to form larger fragments, (3) error correction, (4) quality control, and (5) shipment. Prior to de novo synthesis, an intended nucleic acid sequence or group of nucleic acid sequences is preselected. For example, a group of genes is preselected for generation.

[00193] Once large nucleic acids for generation are selected, a predetermined library of nucleic acids is designed for de novo synthesis. Various suitable methods are known for generating high density polynucleotide arrays. In the workflow example, a device surface layer is provided. In the example, chemistry of the surface is altered in order to improve the polynucleotide synthesis process. Areas of low surface energy are generated to repel liquid while areas of high surface energy are generated to attract liquids. The surface itself may be in the form of a planar surface or contain variations in shape, such as protrusions or microwells which increase surface area. In the workflow example, high surface energy molecules selected serve a dual function of supporting DNA chemistry, as disclosed in International Patent Application Publication WO/2015/021080, which is herein incorporated by reference in its entirety.

[00194] In situ preparation of polynucleotide arrays is generated on a solid support **701** and utilizes single nucleotide extension process to extend multiple oligomers in parallel. A deposition device, such as a material deposition device, is designed to release reagents in a step wise fashion

such that multiple polynucleotides extend, in parallel, one residue at a time to generate oligomers with a predetermined nucleic acid sequence **702**. In some instances, polynucleotides are cleaved from the surface at this stage. Cleavage includes gas cleavage, *e.g.*, with ammonia or methylamine.

[00195] The generated polynucleotide libraries are placed in a reaction chamber. In this exemplary workflow, the reaction chamber (also referred to as “nanoreactor”) is a silicon coated well, containing PCR reagents and lowered onto the polynucleotide library **703**. Prior to or after the sealing **704** of the polynucleotides, a reagent is added to release the polynucleotides from the substrate. In the exemplary workflow, the polynucleotides are released subsequent to sealing of the nanoreactor **705**. Once released, fragments of single stranded polynucleotides hybridize in order to span an entire long range sequence of DNA. Partial hybridization **705** is possible because each synthesized polynucleotide is designed to have a small portion overlapping with at least one other polynucleotide in the pool.

[00196] After hybridization, a PCA reaction is commenced. During the polymerase cycles, the polynucleotides anneal to complementary fragments and gaps are filled in by a polymerase. Each cycle increases the length of various fragments randomly depending on which polynucleotides find each other. Complementarity amongst the fragments allows for forming a complete large span of double stranded DNA **706**.

[00197] After PCA is complete, the nanoreactor is separated from the device **707** and positioned for interaction with a device having primers for PCR **708**. After sealing, the nanoreactor is subject to PCR **709** and the larger nucleic acids are amplified. After PCR **710**, the nanochamber is opened **711**, error correction reagents are added **712**, the chamber is sealed **713** and an error correction reaction occurs to remove mismatched base pairs and/or strands with poor complementarity from the double stranded PCR amplification products **714**. The nanoreactor is opened and separated **715**. Error corrected product is next subject to additional processing steps, such as PCR and molecular bar coding, and then packaged **722** for shipment **723**.

[00198] In some instances, quality control measures are taken. After error correction, quality control steps include for example interaction with a wafer having sequencing primers for amplification of the error corrected product **716**, sealing the wafer to a chamber containing error corrected amplification product **717**, and performing an additional round of amplification **718**. The nanoreactor is opened **719** and the products are pooled **720** and sequenced **721**. After an acceptable quality control determination is made, the packaged product **722** is approved for shipment **723**.

[00199] In some instances, a nucleic acid generated by a workflow such as that in **FIG. 7** is subject to mutagenesis using overlapping primers disclosed herein. In some instances, a library of primers are generated by in situ preparation on a solid support and utilize single nucleotide

extension process to extend multiple oligomers in parallel. A deposition device, such as a material deposition device, is designed to release reagents in a step wise fashion such that multiple polynucleotides extend, in parallel, one residue at a time to generate oligomers with a predetermined nucleic acid sequence **702**.

[00200] Computer systems

[00201] Any of the systems described herein, may be operably linked to a computer and may be automated through a computer either locally or remotely. In various instances, the methods and systems of the disclosure may further comprise software programs on computer systems and use thereof. Accordingly, computerized control for the synchronization of the dispense/vacuum/refill functions such as orchestrating and synchronizing the material deposition device movement, dispense action and vacuum actuation are within the bounds of the disclosure. The computer systems may be programmed to interface between the user specified base sequence and the position of a material deposition device to deliver the correct reagents to specified regions of the substrate.

[00202] The computer system **800** illustrated in **FIG. 8** may be understood as a logical apparatus that can read instructions from media **811** and/or a network port **805**, which can optionally be connected to server **809** having fixed media **812**. The system, such as shown in **FIG. 8** can include a CPU **801**, disk drives **803**, optional input devices such as keyboard **815** and/or mouse **816** and optional monitor **807**. Data communication can be achieved through the indicated communication medium to a server at a local or a remote location. The communication medium can include any means of transmitting and/or receiving data. For example, the communication medium can be a network connection, a wireless connection or an internet connection. Such a connection can provide for communication over the World Wide Web. It is envisioned that data relating to the present disclosure can be transmitted over such networks or connections for reception and/or review by a party **822** as illustrated in **FIG. 8**.

[00203] As illustrated in **FIG. 9**, a high speed cache **904** can be connected to, or incorporated in, the processor **902** to provide a high speed memory for instructions or data that have been recently, or are frequently, used by processor **902**. The processor **902** is connected to a north bridge **906** by a processor bus **908**. The north bridge **906** is connected to random access memory (RAM) **910** by a memory bus **912** and manages access to the RAM **910** by the processor **902**. The north bridge **906** is also connected to a south bridge **914** by a chipset bus **916**. The south bridge **914** is, in turn, connected to a peripheral bus **918**. The peripheral bus can be, for example, PCI, PCI-X, PCI Express, or other peripheral bus. The north bridge and south bridge are often referred to as a processor chipset and manage data transfer between the processor, RAM, and peripheral components on the peripheral bus **918**. In some alternative architectures, the functionality of the

north bridge can be incorporated into the processor instead of using a separate north bridge chip. In some instances, system **900** can include an accelerator card **922** attached to the peripheral bus **918**. The accelerator can include field programmable gate arrays (FPGAs) or other hardware for accelerating certain processing. For example, an accelerator can be used for adaptive data restructuring or to evaluate algebraic expressions used in extended set processing.

[00204] Software and data are stored in external storage **924** and can be loaded into RAM **910** and/or cache **904** for use by the processor. The system **900** includes an operating system for managing system resources; non-limiting examples of operating systems include: Linux, Windows™, MACOS™, BlackBerry OSTM, iOS™, and other functionally-equivalent operating systems, as well as application software running on top of the operating system for managing data storage and optimization in accordance with example instances of the present disclosure. In this example, system **900** also includes network interface cards (NICs) **920** and **921** connected to the peripheral bus for providing network interfaces to external storage, such as Network Attached Storage (NAS) and other computer systems that can be used for distributed parallel processing.

[00205] **FIG. 10** is a diagram showing a network **1000** with a plurality of computer systems **1002a**, and **1002b**, a plurality of cell phones and personal data assistants **1002c**, and Network Attached Storage (NAS) **1004a**, and **1004b**. In example instances, systems **1002a**, **1002b**, and **1002c** can manage data storage and optimize data access for data stored in Network Attached Storage (NAS) **1004a** and **1004b**. A mathematical model can be used for the data and be evaluated using distributed parallel processing across computer systems **1002a**, and **1002b**, and cell phone and personal data assistant systems **1002c**. Computer systems **1002a**, and **1002b**, and cell phone and personal data assistant systems **1002c** can also provide parallel processing for adaptive data restructuring of the data stored in Network Attached Storage (NAS) **1004a** and **1004b**. **FIG. 10** illustrates an example only, and a wide variety of other computer architectures and systems can be used in conjunction with the various instances of the present disclosure. For example, a blade server can be used to provide parallel processing. Processor blades can be connected through a back plane to provide parallel processing. Storage can also be connected to the back plane or as Network Attached Storage (NAS) through a separate network interface. In some example instances, processors can maintain separate memory spaces and transmit data through network interfaces, back plane or other connectors for parallel processing by other processors. In other instances, some or all of the processors can use a shared virtual address memory space.

[00206] **FIG. 11** is a block diagram of a multiprocessor computer system using a shared virtual address memory space in accordance with an example instance. The system includes a plurality of

processors **1102a-f** that can access a shared memory subsystem **1104**. The system incorporates a plurality of programmable hardware memory algorithm processors (MAPs) **1106a-f** in the memory subsystem **1104**. Each MAP **1106a-f** can comprise a memory **1108a-f** and one or more field programmable gate arrays (FPGAs) **1110a-f**. The MAP provides a configurable functional unit and particular algorithms or portions of algorithms can be provided to the FPGAs **1110a-f** for processing in close coordination with a respective processor. For example, the MAPs can be used to evaluate algebraic expressions regarding the data model and to perform adaptive data restructuring in example instances. In this example, each MAP is globally accessible by all of the processors for these purposes. In one configuration, each MAP can use Direct Memory Access (DMA) to access an associated memory **1108a-f**, allowing it to execute tasks independently of, and asynchronously from the respective microprocessor **1102a-f**. In this configuration, a MAP can feed results directly to another MAP for pipelining and parallel execution of algorithms.

[00207] The above computer architectures and systems are examples only, and a wide variety of other computer, cell phone, and personal data assistant architectures and systems can be used in connection with example instances, including systems using any combination of general processors, co-processors, FPGAs and other programmable logic devices, system on chips (SOCs), application specific integrated circuits (ASICs), and other processing and logic elements. In some instances, all or part of the computer system can be implemented in software or hardware. Any variety of data storage media can be used in connection with example instances, including random access memory, hard drives, flash memory, tape drives, disk arrays, Network Attached Storage (NAS) and other local or distributed data storage devices and systems.

[00208] In example instances, the computer system can be implemented using software modules executing on any of the above or other computer architectures and systems. In other instances, the functions of the system can be implemented partially or completely in firmware, programmable logic devices such as field programmable gate arrays (FPGAs) as referenced in **FIG. 9**, system on chips (SOCs), application specific integrated circuits (ASICs), or other processing and logic elements. For example, the Set Processor and Optimizer can be implemented with hardware acceleration through the use of a hardware accelerator card, such as accelerator card **922** illustrated in **FIG. 9**.

[00209] The following examples are set forth to illustrate more clearly the principle and practice of embodiments disclosed herein to those skilled in the art and are not to be construed as limiting the scope of any claimed embodiments. Unless otherwise stated, all parts and percentages are on a weight basis.

EXAMPLES

[00210] The following examples are given for the purpose of illustrating various embodiments of the disclosure and are not meant to limit the present disclosure in any fashion. The present examples, along with the methods described herein are presently representative of preferred embodiments, are exemplary, and are not intended as limitations on the scope of the disclosure. Changes therein and other uses which are encompassed within the spirit of the disclosure as defined by the scope of the claims will occur to those skilled in the art.

[00211] Example 1: Functionalization of a device surface

[00212] A device was functionalized to support the attachment and synthesis of a library of polynucleotides. The device surface was first wet cleaned using a piranha solution comprising 90% H₂SO₄ and 10% H₂O₂ for 20 minutes. The device was rinsed in several beakers with DI water, held under a DI water gooseneck faucet for 5 min, and dried with N₂. The device was subsequently soaked in NH₄OH (1:100; 3 mL:300 mL) for 5 min, rinsed with DI water using a handgun, soaked in three successive beakers with DI water for 1 min each, and then rinsed again with DI water using the handgun. The device was then plasma cleaned by exposing the device surface to O₂. A SAMCO PC-300 instrument was used to plasma etch O₂ at 250 watts for 1 min in downstream mode.

[00213] The cleaned device surface was actively functionalized with a solution comprising N-(3-triethoxysilylpropyl)-4-hydroxybutyramide using a YES-1224P vapor deposition oven system with the following parameters: 0.5 to 1 torr, 60 min, 70 °C, 135 °C vaporizer. The device surface was resist coated using a Brewer Science 200X spin coater. SPRTM 3612 photoresist was spin coated on the device at 2500 rpm for 40 sec. The device was pre-baked for 30 min at 90 °C on a Brewer hot plate. The device was subjected to photolithography using a Karl Suss MA6 mask aligner instrument. The device was exposed for 2.2 sec and developed for 1 min in MSF 26A. Remaining developer was rinsed with the handgun and the device soaked in water for 5 min. The device was baked for 30 min at 100 °C in the oven, followed by visual inspection for lithography defects using a Nikon L200. A descum process was used to remove residual resist using the SAMCO PC-300 instrument to O₂ plasma etch at 250 watts for 1 min.

[00214] The device surface was passively functionalized with a 100 µL solution of perfluorooctyltrichlorosilane mixed with 10 µL light mineral oil. The device was placed in a chamber, pumped for 10 min, and then the valve was closed to the pump and left to stand for 10 min. The chamber was vented to air. The device was resist stripped by performing two soaks for 5 min in 500 mL NMP at 70 °C with ultrasonication at maximum power (9 on Crest system). The device was then soaked for 5 min in 500 mL isopropanol at room temperature with ultrasonication

at maximum power. The device was dipped in 300 mL of 200 proof ethanol and blown dry with N₂. The functionalized surface was activated to serve as a support for polynucleotide synthesis.

[00215] Example 2: Synthesis of a 50-mer sequence on an oligonucleotide synthesis device

[00216] A two dimensional oligonucleotide synthesis device was assembled into a flowcell, which was connected to a flowcell (Applied Biosystems (ABI394 DNA Synthesizer"). The two-dimensional oligonucleotide synthesis device was uniformly functionalized with N-(3-TRIETHOXYSILYLPROPYL)-4-HYDROXYBUTYRAMIDE (Gelest) was used to synthesize an exemplary polynucleotide of 50 bp ("50-mer polynucleotide") using polynucleotide synthesis methods described herein.

[00217] The sequence of the 50-mer was as described in SEQ ID NO.: 2.

5'AGACAATCAACCATTTGGGGTGGACAGCCTTGACCTCTAGACTTCGGCAT##TTTTTTT TTT3' (SEQ ID NO.: 2), where # denotes Thymidine-succinyl hexamide CED phosphoramidite (CLP-2244 from ChemGenes), which is a cleavable linker enabling the release of oligos from the surface during deprotection.

[00218] The synthesis was done using standard DNA synthesis chemistry (coupling, capping, oxidation, and deblocking) according to the protocol in **Table 3** and an ABI synthesizer.

Table 3: Synthesis protocols

Table 3		
General DNA Synthesis Process Name	Process Step	Time (sec)
WASH (Acetonitrile Wash Flow)	Acetonitrile System Flush	4
	Acetonitrile to Flowcell	23
	N ₂ System Flush	4
	Acetonitrile System Flush	4
DNA BASE ADDITION (Phosphoramidite + Activator Flow)	Activator Manifold Flush	2
	Activator to Flowcell	6
	Activator + Phosphoramidite to Flowcell	6
	Activator to Flowcell	0.5
	Activator + Phosphoramidite to Flowcell	5
	Activator to Flowcell	0.5
	Activator + Phosphoramidite to Flowcell	5
	Activator to Flowcell	0.5
	Activator + Phosphoramidite to Flowcell	5
	Activator to Flowcell	0.5
	Incubate for 25sec	25

Table 3		
General DNA Synthesis Process Name	Process Step	Time (sec)
WASH (Acetonitrile Wash Flow)	Acetonitrile System Flush	4
	Acetonitrile to Flowcell	15
	N2 System Flush	4
	Acetonitrile System Flush	4
DNA BASE ADDITION (Phosphoramidite + Activator Flow)	Activator Manifold Flush	2
	Activator to Flowcell	5
	Activator + Phosphoramidite to Flowcell	18
	Incubate for 25sec	25
WASH (Acetonitrile Wash Flow)	Acetonitrile System Flush	4
	Acetonitrile to Flowcell	15
	N2 System Flush	4
	Acetonitrile System Flush	4
CAPPING (CapA+B, 1:1, Flow)	CapA+B to Flowcell	15
WASH (Acetonitrile Wash Flow)	Acetonitrile System Flush	4
	Acetonitrile to Flowcell	15
	Acetonitrile System Flush	4
OXIDATION (Oxidizer Flow)	Oxidizer to Flowcell	18
WASH (Acetonitrile Wash Flow)	Acetonitrile System Flush	4
	N2 System Flush	4
	Acetonitrile System Flush	4
	Acetonitrile to Flowcell	15
	Acetonitrile System Flush	4
	Acetonitrile to Flowcell	15
	N2 System Flush	4
	Acetonitrile System Flush	4
	Acetonitrile to Flowcell	23
	N2 System Flush	4
	Acetonitrile System Flush	4
DEBLOCKING (Deblock Flow)	Deblock to Flowcell	36
WASH (Acetonitrile Wash Flow)	Acetonitrile System Flush	4
	N2 System Flush	4
	Acetonitrile System Flush	4
	Acetonitrile to Flowcell	18
	N2 System Flush	4.13
	Acetonitrile System Flush	4.13
	Acetonitrile to Flowcell	15

[00219] The phosphoramidite/activator combination was delivered similar to the delivery of bulk reagents through the flowcell. No drying steps were performed as the environment stays "wet" with reagent the entire time.

[00220] The flow restrictor was removed from the ABI 394 synthesizer to enable faster flow. Without flow restrictor, flow rates for amidites (0.1M in ACN), Activator, (0.25M Benzoylthiotetrazole ("BTT"; 30-3070-xx from GlenResearch) in ACN), and Ox (0.02M I2 in 20% pyridine, 10% water, and 70% THF) were roughly ~100uL/sec, for acetonitrile ("ACN") and capping reagents (1:1 mix of CapA and CapB, wherein CapA is acetic anhydride in THF/Pyridine and CapB is 16% 1-methylimidazole in THF), roughly ~200uL/sec, and for Deblock (3% dichloroacetic acid in toluene), roughly ~300uL/sec (compared to ~50uL/sec for all reagents with flow restrictor). The time to completely push out Oxidizer was observed, the timing for chemical flow times was adjusted accordingly and an extra ACN wash was introduced between different chemicals. After polynucleotide synthesis, the chip was deprotected in gaseous ammonia overnight at 75 psi. Five drops of water were applied to the surface to recover polynucleotides. The recovered polynucleotides were then analyzed on a BioAnalyzer small RNA chip.

[00221] Example 3: Synthesis of a 100-mer sequence on an oligonucleotide synthesis device

[00222] The same process as described in Example 2 for the synthesis of the 50-mer sequence was used for the synthesis of a 100-mer polynucleotide ("100-mer polynucleotide"; 5' CGGGATCCTTATCGTCATCGTCGTACAGATCCCGACCCATTTGCTGTCCACCAGTCATG CTAGCCATACCATGATGATGATGATGATGAGAACCCCGCAT##TTTTTTTTTTT3', where # denotes Thymidine-succinyl hexamide CED phosphoramidite (CLP-2244 from ChemGenes); SEQ ID NO.: 3) on two different silicon chips, the first one uniformly functionalized with N-(3-TRIETHOXSILYLPROPYL)-4-HYDROXYBUTYRAMIDE and the second one functionalized with 5/95 mix of 11-acetoxundecyltriethoxysilane and n-decyltriethoxysilane, and the polynucleotides extracted from the surface were analyzed on a BioAnalyzer instrument.

[00223] All ten samples from the two chips were further PCR amplified using a forward (5'ATGCGGGGTTCTCATCATC3'; SEQ ID NO.: 4) and a reverse (5'CGGGATCCTTATCGTCATCG3'; SEQ ID NO.: 5) primer in a 50uL PCR mix (25uL NEB Q5 mastermix, 2.5uL 10uM Forward primer, 2.5uL 10uM Reverse primer, 1uL polynucleotide extracted from the surface, and water up to 50uL) using the following thermalcycling program:

98 °C, 30 sec

98 °C, 10 sec; 63 °C, 10 sec; 72 °C, 10 sec; repeat 12 cycles

72 °C, 2min

[00224] The PCR products were also run on a BioAnalyzer, demonstrating sharp peaks at the 100-mer position. Next, the PCR amplified samples were cloned, and Sanger sequenced. **Table 4** summarizes the results from the Sanger sequencing for samples taken from spots 1-5 from chip 1 and for samples taken from spots 6-10 from chip 2.

Table 4: Sequencing results

Spot	Error rate	Cycle efficiency
1	1/763 bp	99.87%
2	1/824 bp	99.88%
3	1/780 bp	99.87%
4	1/429 bp	99.77%
5	1/1525 bp	99.93%
6	1/1615 bp	99.94%
7	1/531 bp	99.81%
8	1/1769 bp	99.94%
9	1/854 bp	99.88%
10	1/1451 bp	99.93%

[00225] Thus, the high quality and uniformity of the synthesized polynucleotides were repeated on two chips with different surface chemistries. Overall, 89% of the 100-mers that were sequenced were perfect sequences with no errors, corresponding to 233 out of 262.

[00226] **Table 5** summarizes error characteristics for the sequences obtained from the polynucleotide samples from spots 1-10.

Table 5: Error characteristics

Sample ID/Spot no.	OSA_0 046/1	OSA_0 047/2	OSA_0 048/3	OSA_0 049/4	OSA_0 050/5	OSA_0 051/6	OSA_0 052/7	OSA_0 053/8	OSA_0 054/9	OSA_00 55/10
Total Sequences	32	32	32	32	32	32	32	32	32	32
Sequencing Quality	25 of 28	27 of 27	26 of 30	21 of 23	25 of 26	29 of 30	27 of 31	29 of 31	28 of 29	25 of 28
Oligo Quality	23 of 25	25 of 27	22 of 26	18 of 21	24 of 25	25 of 29	22 of 27	28 of 29	26 of 28	20 of 25
ROI Match Count	2500	2698	2561	2122	2499	2666	2625	2899	2798	2348
ROI Mutation	2	2	1	3	1	0	2	1	2	1
ROI Multi Base Deletion	0	0	0	0	0	0	0	0	0	0
ROI Small Insertion	1	0	0	0	0	0	0	0	0	0
ROI	0	0	0	0	0	0	0	0	0	0

Single Base Deletion										
Large Deletion Count	0	0	1	0	0	1	1	0	0	0
Mutation: G>A	2	2	1	2	1	0	2	1	2	1
Mutation: T>C	0	0	0	1	0	0	0	0	0	0
ROI Error Count	3	2	2	3	1	1	3	1	2	1
ROI Error Rate	Err: ~1 in 834	Err: ~1 in 1350	Err: ~1 in 1282	Err: ~1 in 708	Err: ~1 in 2500	Err: ~1 in 2667	Err: ~1 in 876	Err: ~1 in 2900	Err: ~1 in 1400	Err: ~1 in 2349
ROI Minus Primer Error Rate	MP Err: ~1 in 763	MP Err: ~1 in 824	MP Err: ~1 in 780	MP Err: ~1 in 429	MP Err: ~1 in 1525	MP Err: ~1 in 1615	MP Err: ~1 in 531	MP Err: ~1 in 1769	MP Err: ~1 in 854	MP Err: ~1 in 1451

[00227] Example 4: Design of antibody scaffolds

[00228] To generate scaffolds, structural analysis, repertoire sequencing analysis of the heavy chain, and specific analysis of heterodimer high-throughput sequencing datasets were performed. Each heavy chain was associated with each light chain scaffold. Each heavy chain scaffold was assigned 5 different long CDRH3 loop options. Each light chain scaffold was assigned 5 different L3 scaffolds. The heavy chain CDRH3 stems were chosen from the frequently observed long H3 loop stems (10 amino acids on the N-terminus and the C-terminus) found both across individuals and across V-gene segments. The light chain scaffold L3s were chosen from heterodimers comprising long H3s. Direct heterodimers based on information from the Protein Data Bank (PDB) and deep sequencing datasets were used in which CDR H1, H2, L1, L2, L3, and CDRH3 stems were fixed. The various scaffolds were then formatted for display on phage to assess for expression.

[00229] Structural Analysis

[00230] About 2,017 antibody structures were analyzed from which 22 structures with long CDRH3s of at least 25 amino acids in length were observed. The heavy chains included the following: IGHV1-69, IGHV3-30, IGHV4-49, and IGHV3-21. The light chains identified included the following: IGLV3-21, IGKV3-11, IGKV2-28, IGKV1-5, IGLV1-51, IGLV1-44, and IGKV1-13. In the analysis, four heterodimer combinations were observed multiple times including: IGHV4-59/61-IGLV3-21, IGHV3-21-IGKV2-28, IGHV1-69-IGKV3-11, and IGHV1-69-IGKV1-5. An analysis of sequences and structures identified intra-CDRH3 disulfide bonds in a few structures with packing of bulky side chains such as tyrosine in the stem providing support for long

H3 stability. Secondary structures including beta-turn-beta sheets and a “hammerhead” subdomain were also observed.

[00231] *Repertoire Analysis*

[00232] A repertoire analysis was performed on 1,083,875 IgM+/CD27-naïve B cell receptor (BCR) sequences and 1,433,011 CD27+ sequences obtained by unbiased 5'RACE from 12 healthy controls. The 12 healthy controls comprised equal numbers of male and female and were made up of 4 Caucasian, 4 Asian, and 4 Hispanic individuals. The repertoire analysis demonstrated that less than 1% of the human repertoire comprises BCRs with CDRH3s longer than 21 amino acids. A V-gene bias was observed in the long CDR3 subrepertoire, with IGHV1-69, IGHV4-34, IGHV1-18, and IGHV1-8 showing preferential enrichment in BCRs with long H3 loops. A bias against long loops was observed for IGHV3-23, IGHV4-59/61, IGHV5-51, IGHV3-48, IGHV3-53/66, IGHV3-15, IGHV3-74, IGHV3-73, IGHV3-72, and IGHV2-70. The IGHV4-34 scaffold was demonstrated to be autoreactive and had a short half-life.

[00233] Viable N-terminal and C-terminal CDRH3 scaffold variation for long loops were also designed based on the 5'RACE reference repertoire. About 81,065 CDRH3s of amino acid length 22 amino acids or greater were observed. By comparing across V-gene scaffolds, scaffold-specific H3 stem variation was avoided as to allow the scaffold diversity to be cloned into multiple scaffold references.

[00234] *Heterodimer Analysis*

[00235] Heterodimer analysis was performed on scaffolds and variant sequences and lengths of the scaffolds were assayed.

[00236] *Structural Analysis*

[00237] Structural analysis was performed using GPCR scaffolds of variant sequences and lengths were assayed.

[00238] **Example 5: Generation of GPCR antibody libraries**

[00239] Based on GPCR-ligand interaction surfaces and scaffold arrangements, libraries were designed and de novo synthesized. *See* Example 4. Referring to **FIG. 5**, 10 variant sequences were designed for the variable domain, heavy chain **503**, 237 variant sequences were designed for the heavy chain complementarity determining region 3 **507**, and 44 variant sequences were designed for the variable domain, light chain **513**. The fragments were synthesized as three fragments as seen in **FIG. 6** following similar methods as described in Examples 1-3.

[00240] Following de novo synthesis, 10 variant sequences were generated for the variable domain, heavy chain **602**, 236 variant sequences were generated for the heavy chain complementarity determining region 3 **604**, and 43 variant sequences were designed for a region

comprising the variable domain **606**, light chain and CDRL3 and of which 9 variants for variable domain, light chain were designed. This resulted in a library with about 10^5 diversity ($10 \times 236 \times 43$). This was confirmed using next generation sequencing (NGS) with 16 million reads. The normalized sequencing reads for each of the 10 variants for the variable domain, heavy chain was about 1. The normalized sequencing reads for each of the 43 variants for the variable domain, light chain was about 1. The normalized sequencing reads for 236 variant sequences for the heavy chain complementarity determining region 3 were about 1.

[00241] The various light and heavy chains were then tested for expression and protein folding. The 10 variant sequences for variable domain, heavy chain included the following: IGHV1-18, IGHV1-69, IGHV1-8 IGHV3-21, IGHV3-23, IGHV3-30/33rn, IGHV3-28, IGHV3-74, IGHV4-39, and IGHV4-59/61. Of the 10 variant sequences, IGHV1-18, IGHV1-69, and IGHV3-30/33rn exhibited improved characteristics such as improved thermostability. 9 variant sequences for variable domain, light chain included the following: IGKV1-39, IGKV1-9, IGKV2-28, IGKV3-11, IGKV3-15, IGKV3-20, IGKV4-1, IGLV1-51, and IGLV2-14. Of the 9 variant sequences, IGKV1-39, IGKV3-15, IGLV1-51, and IGLV2-14 exhibited improved characteristics such as improved thermostability.

[00242] Example 6: Expression of GPCR antibody libraries in HEK293 cells

[00243] Following generation of GPCR antibody libraries, about 47 GPCRs were selected for screening. GPCR constructs about 1.8 kb to about 4.5 kb in size were designed in a pCDNA3.1 vector. The GPCR constructs were then synthesized following similar methods as described in Examples 4-5 including hierarchal assembly. Of the 47 GPCR constructs, 46 GPCR constructs were synthesized.

[00244] The synthesized GPCR constructs were transfected in HEK293 and assayed for expression using immunofluorescence. HEK293 cells were transfected with the GPCR constructs comprising an N-terminally hemagglutinin (HA)-tagged human Y_1 receptor (data not shown). Following 24-48 hours of transfection, cells were washed with phosphate buffered saline (PBS) and fixed with 4% paraformaldehyde. Cells were stained using fluorescent primary antibody directed towards the HA tag or secondary antibodies comprising a fluorophore and DAPI to visualize the nuclei in blue. Human Y_1 receptor was visualized on the cell surface in non-permeabilized cells and on the cell surface and intracellularly in permeabilized cells (data not shown).

[00245] GPCR constructs were also visualized by designing GPCR constructs comprising auto-fluorescent proteins. Human Y_1 receptor comprised EYFP fused to its C-terminus, and human Y_5 receptor comprised ECFP fused to its C-terminus (data not shown). HEK293 cells were transfected with human Y_1 receptor or co-transfected with human Y_1 receptor and human Y_5 receptor.

Following transfection cells were washed and fixed with 4% paraformaldehyde. Cells were stained with DAPI. Localization of human Y₁ receptor and human Y₅ receptor were visualized by fluorescence microscopy.

[00246] Example 7 Design of immunoglobulin library

[00247] An immunoglobulin scaffold library was designed for placement of GPCR binding domains and for improving stability for a range of GPCR binding domain encoding sequences. The immunoglobulin scaffold included a VH domain attached with a VL domain with a linker. Variant nucleic acid sequences were generated for the framework elements and CDR elements of the VH domain and VL domain. The structure of the design is shown in **FIG. 12A**. A full domain architecture is shown in **FIG. 12B**. Sequences for the leader, linker, and pIII are listed in **Table 7**.

Table 7. Nucleotide sequences

SEQ ID NO	Domain	Sequence
6	Leader	GCAGCCGCTGGCTTGCTGCTGCTGGCAGCTCAGCCGGCCATGGCC
7	Linker	GCTAGCGGTGGAGGCGGTTCAGGCGGAGGTGGCTCTGGCGGTGGC GGATCGCATGCATCC
8	pIII	CGCGCGGCCGCTGGAAGCGGCTCCCACCATCACCATCACCAT

[00248] The VL domains that were designed include IGKV1-39, IGKV3-15, IGLV1-51, and IGLV2-14. Each of four VL domains were assembled with their respective invariant four framework elements (FW1, FW2, FW3, FW4) and variable 3 CDR (L1, L2, L3) elements. For IGKV1-39, there was 490 variants designed for L1, 420 variants designed for L2, and 824 variants designed for L3 resulting in a diversity of 1.7×10^8 ($490 * 420 * 824$). For IGKV3-15, there was 490 variants designed for L1, 265 variants designed for L2, and 907 variants designed for L3 resulting in a diversity of 1.2×10^8 ($490 * 265 * 907$). For IGLV1-51, there was 184 variants designed for L1, 151 variants designed for L2, and 824 variants designed for L3 resulting in a diversity of 2.3×10^7 ($184 * 151 * 824$). IGLV2-14, 967 variants designed for L1, 535 variants designed for L2, and 922 variants designed for L3 resulting in a diversity of 4.8×10^8 ($967 * 535 * 922$). **Table 8** lists the amino acid sequences and nucleotide sequences for the four framework elements (FW1, FW2, FW3, FW4) for IGLV1-51. **Table 9** lists the variable 3 CDR (L1, L2, L3) elements for IGLV1-51. Variant amino acid sequences and nucleotide sequences for the four framework elements (FW1, FW2, FW3, FW4) and the variable 3 CDR (L1, L2, L3) elements were also designed for IGKV1-39, IGKV3-15, and IGLV2-14.

Table 8. Sequences for IGLV1-51 framework elements

Element	SEQ ID NO	Amino Acid Sequence	SEQ ID NO	Nucleotide Sequence
IGLV1-51				
FW1	9	QSVLTQPPSVSAAPGQKVTISC	10	CAGTCTGTGTTGACGCAGCCGCCCTCAGT GTCTGCGGCCCCAGGACAGAAGGTCACCA TCTCCTGC
FW2	11	WYQQLPGTAPKLLIY	12	TGGTATCAGCAGCTCCCAGGAACAGCCCC CAAACCTCTCATTTAT
FW3	13	GIPDRFSGSKSGTSATLGITGL QTGDEADYY	14	GGGATTCTGACCGATTCTCTGGCTCCAA GTCTGGCACGTCAGCCACCCTGGGCATCA CCGGACTCCAGACTGGGGACGAGGCCGA TTATTAC
FW4	15	GGGTKLTVL	16	GGCGGAGGGACCAAGCTGACCGTCCTA

Table 9. Sequences for IGLV1-51 CDR elements

SEQ ID NO	Amino Acid Sequence	SEQ ID NO	Nucleotide Sequence
IGLV1-51-L1			
17	SGSSSNIGSNHVS	200	TCTGGAAGCAGCTCCAACATTGGGAGTAATCATGTATCC
18	SGSSSNIGNNYLS	201	TCTGGAAGCAGCTCCAACATTGGGAATAATTATCTATCC
19	SGSSSNIANNYVS	202	TCTGGAAGCAGCTCCAACATTGCGAATAATTATGTATCC
20	SGSSPNIGNNYVS	203	TCTGGAAGCAGCCCCAACATTGGGAATAATTATGTATCG
21	SGSRNIGSNYVS	204	TCTGGAAGCAGATCCAATATTGGGAGTAATTATGTTTCG
22	SGSSSNVGDNYVS	205	TCTGGAAGCAGCTCCAACGTTGGCGATAATTATGTTTCC
23	SGSSSNIGIQYVS	206	TCTGGAAGCAGCTCCAACATTGGGATTCAATATGTATCC
24	SGSSSNVGNFVS	207	TCTGGAAGCAGCTCCAATGTTGGTAACAATTTGTCTCC
25	SGSASNIGNNYVS	208	TCTGGAAGCGCCTCCAACATTGGGAATAATTATGTATCC
26	SGSGSNIGNNDVS	209	TCTGGAAGCGGCTCCAATATTGGGAATAATGATGTGTCC
27	SGSISNIGNNYVS	210	TCTGGAAGCATCTCCAACATTGGTAATAATTATGTATCC
28	SGSISNIGKNYVS	211	TCTGGAAGCATCTCCAACATTGGGAAAAATTATGTGTCC
29	SGSSSNIGHNYVS	212	TCTGGAAGCAGCTCCAACATTGGGCATAATTATGTATCG
30	PGSSSNIGNNYVS	213	CCTGGAAGCAGCTCCAACATTGGGAATAATTATGTATCC
31	SGSTSNIGIHYVS	214	TCTGGAAGCACCTCCAACATTGGAATTCATTATGTATCC
32	SGSSSNIGSHYVS	215	TCTGGAAGCAGCTCCAACATTGGCAGTCATTATGTTTCC
33	SGSSSNIGNEYVS	216	TCCGGAAGCAGCTCCAACATTGGAAATGAATATGTATCC
34	SGSTSNIGNNYIS	217	TCTGGAAGCACCTCCAACATTGGAAATAATTATATATCG
35	SGSSSNIGNHFVS	218	TCTGGAAGCAGCTCCAATATTGGGAATCATTGTGTATCG
36	SGSSSNIGNNYVA	219	TCTGGAAGCAGCTCCAACATTGGGAATAATTATGTGGCC
37	SGSSSNIGSYYVS	220	TCTGGAAGCAGCTCCAACATTGGAAGTTATTATGTATCC
38	SGSGFNIGNNYVS	221	TCTGGAAGTGGTTTCAACATTGGGAATAATTATGTCTCT
39	SGSTSNIGNNYVS	222	TCTGGAAGCACCTCCAACATTGGGAATAATTATGTGTCC
40	SGSSSDIGNNYVS	223	TCTGGAAGCAGCTCCGACATTGGCAATAATTATGTATCC
41	SGSSSNIGNNVVS	224	TCTGGAAGCAGCTCCAACATTGGGAATAATGTTGTATCC
42	SGSKSNIGKNYVS	225	TCTGGAAGCAAGTCTAACATTGGGAAAAATTATGTATCC
43	SGSSTNIGNNYVS	226	TCTGGAAGCAGACCAACATTGGGAATAATTATGTATCC
44	SGSISNIGDNYVS	227	TCTGGAAGCATCTCCAACATTGGGGATAATTATGTATCC

45	SGSSSNIGSKDVS	228	TCTGGAAGCAGCTCCAACATTGGGAGTAAGGATGTATCA
46	SGSSSNIGNNDVS	229	TCTGGAAGCAGCTCCAACATTGAGAATAATGATGTATCG
47	SGSSSNIGNHYVS	230	TCTGGAAGCAGCTCCAACATTGGGAATCATTATGTATCC
48	SGSSSNIGKDFVS	231	TCTGGAAGCAGCTCCAACATTGGGAAGGATTTTGTCTCC
49	SGSTSNIGSNFVS	232	TCTGGCAGTACTTCCAACATCGGAAGTAATTTTGTTC
50	SGSTSNIGHNYVS	233	TCTGGAAGCACCTCCAACATTGGGCATAATTATGTATCC
51	SASSSNIGNNYVS	234	TCTGCAAGCAGCTCCAACATTGGGAATAATTATGTATCC
52	SGSSSNIGNNYVS	235	TCTGGAAGCAGCTCCAGCATTGGCAATAATTATGTATCC
53	SGSSSTIGNNYVS	236	TCTGGAAGCAGCTCCACCATTGGGAATAATTATGTATCC
54	SGSSSNIGNNYVS	237	TCTGGAAGCAGCTCCAACATTGAAAATAATTATGTATCC
55	SGSSSNIGNQYVS	238	TCTGGAAGCAGCTCCAACATTGGGAATCAGTATGTATCC
56	SGSSSNIGNNYVF	239	TCTGGAAGCAGCTCCAACATTGGGAATAATTATGTATTC
57	SGSSSNIGNRYVS	240	TCTGGAAGCAGCTCCAACATTGGGAGGAATTATGTCTCC
58	SGGSSNIGNYYVS	241	TCTGGAGGCAGCTCCAACATTGGAAATTATTATGTATCG
59	SGSSSNIGDNYVS	242	TCTGGAAGCAGCTCCAACATTGGAGATAATTATGTCTCC
60	SGGSSNIGINYVS	243	TCTGGAGGCAGCTCCAACATTGGAAATTAATTATGTATCC
61	SGGSSNIGKNYVS	244	TCTGGAGGCAGCTCCAACATTGGGAAGAATTATGTATCC
62	SGSSSNIGKRSVS	245	TCTGGAAGCAGCTCCAACATTGGGAAGAGATCTGTATCG
63	SGRSNIGNNYVS	246	TCTGGAAGCAGATCCAACATTGGGAATAACTATGTATCC
64	SGSSSNIGNNLVS	247	TCGGGAAGCAGCTCCAACATTGGGAATAATCTTGTTC
65	SGSSSNIGINYVS	248	TCTGGAAGCAGCTCCAACATTGGGATCAATTATGTATCC
66	SGSSSNIGNNFVS	249	TCTGGAAGCAGCTCCAACATCGGAATAATTTTGTATCC
67	SGTSSNIGNRFVS	250	TCTGGAACCAGCTCCAACATTGGCAGAAATTTTGTATCC
68	SGRRSNIGNNYVS	251	TCTGGAAGGAGGTCCAACATTGGAAATAATTATGTGTCC
69	SGGSFNIGNNYVS	252	TCTGGAGGCAGCTTCAATATTGGGAATAATTATGTATCC
70	SGSTSNIGENYVS	253	TCTGGAAGCACTTCCAACATTGGGGAGAATTATGTGTCC
71	SGSSSNIGSDYVS	254	TCTGGAAGCAGCTCCAATATTGGGAGTGATTATGTATCC
72	SGTSSNIGSNYVS	255	TCTGGAACCAGCTCCAACATTGGGAGTAATTATGTATCC
73	SGSSSNIGTNFVS	256	TCTGGAAGCAGCTCCAACATTGGGACTAATTTTGTATCC
74	SGSSSNFGNNYVS	257	TCTGGAAGCAGCTCCAACCTTTGGGAATAATTATGTATCC
75	SGSTSNIGNNHVS	258	TCTGGAAGCACCTCCAACATTGGGAATAATCATGTATCC
76	SGSSSNIGNDFVS	259	TCTGGAAGCAGCTCCAACATTGGGAATGATTTTGTATCC
77	SGSSSDIGDNYVS	260	TCTGGAAGCAGCTCCGACATTGGCGATAATTATGTGTCC
78	SGSSSNIGKYYVS	261	TCTGGAAGCAGCTCCAACATTGGGAAATATTATGTATCC
79	SGSSSNIGGNYVS	262	TCTGGAAGCAGCTCCAACATTGGCGGTAATTATGTATCC
80	SGSSSNIGTNNYVS	263	TCTGGAAGCAGCTCCAACACTGGGAATAATTATGTATCC
81	SGSSSNVGNNYVS	264	TCTGGAAGCAGCTCCAACGTTGGGAATAATTATGTGTCT
82	SGSSSNIANNFVS	265	TCTGGAAGCAGCTCCAACATTGCGAATAATTTTGTATCC
83	SGSSSNIGNDYVS	266	TCTGGAAGCAGCTCCAACATTGGGAATGATTATGTATCC
84	SGSTSNIGNNYVS	267	TCTGGAAGCACCTCCAATATTGAGAATAATTATGTTTC
85	SGGSSNIGNNDVS	268	TCTGGAGGCAGCTCCAATATTGGCAATAATGATGTGTCC
86	SGSTSNIGNHYVS	269	TCTGGAAGCACCTCCAACATTGGGAATCATTATGTATCC
87	SGSSSNIGDNDVS	270	TCAGGAAGCAGCTCCAATATTGGGGATAATGATGTATCC
88	SGYSSNIGNNYVS	271	TCTGGATACAGCTCCAACATTGGGAATAATTATGTATCC
89	SGSGSNIGNNFVS	272	TCTGGAAGCGGCTCCAACATTGGAAATAATTTTGTATCC

90	SGSSSNIWNNYVS	273	TCTGGAAGCAGCTCCAACATTGGAATAATTATGTATCC
91	FGSSSNIGNNYVS	274	TTTGAAGCAGCTCCAACATTGGGAATAATTATGTATCC
92	SGSSSNIENYVS	275	TCTGGAAGCAGCTCCAACATTGAGAAGAATTATGTATCC
93	SGSRSNIGNYYVS	276	TCTGGAAGTAGATCCAATATTGGAAATTATTATGTATCC
94	SGTKSNIGNNYVS	277	TCTGGAACCAAGTCAAACATTGGGAATAATTATGTATCT
95	SGSTSNIGNYYVS	278	TCTGGAAGCACCTCCAACATTGGGAATTATTATGTATCC
96	SGTSSNIGNNYVA	279	TCTGGAACCAGCTCCAACATTGGGAATAATTATGTGGCC
97	PGTSSNIGNNYVS	280	CCTGGAACCAGCTCCAACATTGGGAATAATTATGTATCC
98	SGSTSNIGINYVS	281	TCCGGAAGCACCTCCAACATTGGGATTAATTATGTATCC
99	SGSSSNIGSNLVS	282	TCTGGAAGCAGCTCCAACATTGGGAGTAATCTGGTATCC
100	SGSSSNIENNHVS	283	TCTGGAAGCAGCTCCAACATTGAGAATAATCATGTATCC
101	SGTRSNIGNNYVS	284	TCTGGAACCAGGTCCAACATCGGCAATAATTATGTTTCG
102	SGSTSNIGDNYVS	285	TCTGGAAGCACCTCCAACATTGGGGACAATTATGTTTCC
103	SGGSSNIGKNFVS	286	TCTGGAGGCAGTTCCAACATTGGGAAGAATTTTGTATCC
104	SGSRSDIGNNYVS	287	TCTGGAAGCAGGTCCGACATTGGGAATAATTATGTATCC
105	SGTSSNIGNNDVS	288	TCTGGAAGTAGCTCCAACATTGGGAATAATGATGTATCC
106	SGSSSNIGSKYVS	289	TCTGGAAGCAGCTCCAACATTGGGAGTAAATATGTATCA
107	SGSSFNIGNNYVS	290	TCTGGAAGCAGCTTCAACATTGGGAATAATTATGTATCC
108	SGSSSNIGNTYVS	291	TCTGGAAGCAGCTCCAACATTGGGAATACTTATGTATCC
109	SGSSSNIGDNHVS	292	TCTGGAAGCAGCTCCAATATTGGGGATAATCATGTATCC
110	SGSSSNIGNNHVS	293	TCTGGAAGCAGCTCCAACATTGGCAATAATCATGTTTCC
111	SGSTSNIGNNDVS	294	TCTGGAAGCACCTCCAACATTGGGAATAATGATGTATCC
112	SGSRSNVGNYYVS	295	TCTGGAAGCAGATCCAACGTTGGCAATAATTATGTTTCA
113	SGGTSNIGKNYVS	296	TCCGGAGGCACCTCCAACATTGGGAAGAATTATGTGTCT
114	SGSSSNIADNYVS	297	TCTGGAAGCAGCTCCAACATTGCCGATAATTATGTTTCC
115	SGSSSNIGANYVS	298	TCTGGAAGCAGCTCCAACATTGGCGCCAATTATGTATCC
116	SGSSSNIGSNYVA	299	TCTGGAAGCAGCTCCAACATTGGGAGTAATTATGTGGCC
117	SGSSSNIGNNFLS	300	TCTGGAAGCAGCTCCAACATTGGGAACAATTTCTCTCC
118	SGRSSNIGKNYVS	301	TCTGGAAGAAGCTCCAACATTGGGAAGAATTATGTATCC
119	SGSSPNIGANYVS	302	TCTGGAAGCAGCCCCAACATTGGGGCTAATTATGTATCC
120	SGSSSNIGPNYVS	303	TCCGGAAGCAGCTCCAACATTGGGCCTAATTATGTGTCC
121	SGSSSTIGNNYIS	304	TCTGGAAGCAGCTCCACCATTGGGAATAATTATATATCC
122	SGSSSNIGNYFVS	305	TCTGGAAGCAGCTCCAACATTGGGAATTATTTTGTATCC
123	SGSRSNIGNNFVS	306	TCTGGAAGCCGCTCCAACATTGGTAATAATTTTGTATCC
124	SGGSSNIGSNFVS	307	TCTGGAGGCAGCTCCAACATTGGGAGTAATTTTGTATCC
125	SGSSSNIGYNYVS	308	TCTGGAAGCAGCTCCAACATTGGGTATAATTATGTATCC
126	SGTSSNIENNYVS	309	TCTGGAACCAGCTCGAACATTGAGAACAATTATGTATCC
127	SGSSSNIGNYYVS	310	TCTGGAAGTAGCTCCAACATTGGGAATTATTATGTATCC
128	SGSTSNIGKNYVS	311	TCTGGAAGCACCTCCAACATTGGGAAGAATTATGTATCC
129	SGSSSNIGTYYYVS	312	TCTGGAAGCAGTTCCAACATTGGGACTTATTATGTCTCT
130	SGSSSNVGNKYVS	313	TCTGGAAGCAGCTCCAACGTTGGGAAAAATTATGTATCT
131	SGSTSNIGDNFVS	314	TCTGGAAGCACCTCCAACATTGGGGATAATTTTGTATCC
132	SGSTSNIGTNYVS	315	TCTGGAAGCACCTCCAACATTGGAACTAATTATGTTTCC
133	SGGTSNIGNNYVS	316	TCTGGAGGTACTTCCAACATTGGGAATAATTATGTCTCC
134	SGSYSNIGNNYVS	317	TCTGGAAGCTACTCCAATATTGGGAATAATTATGTATCC

135	SGSSSNIEDNYVS	318	TCTGGAAGCAGCTCCAACATTGAAGATAATTATGTATCC
136	SGSSSNIGKHYVS	319	TCTGGAAGCAGCTCCAACATTGGGAAACATTATGTATCC
137	SGSGSNIGSNYVS	320	TCCGGTTCCGGCTCAAACATTGGAAGTAATTATGTCTCC
138	SGSSSNIGNNYIS	321	TCTGGAAGCAGCTCCAACATTGGAAATAATTATATATCA
139	SGASSNIGNNYVS	322	TCTGGAGCCAGTTCCAACATTGGGAATAATTATGTTTCC
140	SGRTSNIGNNYVS	323	TCTGGACGCACCTCCAACATCGGGAACAATTATGTATCC
141	SGGSSNIGSNYVS	324	TCTGGAGGCAGCTCCAATATTGGGAGTAATTACGTATCC
142	SGSGSNIGNNYVS	325	TCTGGAAGCGGCTCCAACATTGGGAATAATTATGTATCC
143	SGSTSNIGSNYVS	326	TCTGGAAGCACCTCCAACATTGGGAGTAATTATGTATCC
144	SGSSSIGNNYVA	327	TCTGGAAGCAGCTCCAGCATTGGGAATAATTATGTGGCG
145	SGSSSNLGNMYVS	328	TCTGGAAGCAGTTCCAACCTTGGAAATAATTATGTATCC
146	SGTSSNIGKNYVS	329	TCTGGAACCAGCTCCAACATTGGGAAAAATTATGTATCC
147	SGSSSDIGNKYIS	330	TCTGGAAGCAGCTCCGATATTGGGAACAAGTATATATCC
148	SGSSSNIGSNYIS	331	TCTGGAAGCAGCTCCAACATTGGAAGTAATTACATATCC
149	SGSTSNIGANYVS	332	TCTGGAAGCACCTCCAACATTGGGGCTAACTATGTGTCC
150	SGSSSNIGNKYVS	333	TCTGGAAGCAGCTCCAACATTGGGAATAAGTATGTATCC
151	SGSSSNIGNNYGS	334	TCTGGAAGCAGCTCCAACATTGGGAATAATTATGGATCC
152	SGSTSNIANMYVS	335	TCTGGAAGCACCTCCAACATTGCGAATAATTATGTATCC
153	SGSYSNIGSNYVS	336	TCTGGAAGCTACTCCAATATTGGGAGTAATTATGTATCC
154	SGSSSNIGSNFVS	337	TCTGGAAGCAGCTCCAACATTGGGAGTAATTTGTATCC
155	SGSSSNLENNYVS	338	TCTGGAAGCAGCTCCAATCTTGAGAATAATTATGTATCC
156	SGSISNIGSNYVS	339	TCTGGAAGCATCTCCAATATTGGCAGTAATTATGTATCC
157	SGSSSDIGSNYVS	340	TCTGGAAGCAGCTCCGACATTGGGAGTAATTATGTATCC
158	SGSSSNIGTNYVS	341	TCTGGAAGCAGCTCCAACATTGGGACTAATTATGTATCC
159	SGSSSNIGKNFVS	342	TCTGGAAGCAGCTCCAACATTGGGAAGAATTTGTATCC
160	SGSSSNIGNNFIS	343	TCTGGAAGCAGCTCCAACATTGGGAATAATTTTATATCC
161	SGGSSNIGNNYVS	344	TCTGGAGGCAGCTCCAACATTGGCAATAATTATGTTTCC
162	SGSSSNIGENYVS	345	TCTGGAAGCAGCTCCAACATTGGGGAGAATTATGTATCC
163	SGSSSNIGNNFVA	346	TCTGGAAGCAGCTCCAATATTGGGAATAATTTTGTGGCC
164	SGGSSNIGNNYVA	347	TCTGGAGGCAGCTCCAACATTGGGAATAATTATGTAGCC
165	SGSSSHIGNNYVS	348	TCTGGAAGCAGCTCCACATTGGAAATAATTATGTATCC
166	SGSSSNIGSNDVS	349	TCTGGAAGCAGCTCCAATATTGGAAGTAATGATGTATCG
167	SGSSSNIGNNYVT	350	TCTGGAAGCAGCTCCAACATTGGGAATAATTATGTAACC
168	SGSSSNIGNNPVS	351	TCTGGAAGCAGCTCCAACATTGGGAATAATCCTGTATCC
169	SGGSSNIGNHYVS	352	TCTGGAGGCAGCTCCAATATTGGGAATCATTATGTATCC
170	SGTSSNIGNNYVS	353	TCTGGAACCAGCTCCAACATTGGGAATAATTATGTATCC
171	SGSSSNIGSNYVS	354	TCTGGAAGCAGCTCCAACATTGGAAGTAATTATGTCTCG
172	SGGTSNIGSNYVS	355	TCTGGAGGCACCTCCAACATTGGAAGTAATTATGTATCC
173	SGSKSNIGNNYVS	356	TCTGGAAGCAAGTCCAACATTGGGAATAATTATGTATCC
174	SGRSSNIGNNYVS	357	TCTGGAAGAAGCTCCAACATTGGGAATAATTATGTATCG
175	SGSSSNVGSNYVS	358	TCTGGAAGCAGCTCCAACGTTGGGAGTAATTATGTTTCC
176	SGSTSNIGNNFVS	359	TCTGGAAGCACCTCCAATATTGGGAATAATTTGTATCC
177	SGSNFNIGNNYVS	360	TCTGGAAGCAACTTCAACATTGGGAATAATTATGTCTCC
178	SGSTSNIGYNYVS	361	TCTGGAAGCACCTCCAATATTGGATATAATTATGTATCC
179	SGSSSNIVSNYVS	362	TCTGGAAGCAGCTCCAATATTGTAAGTAATTATGTATCC

180	SGTSSNIGNNFVS	363	TCTGGAACCAGCTCCAACATTGGGAATAATTTGTATCC
181	SGSSSNIGRNFVS	364	TCTGGAAGCAGCTCCAACATTGGGAGGAATTTGTGTCC
182	SGTTSNIGNNYVS	365	TCTGGAACGACCTCCAACATTGGGAATAATTATGTCTCC
183	SGSSSNIGNNDVS	366	TCTGGAAGCAGCTCCAACATTGGGAATAATGATGTATCC
184	SGSSSNIGNHDVS	367	TCTGGAAGCAGCTCCAACATTGGGAATCATGATGTATCC
185	SGSSSNIGSSHVS	368	TCTGGAAGCAGCTCCAACATTGGAAGTAGTCATGTATCC
186	SGSSSNIGIHYVS	369	TCTGGAAGCAGCTCCAACATTGGGATTTCATTATGTATCC
187	SGGGSNIGYNYVS	370	TCTGGAGGCGGCTCCAACATTGGCTATAATTATGTCTCC
188	SGSSSNIGDHYVS	371	TCTGGAAGCAGCTCCAACATTGGGGATCATTATGTGTCC
189	SGSSSNLGNKYVS	372	TCTGGAAGCAGCTCCAACCTTGGGAAGAATTATGTATCT
190	SGSSSNIGDNFVS	373	TCTGGAAGCAGCTCCAACATTGGCGATAATTTGTATCC
191	SGSTSNIENKYVS	374	TCTGGAAGCACCTCCAACATTGAGAAAACTATGTATCC
192	SGSSSNIGKDYVS	375	TCTGGAAGCAGCTCCAACATTGGGAAGGATTATGTATCC
193	SGSSSNIGKNYVS	376	TCTGGAAGCAGCTCCAACATTGGGAAGAATTATGTATCC
194	SGSSSNIGNNYVS	377	TCTGGAAGCAGCTCCAACATTGGGAATAATTATGTATCC
195	SGSSSNIGNNYAS	378	TCTGGAAGCAGCTCCAACATTGGGAATAATTATGCCTCC
196	SGISSNIGNNYVS	379	TCTGGAATCAGCTCCAACATTGGGAATAATTATGTATCC
197	TGSSSNIGNNYVS	380	ACTGGAAGCAGCTCCAACATTGGGAATAATTATGTATCC
198	SGTSSNIGNNHVS	381	TCTGGAACCAGCTCCAACATTGGGAATAATCATGTTTCC
199	SGSRSNIGKNYVS	382	TCTGGAAGTCGTTCCAACATTGGGAAAAATTATGTATCC
IGLV1-51-L2			
383	DNNKRPP	534	GACAATAATAAGCGACCCCCA
384	ENNRPS	535	GAGAATAATAGCGACCCTCA
385	DNNKQPS	536	GACAATAATAAGCAACCCTCA
386	DNNKRPL	537	GACAATAACAAGCGACCCTTG
387	DNDKRPA	538	GACAATGATAAGCGACCCGCA
388	DNHERPS	539	GACAATCATGAGCGACCCTCA
389	ENRKRPS	540	GAAAACCGTAAGCGACCCTCA
390	DNDQRPS	541	GACAATGATCAGCGACCCTCA
391	ENYKRPS	542	GAGAATTATAAGCGACCCTCA
392	ENTKRPS	543	GAAAATACTAAGCGACCCTCA
393	DTEKRPS	544	GACACTGAGAAGAGGCCCTCA
394	DNDKRPP	545	GACAATGATAAGCGACCCCCA
395	DHNRKPS	546	GACCATAATAAGCGACCCTCA
396	GNNRPS	547	GGCAATAATGAGCGACCCTCA
397	DTSKRPS	548	GACACTAGTAAGCGACCCTCA
398	EYNKRPS	549	GAATATAATAAGCGCCCCTCA
399	ENIKRPS	550	GAAAATATTAAGCGACCCTCA
400	DNVKRPS	551	GACAATGTTAAGCGACCCTCA
401	ENDKRSS	552	GAAAACGATAAACGATCCTCA
402	ENNRKHS	553	GAAAATAATAAGCGGACTCA
403	GNDQRPS	554	GGAAATGATCAGCGACCCTCA
404	DNDRRPS	555	GACAATGATAGCGACCCTCA
405	DNHKRPS	556	GACAATCATAAGCGGCCCTCA
406	DNNDRPS	557	GACAATAATGACCGACCCTCA

407	ENNQRPS	558	GAGAATAATCAGCGACCCTCA
408	DNNQRPS	559	GACAATAATCAGCGACCCTCA
409	ENVKRPS	560	GAGAATGTAAAGCGACCCTCA
410	DTYKRPS	561	GACACTTATAAGAGACCCTCA
411	NNNNRPS	562	AACAATAATAACCGACCCTCA
412	GNNNRPS	563	GGCAATAATAATCGACCCTCA
413	ENDQRPS	564	GAAAATGATCAGCGACCCTCA
414	DNNKRAS	565	GACAATAATAAGCGAGCCTCA
415	DNDKRPL	566	GACAATGATAAGCGACCCTTA
416	DTDERPS	567	GACACTGATGAGCGACCTTCA
417	DNRKRPS	568	GACAATAGGAAGCGACCCTCA
418	DNDARPS	569	GACAATGATGCTCGACCCTCA
419	DNNKRLS	570	GACAATAATAAGCGACTCTCA
420	DNDKRAS	571	GACAATGATAAGCGAGCCTCA
421	DNTERPS	572	GACAATACTGAGCGACCCTCA
422	DNNIRPS	573	GACAATAATATTTCGACCCTCA
423	DNKRRPS	574	GACAATAAGAGGCGACCCTCA
424	DDNNRPS	575	GACGATAATAACCGACCCTCA
425	ANNRRPS	576	GCGAATAATCGACGACCCTCA
426	DNDKRLS	577	GACAATGATAAGCGACTGTCA
427	DNNKRPA	578	GACAATAATAAGCGACCCGCA
428	DNYRRPS	579	GACAATTATAGACGTCCCTCA
429	ANDQRPS	580	GCCAATGATCAGCGACCCTCA
430	DNDKRRS	581	GACAATGATAAGCGACGCTCA
431	DKNERPS	582	GACAAGAATGAGCGACCCTCA
432	DNKERPS	583	GACAATAAGGAGCGACCCTCA
433	DNNKGPS	584	GACAATAATAAGGGACCCTCA
434	ENDRRPS	585	GAAAATGATAGACGACCCTCA
435	ENDERPS	586	GAAAATGATGAGCGACCCTCA
436	QNNKRPS	587	CAAAATAATAAGCGACCCTCA
437	DNRERPS	588	GACAATCGTGAGCGACCCTCA
438	DNNRRPS	589	GACAATAATAGACGACCCTCA
439	GNNRRPS	590	GGAAATAATAGGCGACCCTCA
440	DNDNRPS	591	GACAATGATAACCGACCCTCA
441	EDNKRPS	592	GAAGATAATAAGCGACCCTCA
442	DDDERPS	593	GACGATGATGAGCGGCCCTCA
443	ASNKRPS	594	GCAAGTAATAAGCGACCCTCA
444	DNNKRSS	595	GACAATAATAAGCGATCCTCA
445	QNNERPS	596	CAAAATAATGAGCGACCCTCA
446	DDDRRPS	597	GACGATGATAGGCGACCCTCA
447	NNDKRPS	598	AACAATGATAAGCGACCCTCA
448	DNNNRPS	599	GACAATAATAACCGACCCTCA
449	DNNVRPS	600	GACAATAATGTGCGACCCTCA
450	ENNERPS	601	GAAAATAATGAGCGACCCTCA
451	DNNHRPS	602	GACAATAATCACCGACCCTCA

452	DNDRPS	603	GACAATGATGAGCGCCCCTCG
453	DNIRRPS	604	GACAATATCCGGCGACCCTCA
454	DFNKRPS	605	GACTTTAATAAGCGACCCTCA
455	ETNKRPS	606	GAAACTAATAAGCGACCCTCA
456	NDNKRPS	607	AACGATAATAAGCGACCCTCA
457	DDNKRPS	608	GACGATAATAAGCGACCCTCA
458	DNYKRPS	609	GACAATTATAAGCGACCCTCA
459	HNNKRPS	610	CACAATAATAAGCGACCCTCA
460	DNHQRPS	611	GACAATCATCAGCGACCCTCA
461	DNYKRAS	612	GACAATTATAAGCGAGCCTCA
462	DNIKRPS	613	GACAATATTAAGCGACCCTCA
463	DTHKRPS	614	GCACTCATAAGCGACCCTCA
464	DTNRRPS	615	GACACTAATAGGCGACCCTCT
465	DTNQRPS	616	GACACTAATCAGCGACCCTCA
466	ESDKRPS	617	GAAAGTGATAAGCGACCCTCA
467	DNDKRSS	618	GACAATGATAAGCGATCTTCG
468	GSNKRPS	619	GGCAGTAATAAGCGACCCTCA
469	DNNKRVS	620	GACAATAACAAGCGAGTTTCA
470	NNNRRPS	621	AACAATAATAGGCGACCCTCA
471	DNFKRPS	622	GACAATTTTAAGCGACCCTCA
472	ENDKRPS	623	GAAAATGATAAACGACCCTCA
473	ENNKRLS	624	GAAAATAATAAGCGACTCTCA
474	ADNKRPS	625	GCAGATAATAAGCGACCCTCA
475	EDNERPS	626	GAAGATAATGAGCGCCCCTCA
476	DTDQRPS	627	GCACTGATCAGCGACCCTCA
477	DNYQRPS	628	GACAATTATCAGCGACCCTCA
478	DENKRPS	629	GACGAGAATAAGCGACCCTCA
479	DTNKRPS	630	GACACTAATAAGCGACCCTCA
480	DDYRRPS	631	GACGATTATCGGCGACCCTCA
481	DNDKRHS	632	GACAACGATAAGCGGCACTCA
482	ENDNRPS	633	GAAAATGATAATCGACCCTCA
483	DDNERPS	634	GACGATAATGAGCGCCCCTCA
484	DNKKRPS	635	GACAATAAGAAGCGACCCTCA
485	DVDKRPS	636	GACGTTGATAAGCGACCCTCA
486	ENKKRPS	637	GAAAATAAAAAACGACCCTCT
487	VNDKRPS	638	GTCAATGATAAGCGACCCTCA
488	DNDHRPS	639	GACAATGATCACCGACCCTCA
489	DINKRPS	640	GACATTAATAAGCGACCCTCA
490	ANNERPS	641	GCCAATAATGAGCGACCCTCA
491	DNENRPS	642	GACAATGAAAACCGACCGTCA
492	GDDKRPS	643	GGCGATGATAAGCGACCCTCA
493	ANNQRPS	644	GCCAATAATCAGCGACCTTCA
494	DDDKRPS	645	GACGATGATAAGCGACCCTCA
495	YNNKRPS	646	TACAATAATAAGCGGCCCTCA
496	EDDKRPS	647	GAAGATGATAAGCGACCCTCA

497	ENNNRPS	648	GAAAACAATAACCGACCCTCG
498	DNNLRPS	649	GACAATAATCTGCGACCCTCA
499	ESNKRPS	650	GAGAGTAACAAGCGACCCTCA
500	DTDKRPS	651	GACACTGATAAGCGGCCCTCA
501	DDDQRPS	652	GACGATGATCAGCGACCCTCA
502	VNNKRPS	653	GTGAATAATAAGAGACCCTCC
503	DDYKRPS	654	GACGATTATAAGCGACCCTCA
504	DNTKRPS	655	GACAATACTAAGCGACCCTCA
505	DDTERPS	656	GACGATACTGAGCGACCCTCA
506	GNDKRPS	657	GGCAATGATAAGCGACCCTCA
507	DNEKRPS	658	GACAATGAAAAGCGACCCTCA
508	DNDDRPS	659	GACAATGATGACCGACCCTCA
509	DDNRRPS	660	GACGATAATAGGCGTCCCTCA
510	GNNKRPS	661	GGCAATAATAAGCGACCCTCA
511	ANDKRPS	662	GCCAATGATAAGCGACCCTCA
512	DNNKRHS	663	GACAATAATAAGCGACACTCA
513	DDNQRPS	664	GACGACAATCAGCGACCCTCA
514	GNDRRPS	665	GGCAATGATAGGCGACCCTCA
515	DNHNRPS	666	GACAATCATAACCGACCCTCA
516	DNYERPS	667	GACAATTATGAGCGACCCTCA
517	ENNKRSS	668	GAAAATAATAAGCGATCCTCA
518	DDHKRPS	669	GACGATCATAAGCGGCCCTCA
519	DNNKRRS	670	GACAATAATAAACGACGTTCA
520	DNDKRPS	671	GACAATGATAAGCGACCGTCA
521	DKNKRPS	672	GACAAGAATAAGCGACCCTCA
522	DNNKRPS	673	GACAATAATAAGCGACCCTCA
523	DIDKRPS	674	GACATTGATAAGCGACCCTCA
524	DDKKRPS	675	GACGATAAGAAGCGACCCTCA
525	ANNKRPS	676	GCCAATAATAAGCGACCCTCA
526	DNDKGPS	677	GACAATGATAAGGGACCCTCA
527	EDNRRPS	678	GAAGATAATAGGCGACCCTCA
528	ENNKRPS	679	GAGAATAATAAGCGACCCTCA
529	NNNKRPS	680	AACAATAATAAGCGACCCTCA
530	DNNERPS	681	GACAATAATGAGCGACCCTCA
531	DNIQRPS	682	GACAATATTGAGCGACCCTCA
532	DNNYRPS	683	GACAATAATTACCGACCCTCA
533	DNYNRPS	684	GACAATTATAACCGACCCTCA
IGLV1-51-L3			
685	CGTWDTSLSAVVF	1509	TGCGGAACATGGGATACCAGCCTGAGTGCTGTGGTGTTC
686	CGTWDTSLSAGVF	1510	TGCGGAACATGGGATACCAGCCTGAGTGCTGGGGTGTTC
687	CGTWDTSLSAWVF	1511	TGCGGAACATGGGATACCAGCCTGAGTGCTTGGGTGTTC
688	CGTWDRSLSAGVF	1512	TGCGGAACATGGGATAGGAGCCTGAGTGCGGGGGTGTTC
689	CGTWDRSLSAWVF	1513	TGCGGAACATGGGATAGGAGCCTGAGTGCTTGGGTATTT
690	CGTWDTSLSGGVF	1514	TGCGGAACATGGGATACCAGCCTGAGTGCTGGGGTGTTC
691	CGTWDTSLRAGVF	1515	TGCGGAACATGGGATACTAGCCTGCGTGCTGGCGTCTTC

692	CGTWDRSLSVWVF	1516	TGCGGAACATGGGATAGGAGCCTGAGTGTTTGGGTGTTC
693	CGTWDTSLSVVVF	1517	TGCGGAACATGGGATACCAGTCTGAGTGTTGTGGTCTTC
694	CGTWDTSLSAAVF	1518	TGCGGAACGTGGGATACCAGCCTGAGTGCTGCGGTGTTC
695	CGAWDTSLSAGVF	1519	TGCGGAGCATGGGATACCAGCCTGAGTGCTGGAGTGTTC
696	CATWDTSLSAVVF	1520	TGCGCAACATGGGATACCAGCCTGAGTGCTGTGGTATTC
697	CATWDTSLSAGVF	1521	TGCGCAACATGGGATACCAGCCTGAGTGCTGGTGTGTTC
698	CGTWESSLSAWVF	1522	TGTGGAACATGGGAGAGCAGCCTGAGTGCTTGGGTGTTC
699	CGTWDTTSLAGVF	1523	TGCGGAACATGGGATACCACCCTGAGTGCGGGTGTCTTC
700	CGTWDTSLSVWVF	1524	TGCGGAACATGGGATACTAGCCTGAGTGTTGTGGGTGTTC
701	CGTWDTSLSVG VF	1525	TGCGGAACATGGGATACTAGCCTGAGTGTTGGGGTGTTC
702	CGTWDTSLSTGVF	1526	TGCGGAACATGGGACACCAGTCTGAGCACTGGCGTCTTC
703	CGTWDTSLSGVVF	1527	TGCGGAACATGGGATACCAGCCTGAGTGGTGTGGTCTTC
704	CGTWDTSLSAYVF	1528	TGCGGAACATGGGATACCAGCCTGAGTGCTTATGTCTTC
705	CGTWDTSLSAEVF	1529	TGCGGAACATGGGATACCAGCCTGAGTGCTGAGGTGTTC
706	CGTWDTGLSAGVF	1530	TGCGGAACATGGGATACCGGCCTGAGTGCTGGGGTATTC
707	CGTWDRSLSAYVF	1531	TGCGGAACGTGGGATAGGAGCCTGAGTGCTTATGTCTTC
708	CGTWDRSLSAVVF	1532	TGCGGAACATGGGATAGGAGCCTCAGTGCCGTGGTATTC
709	CGTWDNTLSAWVF	1533	TGCGGAACATGGGATAACACCCTGAGTGCGTGGGTGTTC
710	CGTWDNRLSAGVF	1534	TGCGGAACATGGGATAACAGGCTGAGTGCTGGGGTGTTC
711	CGTWDISLSAWVF	1535	TGCGGAACATGGGACATCAGCCTGAGTGCTTGGGTGTTC
712	CGTWHSSLSAGVF	1536	TGCGGAACATGGCATAGCAGCCTGAGTGCTGGGGTATTC
713	CGTWGSSLSAWVF	1537	TGCGGAACATGGGGTAGCAGTTTGAGTGCTTGGGTGTTC
714	CGTWESSLSGWVF	1538	TGCGGAACATGGGAGAGCAGCCTGAGTGCTTGGGTGTTC
715	CGTWESSLSAVVF	1539	TGCGGAACATGGGAGAGCAGCCTGAGTGCTGTGGTTTTTC
716	CGTWDYSLSAVVF	1540	TGCGGAACATGGGATTACAGCCTGAGTGCTGTGGTATTC
717	CGTWDYSLSAGVF	1541	TGCGGAACATGGGATTACAGCCTGAGTGCTGGGGTATTC
718	CGTWDVSLSVG VF	1542	TGCGGAACATGGGATGTCAGCCTGAGTGTTGGAGTGTTC
719	CGTWDTTSLAVVF	1543	TGCGGAACATGGGATACCACCCTGAGTGCTGTGGTTTTTC
720	CGTWDTTLNIGVF	1544	TGCGGAACATGGGATACCACTCTGAATATTGGGGTGTTC
721	CGTWDTSLTAVVF	1545	TGCGGAACATGGGATACCAGCCTGACTGCTGTGGTATTC
722	CGTWDTSLTAAVF	1546	TGCGGAACCTGGGATACCAGCCTGACTGCTGTGTGTTC
723	CGTWDTSLSVGLF	1547	TGCGGCACATGGGATACCAGCCTGAGTGTTGGGGCTATTC
724	CGTWDTSLSGRVF	1548	TGCGGAACCTGGGATACCAGCCTGAGTGGTAGGGTGTTC
725	CGTWDTSLSGAVF	1549	TGCGGAACATGGGATACCAGCCTGAGTGGTGCAGTGTTC
726	CGTWDTSLSAGLF	1550	TGCGGAACATGGGATACCAGCCTGAGTGCTGGCCTGTTC
727	CGTWDTSLSAGGVF	1551	TGCGGAACATGGGATACCAGCCTGAGTGCTGGAGGGGTCTTC
728	CGTWDTSLRAYVF	1552	TGCGGAACATGGGATACCAGCCTGCGTGCTTATGTCTTC
729	CGTWDTSLRAWVF	1553	TGCGGAACATGGGATACTAGTTTGCGTGCTTGGGTATTC
730	CGTWDTSLNTGVF	1554	TGCGGAACATGGGATACCAGCCTGAATACTGGGGTATTC
731	CGTWDTSLNIWVF	1555	TGCGGAACATGGGATACCAGCCTGAATATTTGGGTGTTC
732	CGTWDTSLNIGVF	1556	TGCGGAACATGGGATACAAGCCTGAATATTGGGGTGTTC
733	CGTWDTSLIAVVF	1557	TGCGGAACATGGGATACCAGCCTGATTGCTGTGGTGTTC
734	CGTWDRSLSGWVF	1558	TGCGGAACGTGGGATAGGAGCCTGAGTGCTTGGGTGTTC
735	CGTWDNRLSGWVF	1559	TGCGGAACATGGGATAACAGGCTGAGTGCTTGGGTGTTC
736	CGTWDKSLSAVVF	1560	TGCGGAACGTGGGATAAGAGCCTGAGTGCTGTGGTCTTC

737	CGTWDKGLSAWVF	1561	TGCGGAACATGGGATAAAGGCCTGAGTGCTTGGGTGTTC
738	CGTWDISLSAGVF	1562	TGCGGAACATGGGATATCAGCCTGAGTGCTGGGGTGTTC
739	CGTWDESLSGEVVF	1563	TGCGGAACATGGGATGAGAGCCTGAGTGGTGGCGAGGTGGTCTTC
740	CGTWDASLSAWVF	1564	TGCGGAACATGGGATGCCAGCCTGAGTGCTTGGGTGTTC
741	CGTWDAGLSAWVF	1565	TGCGGAACATGGGATGCCGGCCTGAGTGCTTGGGTGTTC
742	CGAWDTSLSAWVF	1566	TGCGGAGCATGGGATACCAGCCTGAGTGCTTGGGTGTTC
743	CGAWDTSLSAVVF	1567	TGCGGAGCATGGGATACCAGCCTGAGTGCTGTGGTGTTC
744	CGAWDTSLRAGVF	1568	TGCGGAGCATGGGATACCAGCCTGCGTGCTGGGGTTTTTC
745	CATWDTSVSAWVF	1569	TGCGCAACATGGGATACCAGCGTGAGTGCTTGGGTGTTC
746	CATWDTSLSAWVF	1570	TGCGCAACATGGGATACCAGCCTGAGTGCGTGGGTGTTC
747	CATWDNTLSAGVF	1571	TGCGCAACATGGGACAACACCCTGAGTGCTGGGGTGTTC
748	CAAWDRSLSVWVF	1572	TGCGCAGCATGGGATAGGAGCCTGAGTGTTTGGGTGTTC
749	CYTWSSSLRGGVF	1573	TGCTACACATGGCATTCCAGTCTGCGTGGTGGGGTGTTC
750	CVTWTSSPSAWVF	1574	TGCGTAACGTGGACTAGTAGCCCGAGTGCTTGGGTGTTC
751	CVTWRGGLVLF	1575	TGCGTGACATGGCGTGGTGGCCTTGTTGTGTTC
752	CVTWDTSLSVVL	1576	TGCGTAACATGGGATACCAGCCTGACTTCTGTGGTACTC
753	CVTWDTSLSVYWVF	1577	TGCGTAACATGGGATACCAGCCTGAGTGTTTATTGGGTGTTC
754	CVTWDTSLSAWVF	1578	TGCGTTACATGGGATACCAGCCTGAGTGCTTGGGTGTTC
755	CVTWDTDLSVALF	1579	TGCGTCACATGGGATACCGACCTCAGCGTTGCGCTCTTC
756	CVTWDRSLSGWVF	1580	TGCGTAACATGGGATAGGAGCCTGAGTGTTTGGGTGTTC
757	CVTWDRSLREVLF	1581	TGCGTAACATGGGATCGCAGCCTGAGAGAGGTGTTATTC
758	CVTWDRSLRAVVF	1582	TGCGTAACATGGGATCGCAGCCTGAGAGCGGTGGTATTC
759	CVTWDRSLDAGVF	1583	TGCGTAACATGGGACAGGAGCCTCGATGCTGGGGTTTTTC
760	CVTWDNTLSAGVF	1584	TGCGTGACATGGGATAACACCCTGAGTGCTGGGGTCTTC
761	CVTWDNNLFGVVF	1585	TGCGTAACATGGGATAACAACCTGTTTGGTGTGGTCTTC
762	CVSWDTSLSGAVF	1586	TGCGTATCATGGGATACCAGCCTGAGTGGTGGCGTATTC
763	CVSWDTSLSAGVF	1587	TGCGTCTCATGGGATACCAGCCTGAGTGCTGGGGTATTC
764	CTTWFRTPSDVVF	1588	TGCACAACATGGTTTAGGACTCCGAGTGATGTGGTCTTC
765	CTTWFRASDVVF	1589	TGCACAACATGGTTTAGGACTGCGAGTGATGTGGTCTTC
766	CTTWYGLSVVF	1590	TGCACAACGTGGGATTACGGTCTGAGTGTCGTCTTC
767	CTARDTSLSPGGVF	1591	TGCACAGCAAGGGATACCAGCCTGAGTCTTGGCGGGTCTTC
768	CSTWNTRPSDVVF	1592	TGCTCAACATGGAATACGAGGCCGAGTGATGTGGTGTTC
769	CSTWESSLTTVVF	1593	TGTTCAACATGGGAGAGCAGTTTGACTACTGTGGTCTTC
770	CSTWDTSLTNVLF	1594	TGCTCAACATGGGATACCAGCCTCATAATGTGCTATTC
771	CSTWDTSLSGVVF	1595	TGCTCAACATGGGATACCAGCCTGAGTGGAGTAGTCTTC
772	CSTWDHSLKAALF	1596	TGCTCAACATGGGATCACAGCCTGAAAGCTGCACTGTTC
773	CSTWDARLSVRVF	1597	TGCTCAACCTGGGATGCGAGGCTGAGTGTCGGGGTGTTC
774	CSSYTSSSTWVF	1598	TGCTCCTCATATACAAGCAGCAGCACTTGGGTGTTC
775	CSSYATRGLRVLF	1599	TGCAGCTCATACGCAACCCGCGGCCTTCGTGTGTGTTC
776	CSSWDATLSVRIF	1600	TGTTTCATCATGGGACGCCACCCTGAGTGTTTCGCATATTC
777	CQVWEGSSDHVVF	1601	TGTCAGGTGTGGGAGGGTAGTAGTGATCATTGGGTGTTC
778	CQTWDNRLSAVVF	1602	TGCCAAACCTGGGATAACAGACTGAGTGCTGTGGTGTTC
779	CQTWDHSLHVGVF	1603	TGTCAAACGTGGGATCACAGCCTGCATGTTGGGGTGTTC
780	CQSYDDILNVVVL	1604	TGCCAGTCTATGACGACATCTTGAATGTTGGGTCCTT
781	CNTWDKSLTSELF	1605	TGCAATACATGGGATAAGAGTTTGACTTCTGAACTCTTC

782	CLTWDRSLNVRVF	1606	TGCTTAACATGGGATCGCAGCCTGAATGTGAGGGTGTTTC
783	CLTWDHSLTAYVF	1607	TGCCTAACATGGGACCACAGCCTGACTGCTTATGTCTTC
784	CLTRDTSLSAPVF	1608	TGCTTAACAAGGGATACCAGTCTGAGTGCCCCGTGTGTTTC
785	CKTWESGLNFGHVF	1609	TGCAAAACATGGGAAAGTGGCCTTAATTTTGGCCACGTCTTC
786	CKTWDTSLSAVVF	1610	TGCAAAACATGGGATACCAGCCTGAGTGCTGTGGTCTTC
787	CGVWDVSLGAGVF	1611	TGCGGAGTCTGGGATGTCAGTCTGGGTGCTGGGGTGTTTC
788	CGVWDTPSAVLF	1612	TGCGGAGTCTGGGATACCACCCGAGTGCCGTTCTTTTC
789	CGVWDTTLSAVLF	1613	TGCGGAGTCTGGGATACCACCCTGAGTGCCGTTCTTTTC
790	CGVWDTS LGVF	1614	TGCGGAGTATGGGATACCAGCCTGGGGGTCTTC
791	CGVWDTNLGKVV	1615	TGCGGGGTATGGGATACCAACCTGGGTAAATGGGTTTTC
792	CGVWDTG LDAGVVF	1616	TGTGGAGTTTGGGATACTGGCCTGGATGCTGGTTGGGTGTTTC
793	CGVWDNVLEAYVF	1617	TGCGGAGTGTGGGATAACGTCCTGGAGGCCTATGTCTTC
794	CGVWDISLSANWVF	1618	TGCGGAGTCTGGGATATCAGCCTGAGTGCTAATTGGGTGTTTC
795	CGVWDHSLGIWAF	1619	TGCGGAGTATGGGATCACAGCCTGGGGATTTGGGCCTTC
796	CGVWDDILTAEVF	1620	TGCGGAGTTTGGGATGATATTCTGACTGCTGAAGTGTTTC
797	CGVRDTS LGVF	1621	TGCGGAGTTCGGGATACCAGCCTGGGGGTCTTC
798	CGTYDTS LPAWVF	1622	TGCGGAACATACGATACGAGCCTGCCTGCTTGGGTGTTT
799	CGTYDNLVFGYVF	1623	TGCGGAACCTACGATAATCTTGTATTTGGTTATGTCTTC
800	CGTYDDRLREVF	1624	TGCGGAACATACGATGATAGACTCAGAGAGGTGTTTC
801	CGTWVTSLSAGVF	1625	TGCGGAACGTGGGTTACCAGCCTGAGTGCTGGGGTGTTTC
802	CGTWVSSLTTVVVF	1626	TGCGGAACATGGGTTAGCAGCCTGACTACTGTAGTATTC
803	CGTWVSSLNVWVF	1627	TGCGGAACATGGGTTAGCAGCCTGAACGTCTGGGTGTTTC
804	CGTWVGRFWVF	1628	TGCGGAACATGGGTTGGCAGGTTTTGGGTATTC
805	CGTWSGGPSGHWLF	1629	TGCGGAACATGGTCTGGCGGCCCGAGTGCCATTGGTTGTTTC
806	CGTWSGGLSGHWLF	1630	TGCGGAACATGGTCTGGCGGCCTGAGTGCCATTGGTTGTTTC
807	CGTWQTGREAVLF	1631	TGCGGAACGTGGCAGACCGGCCGGGAGGCTGTCTATTT
808	CGTWQSRLRWVF	1632	TGCGGAACGTGGCAGAGCAGGCTGAGGTGGGTGTTTC
809	CGTWQSRLGWVF	1633	TGCGGAACGTGGCAGAGCAGGCTGGGGTGGGTGTTTC
810	CGTWPRSLSAVWVF	1634	TGCGGAACATGGCCTAGGAGCCTGAGTGCTGTTTGGGTGTTTC
811	CGTWNNYLSAGDVVF	1635	TGCGGAACATGGAATAACTACCTGAGTGCTGGCGATGTGGTTTTC
812	CGTWLGSQSPYWVF	1636	TGCGGAACATGGCTTGGCAGCCAGAGTCCTTATTGGGTCTTC
813	CGTWHTGLSAYVF	1637	TGCGGAACATGGCATAACCGGCCTGAGTGCTTATGTCTTC
814	CGTWHS TLSAGHWVF	1638	TGCGGAACATGGCATAAGTACCCTGAGTGCTGGCCATTGGGTGTTC
815	CGTWHSSLSTWVF	1639	TGCGGAACATGGCATAAGTAGCCTGAGTACTTGGGTGTTTC
816	CGTWHSSL SAYVF	1640	TGCGGAACATGGCATAAGCAGCCTGAGTGCTTATGTCTTC
817	CGTWHSSL SAVVF	1641	TGCGGAACATGGCATAAGCAGCCTGAGTGCTGTGGTATTC
818	CGTWHSGLSGWVF	1642	TGCGGAACGTGGCATTCCGGCCTGAGTGGGTGGGTTTTTC
819	CGTWHNTLRNVIF	1643	TGCGGAACATGGCATAACACCCTGCGTAATGTGATATTC
820	CGTWHASLTAVF	1644	TGCGGAACATGGCATGCCAGCCTGACTGCTGTGTTTC
821	CGTWGWYGSQRGVVF	1645	TGCGGGACATGGGGATGGTATGGCAGCCAGAGAGGCGTCGTCTTC
822	CGTWGWYGGQRGVVF	1646	TGCGGGACATGGGGATGGTATGGCGGCCAGAGAGGCGTCGTCTTC
823	CGTWGTSLSAWVF	1647	TGCGGAACCTGGGGAACCAGCCTGAGTGCTTGGGTGTTTC
824	CGTWGSSLTTGLF	1648	TGCGGAACCTGGGGTAGCAGCCTGACTACTGGCCTGTTTC

825	CGTWGSSLTAYVF	1649	TGCGGAACATGGGGTAGCAGCCTGACTGCCTATGTCTTC
826	CGTWGSSLSVVF	1650	TGCGGAACATGGGGTAGCAGCCTGAGTGTTGTGTTC
827	CGTWGSSLSGGVF	1651	TGCGGAACATGGGGTAGCAGCCTGAGTGGTGGGGTGTTTC
828	CGTWGSSLSAYWVF	1652	TGCGGAACATGGGGTAGCAGCCTGAGTGCTTATTGGGTGTTTC
829	CGTWGSSLSAYVVF	1653	TGCGGAACATGGGGTAGCAGCCTGAGTGCTTATGTGGTGTTTC
830	CGTWGSSLSAYVF	1654	TGCGGAACATGGGGTAGCAGCCTGAGTGCTTATGTCTTC
831	CGTWGSSLSAVVF	1655	TGCGGAACGTGGGGTAGTAGCCTGAGTGCTGTGGTGTTTC
832	CGTWGSSLSAPYVF	1656	TGCGGAACATGGGGTAGCAGCCTGAGTGCTCCTTATGTCTTC
833	CGTWGSSLSAPVF	1657	TGCGGAACATGGGGTAGCAGCCTGAGTGCCCCGGTGTTTC
834	CGTWGSSLSAGVF	1658	TGCGGAACATGGGGTAGCAGCCTGAGTGCTGGGGTGTTTC
835	CGTWGSSLSAGLF	1659	TGCGGAACCTGGGGTAGCAGCCTGAGTGCTGGACTGTTTC
836	CGTWGSSLSAGALF	1660	TGCGGAACATGGGGTAGCAGCCTGAGTGCTGGGGCACTCTTC
837	CGTWGSSLRAWVF	1661	TGCGGAACATGGGGCAGTAGCCTGCGTGCTTGGGTGTTTC
838	CGTWFTSLASGVF	1662	TGCGGAACCTGGTTTACTAGTCTGGCTAGTGGGGTTTTTC
839	CGTWETSLSVVI	1663	TGCGGAACCTGGGAGACCAGTCTGAGTGTCGTGGTCATC
840	CGTWETSLSGVF	1664	TGCGGAACATGGGAGACCAGCCTGAGTGGTGTCTTC
841	CGTWETSLSDWVF	1665	TGCGGAACATGGGAAACCAGCCTGAGTGATTGGGTATTC
842	CGTWETSLSAGVF	1666	TGCGGAACATGGGAGACCAGCCTGAGTGCTGGGGTATTC
843	CGTWETSLNYVAF	1667	TGCGGAACATGGGAAACCAGCCTTAATTATGTGGCCTTC
844	CGTWETSLNTWLL	1668	TGCGGAACATGGGAGACCAGCCTGAATACTTGGTTGCTC
845	CGTWETSESGNYIF	1669	TGCGGAACATGGGAGACCAGCGAGAGTGGAATTACATCTTC
846	CGTWETRLGTWVI	1670	TGCGGAACATGGGAAACCAGACTGGGTACTTGGGTGATC
847	CGTWETQLYWVF	1671	TGCGGAACATGGGAGACCCAGTTATATTGGGTGTTTC
848	CGTWETGLSAGEVF	1672	TGCGGAACATGGGAGACTGGCCTAAGTGCTGGAGAGGTGTTTC
849	CGTWESTLSVFLF	1673	TGCGGAACCTGGGAAAGCACCTGAGTGTTTTCTTATTC
850	CGTWESSLTVVVF	1674	TGCGGGACATGGGAAAGTAGCCTGACTGTTGTGGTCTTC
851	CGTWESSLTGVVF	1675	TGCGGAACATGGGAAAGTAGCCTGACTGGAGTGGTATTC
852	CGTWESSLTGFVF	1676	TGCGGAACATGGGAAAGCAGCCTGACTGGTTTTGTCTTC
853	CGTWESSLSVGVF	1677	TGTGGAACATGGGAGAGCAGCCTGAGTGTTGGGGTGTTTC
854	CGTWESSLSEWVF	1678	TGCGGAACCTGGGAAAGTAGCCTCAGTGAATGGGTGTTTC
855	CGTWESSLSAVF	1679	TGCGGAACATGGGAGAGCAGCCTGAGTGCTGTATTC
856	CGTWESSLSAGYIF	1680	TGCGGAACATGGGAGAGCAGCCTGAGTGCTGGTTATATCTTC
857	CGTWESSLSAGVF	1681	TGCGGAACATGGGAGAGCAGCCTGAGTGCTGGAGTGTTTC
858	CGTWESSLSAGPVF	1682	TGCGGAACATGGGAAAGCAGCCTGAGCGCTGGCCCCGGTGTTTC
859	CGTWESSLSAGGQVF	1683	TGCGGAACATGGGAAAGCAGCCTGAGTGCTGGAGGCCAGGTGT TC
860	CGTWESSLSAFGGYV F	1684	TGCGGAACATGGGAGAGCAGCCTGAGTGCCCTCGGCGGTTATG TCTTC
861	CGTWESSLRVWVF	1685	TGCGGAACATGGGAAAGCAGCCTGAGGGTTTGGGTGTTTC
862	CGTWESSLFTGPWVF	1686	TGCGGAACATGGGAAAGCAGCCTCTTTACTGGGCCTTGGGTGT TC
863	CGTWESLSATYVF	1687	TGCGGAACATGGGAGAGCCTGAGTGCCACCTATGTCTTC
864	CGTWESGLSAGVF	1688	TGCGGAACATGGGAGAGCGCCTGAGTGCTGGTGTCTTC
865	CGTWESDFWVF	1689	TGCGGAACATGGGAAAGCGACTTTTGGGTGTTT
866	CGTWENRLSAVVF	1690	TGCGGTACATGGGAAAACAGACTGAGTGCTGTGGTCTTC
867	CGTWENRLSAGVF	1691	TGCGGAACATGGGAAAACAGACTGAGTGCCGGGGTATTC
868	CGTWEISLTTSVVF	1692	TGCGGAACATGGGAAATCAGCCTGACTACTTCTGTGGTATTC

869	CGTWEISLSTSVVF	1693	TGCGGAACATGGGAAATCAGCCTGAGTACTTCTGTGGTATTC
870	CGTWEGLSVVF	1694	TGCGGAACATGGGAAGGCAGCCTCAGTGTGTTTTTC
871	CGTWEGLRVF	1695	TGCGGAACATGGGAAGGCAGCCTGAGGGTGTTC
872	CGTWEGLRHVF	1696	TGCGGAACATGGGAGGGCAGCCTGAGGCACGTGTTC
873	CGTWDYSPVRAGVF	1697	TGCGGAACATGGGATTACAGCCCTGTACGTGCTGGGGTGTTC
874	CGTWDYSLSVYLF	1698	TGCGGAACGTGGGATTACAGCCTGAGTGTGTTATCTCTTC
875	CGTWDYSLSSGVVF	1699	TGCGGAACATGGGATTACAGCCTGAGTTCTGGCGTGGTATTC
876	CGTWDYSLSAWVF	1700	TGCGGAACATGGGATTACAGCCTGAGTGCCTGGGTGTTC
877	CGTWDYSLSAEVF	1701	TGCGGAACATGGGATTACAGTCTGAGTGTGAGGTGTTC
878	CGTWDYSLRRAIF	1702	TGCGGAACATGGGATTACAGCCTGCGTCGTGCGATATTC
879	CGTWDWSLILQLF	1703	TGCGGAACATGGGATTGGAGCCTCATTCTTCAATTGTTC
880	CGTWDVTLHTGVF	1704	TGCGGAACATGGGATGTCACCTTGCATACTGGGGTGTTC
881	CGTWDVTLHIGVF	1705	TGCGGAACATGGGATGTCACCTTGCATATTGGGGTGTTC
882	CGTWDVTLHAGVF	1706	TGCGGAACATGGGATGTCACCTTGCATGCTGGGGTGTTC
883	CGTWDVSLYSGGVF	1707	TGCGGAACATGGGATGTCAGTTTGTATAGTGGCGGGGTCTTC
884	CGTWDVSLTSFVF	1708	TGTGGAACATGGGATGTCAGCCTGACTTCTTTCTGTCTTC
885	CGTWDVSLSVGL	1709	TGCGGAACATGGGATGTCAGCCTGAGTGTGAGGGTGTTC
886	CGTWDVSLSAGDVVF	1710	TGCGGAACGTGGGATGTCAGCCTGAGTGTGCGCATGTAGTTTTC
887	CGTWDVSLNVVVF	1711	TGCGGAACATGGGATGTCAGCCTGAATGTCGTGGTTTTTC
888	CGTWDVSLNTQVF	1712	TGCGGAACATGGGATGTCAGCCTGAATACTCAGGTGTTC
889	CGTWDVSLGALF	1713	TGCGGCACATGGGATGTGAGCCTGGGTGCGCTGTTC
890	CGTWDVNLKTVVF	1714	TGCGGAACGTGGGACGTTAATCTGAAAACGTGCTTTTTTC
891	CGTWDVILSAEVF	1715	TGCGGAACATGGGATGTCATCCTGAGTGTGAGGTATTC
892	CGTWDTTVSADVVF	1716	TGCGGAACATGGGATACACCGTGAGTGTGAGTTTTTC
893	CGTWDTTLTAWVF	1717	TGCGGAACATGGGATACACCGTGACTGCCTGGGTGTTC
894	CGTWDTTLSVFLF	1718	TGCGGAACATGGGACACCACCTTGAGTGTGTTTTCTATTC
895	CGTWDTSVSAGVF	1719	TGCGGGACTTGGGATACCAAGTGTGAGTGTGAGGTGTTC
896	CGTWDTSVISWVF	1720	TGCGGAACATGGGATACCAAGTGTGATTTCTTGGGTTTTTTC
897	CGTWDTSRSLYVVF	1721	TGCGGAACATGGGATACCAAGTCGGAGTTCTCTATGTGGTCTTC
898	CGTWDTSRSAWVF	1722	TGCGGAACATGGGATACCAAGCCGAGTGTGAGGTATTC
899	CGTWDTSRNPGGIF	1723	TGCGGAACATGGGATACCAAGCCGGAATCCTGGAGGAATTTTTTC
900	CGTWDTSRGHVF	1724	TGCGGAACATGGGACACCAGTCGGGGTCATGTTTTTC
901	CGTWDTSPTGQVLF	1725	TGCGGAACATGGGATACCAAGCCCGAGTACTGGCCAGGTGCTTTTC
902	CGTWDTSPSAWVF	1726	TGCGGAACATGGGATACCAAGCCCGAGTGCCTGGGTGTTC
903	CGTWDTSLTWVF	1727	TGCGGAACATGGGATACTAGCCTGACCTGGGTGTTC
904	CGTWDTSLTWFAVF	1728	TGCGGAACATGGGATACCAAGCCTGACGTGGTTTCGAGTGTTC
905	CGTWDTSLTVVVF	1729	TGCGGAACATGGGATACCAAGCCTGACTGTTGTGGTATTC
906	CGTWDTSLTTSWVF	1730	TGCGGAACATGGGATACCAAGCCTGACTACTTCTTGGGTGTTC
907	CGTWDTSLTTPFWVF	1731	TGCGGAACATGGGATACCAAGCCTGACCACTGGTCCTTTTTGGGTGTTC
908	CGTWDTSLTTPFYVF	1732	TGCGGAACATGGGATACCAAGCCTGACTCCTTTTTATGTCTTC
909	CGTWDTSLTAYVF	1733	TGCGGAACATGGGATACCAAGCCTGACTGCTTATGTCTTC
910	CGTWDTSLTAWVF	1734	TGCGGAACATGGGATACCAAGCCTGACTGCTTGGGTGTTC
911	CGTWDTSLTAWGVF	1735	TGCGGAACATGGGATACCAAGCCTGACTGCGTGGGGGTGTTC

912	CGTWDTSLTAVVL	1736	TGCGGCACATGGGATACCAGCCTGACTGCGGTGGTTCTC
913	CGTWDTSLTARVF	1737	TGCGGAACCTGGGATACCAGCCTGACTGCTCGGGTTTTTC
914	CGTWDTSLTAI VF	1738	TGCGGAACATGGGATACCAGCCTGACTGCGATTGTCTTC
915	CGTWDTSLTAGVF	1739	TGCGGAACATGGGATACCAGCCTGACTGCTGGTGTCTTC
916	CGTWDTSLSVYVF	1740	TGCGGAACATGGGATACCAGCCTGAGTGTTTATGTCTTC
917	CGTWDTSLSVVF	1741	TGCGGAACATGGGATACCAGCCTGAGTGTTGGTGTTC
918	CGTWDTSLSVGEF	1742	TGCGGGACATGGGATACCAGCCTGAGTGTTGGGGAATTC
919	CGTWDTSLSWVF	1743	TGCGGAACATGGGATACCAGCCTGAGTACTTGGGTGTTC
920	CGTWDTSLSVVF	1744	TGCGGAACATGGGATACCAGCCTGAGTACTGTGGTATTC
921	CGTWDTSLSGQVLF	1745	TGCGGAACATGGGATACCAGCCTGAGTACTGGCCAGGTGCTTT TC
922	CGTWDTSLSGPLW VF	1746	TGCGGCACATGGGATACCAGCCTGAGCACTGGTCTCTTTGGGT GTTC
923	CGTWDTSLSYVF	1747	TGCGGAACCTGGGATACCAGCCTGAGTTCTTATGTCTTC
924	CGTWDTSLSVVF	1748	TGCGGAACATGGGATACCAGCCTGAGTTCTGTGGTCTTC
925	CGTWDTSLSRYIF	1749	TGCGGAACATGGGATACCAGCCTGAGTTCTAGATACATATTC
926	CGTWDTSLSRFIF	1750	TGCGGAACATGGGATACCAGCCTGAGTTCTAGATTCATATTC
927	CGTWDTSLSGWVF	1751	TGCGGAACATGGGATACCAGCCTGAGTTCTGGGTGGGTGTTC
928	CGTWDTSLSRYVF	1752	TGCGGAACATGGGATACCAGCCTGAGTCGGTATGTGTTC
929	CGTWDTSLSQWLF	1753	TGCGGAACCTGGGATACCAGTCTGAGTCAATGGCTGTTC
930	CGTWDTSLSPLWV F	1754	TGCGGAACATGGGATACCAGCCTGAGTCCTGGCCTTTGGGTGT TC
931	CGTWDTSLSNYVF	1755	TGCGGAACATGGGATACCAGCCTGAGTAATTATGTCTTC
932	CGTWDTSLSIWVF	1756	TGCGGAACATGGGATACCAGCCTAAGTATTTGGGTGTTC
933	CGTWDTSLSIGFWV F	1757	TGCGGCACATGGGATACCAGCCTGAGCATTGGTCTTTTGGGT GTTC
934	CGTWDTSLSGWVF	1758	TGCGGAACATGGGATACCAGCCTGAGTGTTGGGTGTTC
935	CGTWDTSLSGTVF	1759	TGCGGAACATGGGATACCAGCCTGAGTGGTACAGTGTTC
936	CGTWDTSLSGGQVF	1760	TGCGGAACATGGGATACTAGTCTGAGTGTTGGCCAGGTGTTC
937	CGTWDTSLSGGIF	1761	TGCGGAACATGGGATACCAGCCTGAGTGTTGGGATATTC
938	CGTWDTSLSGEDVVI	1762	TGCGGAACATGGGATACCAGCCTGAGTGTTGAGGATGTGGTAA TC
939	CGTWDTSLSFLYAF	1763	TGCGGAACATGGGATACCAGCCTGAGTTTCCTTTATGCTTTC
940	CGTWDTSLSSEVVF	1764	TGCGGAACATGGGATACCAGCCTGAGTGAGGTCGTATTC
941	CGTWDTSLSSEVF	1765	TGCGGAACATGGGATACCAGCCTGAGTGAAGTGTTC
942	CGTWDTSLSSENWVF	1766	TGCGGAACATGGGATACTAGCCTGAGTGAAAATTGGGTGTTC
943	CGTWDTSLSAYIF	1767	TGCGGAACATGGGATACCAGCCTGAGTGCCTACATATTC
944	CGTWDTSLSAVVL	1768	TGCGGAACATGGGATACCAGCCTGAGTGCTGTGGTACTC
945	CGTWDTSLSAVF	1769	TGCGGAACATGGGATACCAGCCTGAGTGCTGTTTTTC
946	CGTWDTSLSARVF	1770	TGCGGAACATGGGATACCAGCCTGAGTGCCCCGGGTGTTC
947	CGTWDTSLSARQVF	1771	TGCGGCACATGGGATACCAGCCTGAGTGCCCCGCCAGGTATTC
948	CGTWDTSLSALVF	1772	TGCGGAACATGGGATACCAGCCTGAGTGCTTTGGTTTTTC
949	CGTWDTSLSAKVF	1773	TGCGGAACATGGGATACCAGCCTGAGTGCTAAGGTGTTC
950	CGTWDTSLSAKIF	1774	TGCGGAACATGGGATACCAGCCTGAGTGCGAAAATCTTC
951	CGTWDTSLSAKAVF	1775	TGCGGAACATGGGATACCAGCCTGAGTGCCAAGGCGGTATTC
952	CGTWDTSLSAHAVF	1776	TGCGGAACATGGGATACCAGCCTGAGTGCCCATGCTGTGTTC
953	CGTWDTSLSAGYVF	1777	TGCGGAACATGGGATACCAGCCTGAGTGCTGGCTATGTCTTC
954	CGTWDTSLSAGRWF	1778	TGCGGAACATGGGACACCAGTCTGAGTGCTGGCCGCTGGGTGT

	F		TC
955	CGTWDTSLSAGIF	1779	TGCGGAACATGGGATACCAGCCTGAGTGCTGGGATATTC
956	CGTWDTSLSAGGFR VF	1780	TGCGGAACATGGGATACCAGCCTGAGTGCTGGTGGGTTCCGGG TCTTC
957	CGTWDTSLSAGAF	1781	TGCGGAACATGGGATACCAGCCTGAGTGCTGGGGCATTC
958	CGTWDTSLSADWFF	1782	TGCGGAACATGGGATACCAGTCTGAGTGCTGATTGGTTTTTC
959	CGTWDTSLSADEYVF	1783	TGCGGAACATGGGATACCAGCCTGAGTGCTGATGAATATGTCT TC
960	CGTWDTSLSAAWVF	1784	TGCGGCACATGGGATACCAGCCTGAGTGCGGCTTGGGTGTTC
961	CGTWDTSLSAALF	1785	TGCGGAACATGGGATACCAGCCTGAGTGCTGCGCTATTC
962	CGTWDTSLSAAGVF	1786	TGCGGAACATGGGATACCAGCCTGAGTGCTGCGGGGGTTTTTC
963	CGTWDTSLRVVVF	1787	TGCGGAACATGGGATACCAGCCTGAGAGTTGTGGTTTTTC
964	CGTWDTSRLTWVF	1788	TGCGGAACATGGGATACCAGCCTGAGAACCTGGGTATTC
965	CGTWDTSLRGAVF	1789	TGCGGAACGTGGGATACCAGCCTGAGGGGTGCAGTGTTTC
966	CGTWDTSRLRAVVF	1790	TGCGGAACATGGGATACCAGCCTGCGTGCTGTGGTATTC
967	CGTWDTSNLVVYVF	1791	TGCGGAACATGGGATACAAGCCTGAATGTAGTTTATGTCTTC
968	CGTWDTSNLTYLF	1792	TGCGGAACATGGGATACCAGCCTCAACACCTACCTGTTC
969	CGTWDTSNLFAWLF	1793	TGCGGAACATGGGATACTAGCCTGAACCTCGCTTGGCTGTTC
970	CGTWDTSLLVWLF	1794	TGCGGCACATGGGATACCAGCCTTCTTGTGTGGCTTTTC
971	CGTWDTSLKTWVF	1795	TGCGGAACATGGGATACCAGTCTGAAGACGTGGGTGTTC
972	CGTWDTSLIVWVF	1796	TGCGGAACATGGGATACCAGTCTGATTGTCTGGGTGTTC
973	CGTWDTSRLITGVF	1797	TGCGGAACATGGGATACCAGCCTAATTACTGGGGTGTTC
974	CGTWDTSRLISVVF	1798	TGCGGAACATGGGATACCAGCCTGATTAGCGTGGTATTC
975	CGTWDTSRLIAYVF	1799	TGCGGAACATGGGATACCAGCCTGATTGCTTATGTCTTC
976	CGTWDTSRLHTELF	1800	TGCGGAACATGGGATACCAGCCTGCACACTGAGTTGTTC
977	CGTWDTSRLGSYVF	1801	TGCGGAACCTGGGATACCAGCCTGGGTTCTTATGTCTTC
978	CGTWDTSRLGSLWVF	1802	TGCGGAACATGGGATACCAGCCTGGGTTCTCTTTGGGTGTTC
979	CGTWDTSRLGSGVF	1803	TGCGGTACATGGGATACCAGCCTGGGTTCTGGGGTATTC
980	CGTWDTSRLGGRGVF	1804	TGCGGAACCTGGGATACCAGTCTGGGTGGTAGAGGGGTCTTC
981	CGTWDTSRLGAWVF	1805	TGCGGAACATGGGATACCAGCCTGGGTGCTTGGGTGTTC
982	CGTWDTSRLGAVVF	1806	TGCGGAACATGGGATACCAGCCTGGGTGCCGTGGTATTC
983	CGTWDTSRLGAGVF	1807	TGCGGAACATGGGATACCAGCCTGGGTGCTGGGGTATTC
984	CGTWDTSRLGAGLF	1808	TGCGGAACATGGGATACCAGCCTGGGTGCTGGCCTATTC
985	CGTWDTSRLDAVVF	1809	TGCGGAACATGGGATACCAGTCTGGATGCTGTGGTTTTTC
986	CGTWDTSRLDAVLF	1810	TGCGGGACTTGGGATACCAGCCTGGATGCTGTGCTGTTC
987	CGTWDTSRLAWVF	1811	TGCGGAACATGGGATACCAGCCTGGCTTGGGTGTTC
988	CGTWDTSRLATGLF	1812	TGCGGAACATGGGATACCAGCCTGGCGACTGGACTGTTC
989	CGTWDTSRLAPVVF	1813	TGCGGGACATGGGATACCAGCCTGGCCCTGTAGTCTTC
990	CGTWDTRLTIVIF	1814	TGCGGAACATGGGACACCCGCTGACTATTGTGATCTTC
991	CGTWDTRLSSVWLF	1815	TGTGGAACATGGGACACCAGGCTGAGTGTTTGGCTGTTC
992	CGTWDTRLSSVGVF	1816	TGCGGAACGTGGGACACCAGACTGAGTGTTGGGGTTTTTC
993	CGTWDTRLSTVIF	1817	TGCGGCACATGGGATACCAGACTGAGTACTGTAATTTTTTC
994	CGTWDTRLSSVVF	1818	TGCGGAACATGGGATACCCGCTGAGTTCTGTGGTCTTC
995	CGTWDTRLSSIVVF	1819	TGCGGAACATGGGATACCCGCTGAGTATTGTGGTTTTTC
996	CGTWDTRLSSAYVVF	1820	TGCGGAACATGGGATACCAGACTGAGTGCCTATGTGGTATTC
997	CGTWDTRLSSAWVF	1821	TGCGGAACCTGGGACACCCGCTGAGTGCGTGGGTGTTC

998	CGTWDTRL SAVVF	1822	TGCGGAACATGGGATACCAGACTGAGTGCTGTGGTGTTTC
999	CGTWDTRL SAGLF	1823	TGCGGAACATGGGATACCCGCCTGAGTGCTGGGTTGTTC
1000	CGTWDTRL SAGGVF	1824	TGCGGAACATGGGATACCAGACTGAGTGCTGGTGGGGTGTTTC
1001	CGTWDTRL NVWLF	1825	TGCGGAACATGGGATACCAGATTGAATGTGTGGCTATTC
1002	CGTWDTNREV VLL	1826	TGCGGAACATGGGATACCAACCGGGAAGTTGTGCTCCTC
1003	CGTWDTNLRAHVF	1827	TGCGGAACATGGGATACCAACCTGCGTGCCCATGTCTTC
1004	CGTWDTNLPAVVF	1828	TGCGGAACATGGGATACTAATCTGCCCGCTGTAGTGTTTC
1005	CGTWDTNLGGVF	1829	TGCGGAACATGGGACACCAATTTGGGTGGGGTGTTTC
1006	CGTWDTIVSIGVF	1830	TGCGGAACATGGGATACCATCGTGAGTATTGGGGTGTTTC
1007	CGTWDTILSAVVF	1831	TGCGGAACATGGGATACCATCCTGAGTGCGGTGGTGTTTC
1008	CGTWDTILSAEVF	1832	TGCGGCACATGGGATACCATCCTGAGTGCTGAGGTGTTTC
1009	CGTWDTHLGVVF	1833	TGCGGAACATGGGATACCCACCTGGGTGTGGTTTTTC
1010	CGTWDTGPSPHWLF	1834	TGCGGAACATGGGATACCGGCCCCGAGCCCTCATTGGCTGTTTC
1011	CGTWDTGLTFGGVF	1835	TGCGGAACATGGGATACCGGCCTGACTTTTGGAGGCGTGTTTC
1012	CGTWDTGLTAFVF	1836	TGCGGAACATGGGATACCGGCCTGACTGCTTTTGTCTTC
1013	CGTWDTGLSVWVF	1837	TGCGGAACATGGGATACCGGCCTGAGTGTTTGGGTGTTTC
1014	CGTWDTGLSTGIF	1838	TGCGGAACATGGGATACCGGCCTGAGTACTGGGATTTTC
1015	CGTWDTGLSLLF	1839	TGCGGAACATGGGATACCGGCCTGAGTTCCTGCTCTTC
1016	CGTWDTGLSIVVF	1840	TGCGGAACGTGGGACACCGGCCTGAGTATTGTGGTGTTTC
1017	CGTWDTGLSFVVF	1841	TGCGGAACGTGGGACACCGGCCTGAGTTTTGTGGTGTTTC
1018	CGTWDTGLSAWVF	1842	TGCGGAACATGGGATACCGGCCTGAGTGCTTGGGTGTTTC
1019	CGTWDTGLSAGVVF	1843	TGCGGAACATGGGATACCGGCCTGAGTGCTGGTGTGGTATTC
1020	CGTWDTGLRGWIF	1844	TGCGGAACATGGGATACCGGTCTGAGGGGTTGGATTTTC
1021	CGTWDTELSAGVF	1845	TGCGGAACATGGGATACCGAGCTAAGTGCGGGGGTCTTC
1022	CGTWDTALTAGVF	1846	TGCGGAACGTGGGATACCGCCCTGACTGCTGGGGTGTTTC
1023	CGTWDTALSIVVF	1847	TGCGGAACATGGGATACTGCCCTGAGTCTTGTGGTCTTC
1024	CGTWDTALS AWLF	1848	TGCGGAACATGGGATACCGCCCTGAGTGCCTGGCTGTTTC
1025	CGTWDTALSAGVF	1849	TGCGGCACATGGGATACCGCCCTGAGTGCTGGGGTGTTTC
1026	CGTWDTALRGVLF	1850	TGCGGAACATGGGATACCGCCCTGCGTGCGGTGCTGTTTC
1027	CGTWDTALKEWLF	1851	TGCGGAACATGGGATACCGCCCTGAAAGAATGGCTGTTTC
1028	CGTWDRTL TAGDVL F	1852	TGCGGAACATGGGATAGGACCCTGACTGCTGGCGATGTGCTCT TC
1029	CGTWDRSVTYVF	1853	TGCGGAACATGGGATAGAAGCGTGACTTATGTCTTC
1030	CGTWDRSRNEWVF	1854	TGCGGAACATGGGATCGCAGCCGAAATGAATGGGTGTTTC
1031	CGTWDRSLTVWVF	1855	TGCGGAACATGGGATCGCAGTCTGACTGTTTGGGTCTTC
1032	CGTWDRSLTPGWLF	1856	TGCGGAACATGGGATCGCAGCCTGACTCCTGGGTGGTTGTTTC
1033	CGTWDRSLTAWVF	1857	TGCGGAACATGGGATAGAAGCCTGACTGCTTGGGTGTTTC
1034	CGTWDRSLSVVVF	1858	TGCGGAACATGGGACCGCAGCCTGAGTGTTGTGGTATTC
1035	CGTWDRSLSVVF	1859	TGCGGCACATGGGATCGCAGCCTGAGTGTAAGTCTTC
1036	CGTWDRSLSVQLF	1860	TGCGGAACATGGGATAGGAGCCTGAGTGTTCAATTGTTTC
1037	CGTWDRSLSVLWVF	1861	TGCGGAACATGGGATCGCAGCCTCAGTGTTCTTTGGGTGTTTC
1038	CGTWDRSLSVGLF	1862	TGCGGAACATGGGATCGCAGCCTGAGTGTTGGATTATTC
1039	CGTWDRSLSTWVF	1863	TGCGGAACATGGGATCGCAGCCTGAGTACTTGGGTGTTTC
1040	CGTWDRSLSTHWVL	1864	TGCGGAACATGGGATAGAAGCCTGAGTACTCATTGGGTGCTC
1041	CGTWDRSLSTHWVF	1865	TGCGGAACATGGGATAGAAGCCTGAGTACTCATTGGGTGTTTC
1042	CGTWDRSLSSAVF	1866	TGCGGAACCTGGGATCGAAGCCTGAGTTCTGCGGTGTTTC

1043	CGTWDRSLSPSYVF	1867	TGCGGAACATGGGACAGAAGCCTGAGTCCCTCTTATGTCTTC
1044	CGTWDRSLSGEVF	1868	TGCGGAACATGGGATAGGAGCCTGAGTGGTGAGGTGTTC
1045	CGTWDRSLSGAVF	1869	TGCGGAACATGGGATAGGAGCCTGAGTGGTGCGGTGTTC
1046	CGTWDRSLSAVAF	1870	TGCGGAACATGGGATCGCAGCCTGAGTGCTGTGGCATTC
1047	CGTWDRSLSAGGEF	1871	TGCGGAACATGGGATAGGAGCCTGAGTGCCGGGGGGGAATTC
1048	CGTWDRSLSAFWVF	1872	TGCGGAACATGGGATCGCAGCCTGAGTGCTTTTTGGGTGTTC
1049	CGTWDRSLSAAVF	1873	TGCGGAACATGGGATAGGAGCCTGAGTGCTGCGGTGTTC
1050	CGTWDRSLSAALF	1874	TGCGGAACATGGGATAGGAGCCTGAGTGCTGCACTCTTC
1051	CGTWDRSLRVF	1875	TGCGGAACATGGGATCGCAGCCTGAGAGTGTTC
1052	CGTWDRSLNWVF	1876	TGCGGTACATGGGACAGAAGCCTTAATTGGGTGTTC
1053	CGTWDRSLNVYVF	1877	TGCGGAACATGGGATCGCAGCCTGAATGTTTATGTCTTC
1054	CGTWDRSLNVGVF	1878	TGCGGAACATGGGATAGGAGCCTGAATGTTGGGTGTTC
1055	CGTWDRSLHVVVF	1879	TGCGGAACATGGGATCGGAGCCTGCATGTGGTCTTC
1056	CGTWDRSLGGWVF	1880	TGTGGAACATGGGATCGCAGCCTGGGTGGTTGGGTGTTC
1057	CGTWDRSLGAFWVF	1881	TGCGGAACATGGGATCGCAGCCTGGGTGCTTTTTGGGTGTTC
1058	CGTWDRSLFWVF	1882	TGCGGAACATGGGATAGAAGCCTGTTTTGGGTGTTC
1059	CGTWDRSLAAGVF	1883	TGCGGAACATGGGATCGCAGCCTGGCTGCTGGGTGTTC
1060	CGTWDRRLSGVVF	1884	TGCGGAACATGGGATAGGAGGTTGAGTGGTGTCGATTC
1061	CGTWDRRLSDVVF	1885	TGCGGAACATGGGATCGCCGCCTAAGTGATGTGGTATTC
1062	CGTWDRRLSAVVF	1886	TGCGGAACATGGGATAGGAGGCTGAGTGCTGTGGTATTC
1063	CGTWDRRLNVAFF	1887	TGCGGAACATGGGATAGACGCCTGAATGTTGCGTCTTC
1064	CGTWDRRLLAFF	1888	TGTGGAACATGGGATAGGAGGCTGCTTGCTGTTTTTC
1065	CGTWDRNLRAVVF	1889	TGCGGAACATGGGATAGGAACCTGCGCGCCGTGGTCTTC
1066	CGTWDRSLAGVF	1890	TGCGGAACATGGGATAGGCTGAGTGCTGGGTGTTC
1067	CGTWDRGPNTGVF	1891	TGCGGAACATGGGATAGAGGCCGAATACTGGGTATTC
1068	CGTWDRGLNTVYVF	1892	TGCGGAACATGGGATAGAGGCCTGAATACTGTTTACGTCTTC
1069	CGTWDNYVSAPWVF	1893	TGCGGAACATGGGATAACTATGTGAGTGCCCCCTGGGTGTTC
1070	CGTWDNYLSAGDVVF	1894	TGCGGAACATGGGATAACTACCTGAGTGCTGGCGATGTGGTTTTC
1071	CGTWDNYLRAGVF	1895	TGCGGAACATGGGATAACTACCTGAGAGCTGGGTCTTC
1072	CGTWDNYLGAVVF	1896	TGCGGAACATGGGACAATTATCTGGGTGCCGTGGTTTTTC
1073	CGTWDNYLGAGVF	1897	TGCGGAACATGGGATAACTACCTGGGTGCGGGGGTGTTC
1074	CGTWDNTVSAPWVF	1898	TGCGGAACATGGGATAACACCGTGAGTGCCCCCTGGGTTTTTTC
1075	CGTWDNTLSLWVF	1899	TGCGGAACATGGGATAACACCCTGAGTCTTTGGGTGTTC
1076	CGTWDNTLSAGVF	1900	TGCGGAACATGGGATAACACCCTGAGTGCTGGGTCTTC
1077	CGTWDNTLLTVLF	1901	TGCGGAACATGGGACAACACTCTGCTTACTGTGTTATTC
1078	CGTWDNRLSSVIF	1902	TGCGGAACATGGGATAACAGACTGAGTAGTGTGATTTTTTC
1079	CGTWDNRLSAVVF	1903	TGCGGAACATGGGATAACAGGTTGAGTGCTGTGGTCTTC
1080	CGTWDNRLSAGGIF	1904	TGCGGAACATGGGATAACAGGCTGAGTGCTGGTGGGATATTC
1081	CGTWDNRLSAEVF	1905	TGCGGAACATGGGATAACAGACTGAGTGCTGAGGTGTTC
1082	CGTWDNRLRVGVL	1906	TGTGGAACATGGGATAACAGACTGCGTGTTGGGGTTCTC
1083	CGTWDNRLLENVVF	1907	TGCGGAACATGGGATAATCGCCTGCTTGAGAATGTCTTC
1084	CGTWDNNLRAVVF	1908	TGCGGAACATGGGATAACAACCTGCGTGCTGTCTTC
1085	CGTWDNNLRAGVF	1909	TGCGGAACATGGGATAATAACCTGCGTGCTGGAGTGTTC
1086	CGTWDNNLGGGRVF	1910	TGCGGAACATGGGACAACAATTTGGGCGGTGGCCGGGTGTTC
1087	CGTWDNNLGAGVL	1911	TGCGGAACATGGGATAACAACCTGGGTGCTGGCGTCTTC

1088	CGTWDNNLGAGVF	1912	TGCGGAACATGGGATAACAACCTGGGTGCTGGCGTCTTC
1089	CGTWDNLSAAVF	1913	TGCGGAACCTGGGATAACATCCTGAGCGCTGCGGTGTTC
1090	CGTWDNILDAGVF	1914	TGCGGAACCTGGGATAACATCTTGGATGCAGGGGTTTTTC
1091	CGTWDNDLSGWLF	1915	TGCGGAACATGGGATAACGACCTGAGTGGTTGGGTGTTC
1092	CGTWDNDLSAWVF	1916	TGCGGAACATGGGATAACGACCTGAGTGCCTGGGTGTTC
1093	CGTWDLTGGVVF	1917	TGCGGAACATGGGATCTCACCTGGGTGGTGTGGTGTTC
1094	CGTWDLSLSAGVF	1918	TGCGGAACATGGGATCTCAGCCTGAGTGCTGGGGTATTC
1095	CGTWDLSLKEWVF	1919	TGCGGAACATGGGATCTCAGCCTGAAAGAATGGGTGTTC
1096	CGTWDLSLDAVVF	1920	TGCGGAACGTGGGATCTCAGCCTGGATGCTGTTGTTTTTC
1097	CGTWDLKVF	1921	TGCGGAACCTGGGACCTGAAGGTTTTTC
1098	CGTWDKTLVWVF	1922	TGCGGAACATGGGATAAGACTCTGAGTGTTTGGGTGTTC
1099	CGTWDKSLVWVF	1923	TGCGGAACATGGGATAAGAGCCTGAGTGTTTGGGTGTTC
1100	CGTWDKSLSGVVF	1924	TGCGGAACATGGGATAAGAGCCTGAGTGGTGTGGTATTT
1101	CGTWDKSLSDWVF	1925	TGCGGAACATGGGATAAGAGCCTGAGTGATTGGGTGTTC
1102	CGTWDKSLSALVF	1926	TGCGGAACATGGGATAAGAGCCTGAGTGCTTTGGTTTTTC
1103	CGTWDKSLSAGVF	1927	TGCGGAACATGGGATAAGAGCCTGAGTGCTGGCGTCTTC
1104	CGTWDKSLSADVF	1928	TGCGGAACATGGGATAAGAGCCTGAGTGCCGACGTCTTC
1105	CGTWDKRLTIVVF	1929	TGCGGAACATGGGATAAACGCCTGACTATTGTGGTCTTC
1106	CGTWDKRLSAWVL	1930	TGCGGAACATGGGATAAACGCCTGAGTGCCTGGGTGCTC
1107	CGTWDKNLRAVVF	1931	TGCGGAACATGGGATAAGAACCTGCGTGCTGTGGTCTTC
1108	CGTWDITLSGFVF	1932	TGCGGAACATGGGATATCACCTGAGTGGGTTTGTCTTC
1109	CGTWDITLHTGVF	1933	TGCGGAACATGGGATATCACCTGCATACTGGAGTATTC
1110	CGTWDISVTVVF	1934	TGCGGAACATGGGATATCAGTGTGACTGTGGTGTTC
1111	CGTWDISVRGYAF	1935	TGCGGAACATGGGATATCAGTGTGAGGGGTTATGCCTTC
1112	CGTWDISRWVF	1936	TGCGGAACATGGGATATCAGCCGTTGGGTTTTTC
1113	CGTWDISPSAVVF	1937	TGCGGAACATGGGATATCAGCCCGAGTGCTTGGGTGTTC
1114	CGTWDISLSVWVF	1938	TGCGGAACATGGGATATTAGCCTGAGTGTCTGGGTGTTC
1115	CGTWDISLSVVF	1939	TGCGGAACATGGGATATCAGCCTGAGTGTTGTATTC
1116	CGTWDISLSSVVF	1940	TGCGGAACCTGGGATATCAGCCTGAGTTCTGTGGTGTTC
1117	CGTWDISLSHWLF	1941	TGCGGAACATGGGATATCAGCCTGAGTCACTGGTTGTTC
1118	CGTWDISLSGWVF	1942	TGCGGAACATGGGATATCAGTCTGAGTGGTTGGGTGTTC
1119	CGTWDISLSGRVF	1943	TGCGGAACATGGGATATCAGCCTGAGTGGTCGAGTGTTC
1120	CGTWDISLSAWAF	1944	TGCGGAACATGGGACATCAGCCTGAGTGCTTGGGCGTTC
1121	CGTWDISLSAVVF	1945	TGCGGAACATGGGATATCAGCCTGAGTGCTGTGGTTTTTC
1122	CGTWDISLSAVIF	1946	TGCGGGACATGGGACATCAGCCTGAGTGCTGTGATATTC
1123	CGTWDISLSAVF	1947	TGCGGAACATGGGATATCAGCCTGAGTGCTGTGTTC
1124	CGTWDISLSARVF	1948	TGCGGAACATGGGATATCAGCCTGAGTGCCCGGGTGTTC
1125	CGTWDISLSALVF	1949	TGCGGAACATGGGATATCAGCCTGAGTGCCCTGGTGTTC
1126	CGTWDISLSAHVF	1950	TGCGGAACATGGGATATTAGCCTGAGTGCCCATGTCTTC
1127	CGTWDISLSAGVVF	1951	TGCGGAACATGGGATATCAGCCTGAGTGCTGGGGTGGTATTC
1128	CGTWDISLSAGPYVF	1952	TGCGGAACATGGGATATCAGCCTGAGTGCCGGCCCTTATGTCTT C
1129	CGTWDISLSAGGVF	1953	TGCGGCACATGGGATATCAGCCTGAGTGCTGGAGGGGTGTTC
1130	CGTWDISLSAEVF	1954	TGCGGAACATGGGATATCAGCCTGAGTGCTGAGGTTTTTC
1131	CGTWDISLSAAVF	1955	TGCGGAACATGGGATATCAGCCTGAGTGCTGTGTGTTC
1132	CGTWDISLRAVF	1956	TGCGGAACATGGGATATCAGCCTGCGTGCTGTGTTC

1133	CGTWDISLNTGVF	1957	TGCGGAACATGGGATATTAGCCTGAATACTGGGGTGTTTC
1134	CGTWDISLNNYVF	1958	TGCGGAACATGGGATATCAGCCTAAATAATTATGTCTTC
1135	CGTWDISLIAGVF	1959	TGCGGAACATGGGATATCAGCCTAATTGCTGGGGTATTC
1136	CGTWDISLHTWLF	1960	TGCGGAACATGGGATATCAGCCTGCATACTTGGCTGTTTC
1137	CGTWDIRLTDELLF	1961	TGCGGAACATGGGATATCCGCCTGACCGATGAGCTGTTATTC
1138	CGTWDIRLSGFVF	1962	TGCGGAACATGGGATATCAGACTGAGCGGTTTTGTTTTTC
1139	CGTWDINLGAGGLY VF	1963	TGCGGAACATGGGATATCAACCTGGGTGCTGGGGGCCTTTATG TCTTC
1140	CGTWDIILSAEVF	1964	TGCGGAACATGGGATATCATCCTGAGTGCTGAGGTATTC
1141	CGTWDHTLSAVF	1965	TGCGGAACATGGGATCACACCCTGAGTGCTGTCTTC
1142	CGTWDHTLLTVLF	1966	TGCGGAACATGGGACCACACTCTGCTTACTGTGTTATTC
1143	CGTWDHSLTAVVF	1967	TGCGGAACATGGGATCACAGCCTGACTGCTGTGGTATTC
1144	CGTWDHSLTAGIF	1968	TGCGGAACCTGGGATCACAGCCTGACTGCTGGGATATTC
1145	CGTWDHSLSVVLF	1969	TGCGGAACATGGGATCACAGCCTGAGTGTTGTATTATTC
1146	CGTWDHSLSLVF	1970	TGCGGAACATGGGATCACAGCCTGAGTTTGGTATTC
1147	CGTWDHSLSIGVF	1971	TGCGGAACATGGGATCACAGCCTGTCTATTGGGGTTTTTC
1148	CGTWDHSLSAGVF	1972	TGCGGAACATGGGATCACAGCCTGAGTGCTGGGGTGTTTC
1149	CGTWDHSLSAFVF	1973	TGTGGAACCTGGGATCACAGCCTGAGTGCTTTCTGTGTTTC
1150	CGTWDHSLSAAVF	1974	TGCGGAACATGGGATCACAGTCTGAGTGCTGCTGTTTTTC
1151	CGTWDHNLRAVF	1975	TGCGGAACATGGGACCACAATCTGCGTGCTGTCTTC
1152	CGTWDFTLSVGRF	1976	TGCGGGACATGGGATTTTACCCTGAGTGTTGGGCGCTTC
1153	CGTWDFTLSAPVF	1977	TGCGGAACATGGGATTTTACCCTGAGTGCTCCTGTCTTC
1154	CGTWDFSVSAGWVF	1978	TGCGGAACCTGGGATTTTACGCTGAGTGCTGGGTGGGTGTTTC
1155	CGTWDFTSLTTWLF	1979	TGCGGAACCTGGGATTTTCACTTACTACCTGGTTATTC
1156	CGTWDFTSLSVWVF	1980	TGCGGAACATGGGATTTTCACTGAGTGTTTGGGTGTTTC
1157	CGTWDFTSLSTGVF	1981	TGCGGAACATGGGATTTTCACTGAGTACTGGGGTTTTTC
1158	CGTWDFTSLSGVVF	1982	TGCGGCACATGGGATTTTCACTGAGTGGTGTGGTATTC
1159	CGTWDFTSLSGVF	1983	TGCGGAACATGGGATTTTCACTGAGTGTTTTCTGTGTTTC
1160	CGTWDFTSLSAGVF	1984	TGCGGAACATGGGATTTTCACTGAGTGCTGGGGTGTTTC
1161	CGTWDETVRGWVF	1985	TGCGGAACATGGGATGAAACCGTGAGAGGTTGGGTGTTTC
1162	CGTWDESLRSWVF	1986	TGCGGAACATGGGATGAAAGTCTGAGAAGCTGGGTGTTTC
1163	CGTWDERQTDES YVF	1987	TGCGGAACCTGGGATGAGAGGCAGACTGATGAGTCCTATGTCT TC
1164	CGTWDERLVAGQVF	1988	TGCGGAACATGGGATGAGAGACTCGTTGCTGGCCAGGTCTTC
1165	CGTWDERLSPGAFF	1989	TGCGGAACATGGGATGAGAGACTGAGTCTGGAGCTTTTTTC
1166	CGTWDEKVF	1990	TGCGGAACATGGGATGAGAAGGTGTTTC
1167	CGTWDEGQTTDFVVF	1991	TGCGGAACCTGGGATGAAGGCCAGACTACTGATTTCTTTGTCTT C
1168	CGTWDDTLAGVVF	1992	TGCGGAACATGGGATGACACCCTGGCTGGTGTGGTCTTC
1169	CGTWDDRLTSVVF	1993	TGCGGAACATGGGATGACAGGCTGACTTCTGCGGTCTTC
1170	CGTWDDRLFVVVF	1994	TGCGGAACATGGGATGACAGACTGTTTGTGTGGTATTC
1171	CGTWDDNLRGWVF	1995	TGCGGAACATGGGATGATAACCTGAGAGGTTGGGTGTTTC
1172	CGTWDDNLRGVVF	1996	TGCGGAACATGGGATGACAACCTGCGTGGTGTCTGTGTTTC
1173	CGTWDDNLNIGRVF	1997	TGCGGAACCTGGGATGACAATTTGAATATTGGAAGGGTGTTTC
1174	CGTWDDILSAVIF	1998	TGCGGAACATGGGATGACATCCTGAGTGCTGTGATATTC
1175	CGTWDDILRGWVF	1999	TGCGGAACATGGGATGATATCCTGAGAGGTTGGGTGTTTC
1176	CGTWDA TLSPGWLF	2000	TGCGGAACATGGGATGCCACCCTGAGTCTGGGTGGTTATTC

1177	CGTWDASVTSWVF	2001	TGCGGAACATGGGATGCCAGCGTGACTTCTTGGGTGTTC
1178	CGTWDASLTSVVF	2002	TGCGGAACATGGGATGCCAGCCTGACTTCTGTGGTCTTC
1179	CGTWDASLSVWVF	2003	TGCGGAACATGGGATGCCAGCCTGAGTGTTTGGGTGTTC
1180	CGTWDASLSVPWVF	2004	TGCGGAACATGGGATGCCAGCCTGAGTGTTCTTGGGTGTTC
1181	CGTWDASLSVAVF	2005	TGCGGAACATGGGATGCCAGCCTGAGTGTTGGCGGTATTC
1182	CGTWDASLSTWVF	2006	TGCGGAACATGGGATGCCAGCCTGAGTACCTGGGTATTC
1183	CGTWDASLSGVVF	2007	TGCGGAACATGGGATGCCAGCCTGAGTGTTGGTGTTC
1184	CGTWDASLSGGGEF	2008	TGCGGAACATGGGATGCCAGCCTGAGTGTTGGGGGAGAATTC
1185	CGTWDASLSAGVF	2009	TGCGGAACATGGGATGCCAGCCTGAGTGCTGGGGTGTTC
1186	CGTWDASLSAGLF	2010	TGCGGAACATGGGATGCCAGCCTGAGTGCTGGGCTTTTC
1187	CGTWDASLSAEVF	2011	TGTGGCACATGGGATGCCAGCCTGAGTGCTGAAGTCTTC
1188	CGTWDASLSADFWV F	2012	TGCGGAACATGGGATGCCAGCCTGAGTGCTGACTTTTGGGTGTT C
1189	CGTWDASLRVVF	2013	TGCGGAACATGGGATGCCAGCCTGAGAGTCTTCTTC
1190	CGTWDASLRVVL	2014	TGCGGAACATGGGATGCCAGTCTGAGGGCTGTGGTACTC
1191	CGTWDASLNIWVF	2015	TGCGGAACATGGGATGCCAGCCTGAATATTTGGGTTTTTC
1192	CGTWDASLKNLVF	2016	TGCGGGACATGGGATGCCAGCCTGAAGAATCTGGTCTTC
1193	CGTWDASLGAWVF	2017	TGCGGAACATGGGATGCCAGCCTGGGTGCCTGGGTATTC
1194	CGTWDASLGAVVF	2018	TGCGGAACATGGGATGCCAGCCTGGGTGCTGTGGTCTTC
1195	CGTWDASLGAGVF	2019	TGCGGAACATGGGATGCCAGCCTGGGTGCGGGGTCTTC
1196	CGTWDARLSGLYVF	2020	TGCGGAACATGGGATGCTAGGCTGAGTGCCCTTTATGTCTTC
1197	CGTWDARLGGAVF	2021	TGTGGAACCTGGGATGCGAGACTGGGTGGTGCAGTCTTC
1198	CGTWDANLRAGVF	2022	TGCGGAACATGGGATGCCAATCTGCGTGCTGGGGTCTTC
1199	CGTWDALISGWVF	2023	TGCGGAACATGGGATGCTATCATAAGTGGTTGGGTGTTC
1200	CGTWDAGQSVWVF	2024	TGCGGAACATGGGATGCCGGCCAGAGTGTTTGGGTGTTC
1201	CGTWDAGLTGLYVF	2025	TGCGGCACATGGGATGCCGGGCTGACTGGCCTTTATGTCTTC
1202	CGTWDAGLSVYVF	2026	TGCGGAACCTGGGATGCCGGTCTGAGTGTTTATGTCTTC
1203	CGTWDAGLSTGVF	2027	TGCGGGACATGGGATGCCGGCCTGAGTACTGGGGTCTTC
1204	CGTWDAGLSGDVF	2028	TGCGGAACATGGGATGCCGGCCTGAGTGGGGACGTTTC
1205	CGTWDAGLSAGYVF	2029	TGCGGAACATGGGATGCCGGCCTGAGTGCTGGTTATGTCTTC
1206	CGTWDAGLRVWVF	2030	TGCGGAACATGGGATGCCGGCCTGCGTGTTTGGGTGTTC
1207	CGTWDAGLREIF	2031	TGCGGAACATGGGATGCCGGCCTGAGGGAAATTTTC
1208	CGTWASSLSSWVF	2032	TGCGGAACATGGGCCAGCAGCCTGAGTTCTTGGGTGTTC
1209	CGTWAGSLSGHVF	2033	TGCGGAACATGGGCTGGCAGCCTGAGTGGTTCATGTCTTC
1210	CGTWAGSLSAWVF	2034	TGCGGAACATGGGCTGGCAGCCTGAGTGCCGCTTGGGTGTTC
1211	CGTWAGSLNVYWVF	2035	TGCGGAACATGGGCTGGCAGCCTGAATGTTTATTGGGTGTTC
1212	CGTWAGNLRPNWVF	2036	TGCGGAACATGGGCTGGCAACCTGAGACCTAATTGGGTGTTC
1213	CGTRGSLGGAVF	2037	TGCGGAACAAGGGTAGCCTGGGTGGTGCGGTGTTC
1214	CGTRDTLSVPVF	2038	TGCGGAACAAGGGATACCCACCTGAGTGTCCTGGGTGTTC
1215	CGTRDTSLNIEIF	2039	TGCGGAACACGGGATACCCAGCCTCAATATTGAAATCTTC
1216	CGTRDTSLNDVF	2040	TGTGGAACACGGGATACCCAGCCTGAATGATGTCTTC
1217	CGTRDTRLIVVF	2041	TGCGGAACACGGGATACCCGCCTGAGTATTGTGGTTTTTC
1218	CGTRDTILSAEVF	2042	TGCGGCACACGGGATACCATCCTGAGTGCTGAGGTGTTC
1219	CGTRDRSLSGWVF	2043	TGCGGAACACGGGATAGAAGCCTGAGTGTTGGGTGTTC
1220	CGSWYYNVFLF	2044	TGCGGATCATGGTATTACAATGTCTTCTTTTC
1221	CGSWHSSLNLVVF	2045	TGCGGATCTTGGCATAGCAGCCTCAACCTGTGCTCTTC

1222	CGSWGSGLSAPYVF	2046	TGCGGATCATGGGGTAGTGGCCTGAGTGCCCCCTTATGTCTTC
1223	CGSWESGLGAWLF	2047	TGCGGTTCGTGGGAAAGCGGCCTGGGTGCTTGGCTGTTC
1224	CGSWDYGLLLF	2048	TGCGGATCCTGGGATTACGGCCTCCTACTCTTC
1225	CGSWDVSLTAVF	2049	TGCGGTTCATGGGATGTCAGCCTGACTGCTGTTTTC
1226	CGSWDVSLNVGIF	2050	TGCGGATCCTGGGATGTCAGTCTCAATGTTGGCATTTC
1227	CGSWDTTLRAWVF	2051	TGCGGATCATGGGATACCACCCTGCGTGCTTGGGTGTTC
1228	CGSWDTSPVRAWVF	2052	TGCGGCTCGTGGGATACCAGCCTGTCCGTGCTTGGGTGTTC
1229	CGSWDTSLSVWVF	2053	TGCGGATCATGGGATACCAGCCTGAGTGTTTGGGTGTTC
1230	CGSWDTSLSAEVF	2054	TGCGGATCATGGGATACCAGCCTGAGTGCTGAGGTGTTC
1231	CGSWDTSLRAWVF	2055	TGCGGCTCGTGGGATACCAGCCTGCGTGCTTGGGTGTTC
1232	CGSWDTSLRAWAF	2056	TGCGGCTCGTGGGATACCAGCCTGCGTGCTTGGGCGTTC
1233	CGSWDTSLDARLF	2057	TGCGGATCATGGGATACCAGCCTGGATGCTAGGCTGTTC
1234	CGSWDTILLVYVF	2058	TGCGGATCATGGGATACCATCCTGCTTGTCTATGTCTTC
1235	CGSWDRWQAAVF	2059	TGCGGATCATGGGATCGCTGGCAGGCTGCTGTCTTC
1236	CGSWDRSLSGYVF	2060	TGCGGATCATGGGATAGGAGCCTGAGTGGGTATGTCTTC
1237	CGSWDRSLSAYVF	2061	TGCGGATCATGGGATAGAAGCCTGAGTGCTTATGTCTTC
1238	CGSWDRSLSAVVF	2062	TGCGGATCATGGGATAGGAGCCTGAGTGCCGTGGTTTTTC
1239	CGSWDNTLGVVLF	2063	TGCGGATCATGGGATAACACCTTGGGTGTTGTTCTCTTC
1240	CGSWDNRLSTVIF	2064	TGCGGATCGTGGGATAACAGACTAAGTACTGTCATCTTC
1241	CGSWDNRLNTVIF	2065	TGCGGAAGCTGGGATAATCGATTGAACACTGTGATTTTC
1242	CGSWDLSPVRVLVF	2066	TGCGGTTCATGGGATCTCAGCCCTGTACGTGTCCTTGTGTTC
1243	CGSWDLSLSAVVF	2067	TGCGGATCATGGGATCTCAGCCTGAGTGCTGTCTGTTTTTC
1244	CGSWDKNLRAVLF	2068	TGCGGATCATGGGATAAAAACCTGCGTGCTGTGCTGTTC
1245	CGSWDISLSAGVF	2069	TGCGGCTCATGGGATATCAGCCTGAGTGCTGGGGTGTTC
1246	CGSWDIRLSAEVF	2070	TGCGGATCATGGGATATCAGACTGAGTGCAGAGGTCTTC
1247	CGSWDIKLNIGVF	2071	TGCGGATCATGGGACATCAAACCTGAATATTGGGGTATTC
1248	CGSWDFSLNYFVF	2072	TGCGGATCATGGGATTTTCAGTCTCAATTATTTTGTCTTC
1249	CGSWDASLSTEVF	2073	TGCGGATCATGGGATGCCAGCCTGAGTACTGAGGTGTTC
1250	CGSWDAGLRGWVF	2074	TGCGGATCCTGGGATGCCGGCCTGCGTGGCTGGGTTTTTC
1251	CGRWESSLGAVVF	2075	TGCGGAAGATGGGAGAGCAGCCTGGGTGCTGTGGTTTTTC
1252	CGRWDFSLSAYVF	2076	TGCGGAAGATGGGATTTTGTAGTCTGAGTGCTTATGTCTTC
1253	CGQWDNDLSVWVF	2077	TGCGGACAATGGGATAACGACCTGAGTGTTTGGGTGTTC
1254	CGPWHSSVTSGHVL	2078	TGCGGACCCTGGCATAGCAGCGTGACTAGTGGCCACGTGCTC
1255	CGLWDASLSAPTWF	2079	TGCGGATTATGGGATGCCAGCCTGAGTGCTCCTACTTGGGTGTT C
1256	CGIWHTSLSAWVF	2080	TGTGGAATATGGCACACTAGCCTGAGTGCTTGGGTGTTC
1257	CGIWDYSLDTWVF	2081	TGCGGAATATGGGATTACAGCCTGGATACTTGGGTGTTC
1258	CGIWDTSLSAWVF	2082	TGCGGCATATGGGATACCAGCCTGAGTGCTTGGGTGTTC
1259	CGIWDTRL SVYVF	2083	TGCGGAATTTGGGATACCAGGCTGAGTGTTTATGTCTTC
1260	CGIWDTRL SVYIF	2084	TGCGGAATTTGGGATACCAGGCTGAGTGTTTATATCTTC
1261	CGIWDTNLGYLF	2085	TGTGGAATATGGGATACGAATCTGGGTTATCTCTTC
1262	CGIWDTGLSAVVF	2086	TGCGGTATATGGGATACCGGCCTGAGTGCTGTGGTATTC
1263	CGIWDRSLSAWVF	2087	TGCGGAATATGGGATCGCAGCCTGAGTGCTTGGGTGTTT
1264	CGIRDTRL SVYVF	2088	TGCGGAATTCGGGATACCAGGCTGAGTGTTTATGTCTTC
1265	CGGWSSRLGVGPVF	2089	TGCGGAGGATGGAGTAGCAGACTGGGTGTTGGCCCAGTGTTT
1266	CGGWGSGLSAWVF	2090	TGCGGAGGATGGGGTAGCGGCCTGAGTGCTTGGGTGTTC

1267	CGGWDTSLSAWVF	2091	TGCGGAGGATGGGATACCAGCCTGAGTGCTTGGGTGTTC
1268	CGGWDRGLDAWVF	2092	TGCGGAGGATGGGATAGGGGCCTGGATGCTTGGGTTTTC
1269	CGAWRNNVWVF	2093	TGCGGAGCATGGCGTAATAACGTGTGGGTGTTC
1270	CGAWNRRNLNPHSHWVF	2094	TGCGGAGCATGGAACAGGCGCCTGAATCCTCATTCTCATTGGGTGTTC
1271	CGAWHNKLSAVF	2095	TGCGGAGCCTGGCACAACAACTGAGCGCGGTCTTC
1272	CGAWGSSLRASVF	2096	TGCGGAGCATGGGGTAGCAGCCTGAGAGCTAGTGTCTTC
1273	CGAWGSGLSAWVF	2097	TGCGGAGCATGGGGTAGCGGCCTGAGTGCTTGGGTGTTC
1274	CGAWESSLSAPYVF	2098	TGCGGAGCATGGGAAAGTAGCCTGAGTGCCCCCTTATGTCTTC
1275	CGAWESSLNVGLI	2099	TGCGGAGCATGGGAGAGCAGCCTCAATGTTGGACTGATC
1276	CGAWESGRSAGVVF	2100	TGCGGAGCATGGGAGAGCGCCGGAGTGCTGGGGTGGTGTTC
1277	CGAWDYSVSGWVF	2101	TGCGGAGCTTGGGATTACAGTGTGAGTGGTTGGGTGTTC
1278	CGAWDYSLTAGVF	2102	TGCGGAGCATGGGATTACAGCCTGACTGCCGGAGTATTC
1279	CGAWDYRLSAVLF	2103	TGCGGAGCCTGGGATTACAGACTGAGTGCCGTGCTATTC
1280	CGAWDVRLDVGVF	2104	TGCGGAGCGTGGGATGTTCTGTCTGGATGTTGGGGTGTTC
1281	CGAWDTYSYVF	2105	TGCGGAGCATGGGATACCTACAGTTATGTCTTC
1282	CGAWDTTSLSGVVF	2106	TGCGGAGCATGGGATACGACCCTGAGTGGTGTGGTATTC
1283	CGAWDTTSLSAVIF	2107	TGCGGAGCGTGGGATACTACCCTGAGTGCTGTGATATTC
1284	CGAWDTSQGASYVF	2108	TGCGGCGCATGGGATACCAGCCAGGGTGCGTCTTATGTCTTT
1285	CGAWDTSPVRAGVF	2109	TGCGGAGCATGGGATACCAGCCCTGTACGTGCTGGGGTGTTC
1286	CGAWDTSWLWF	2110	TGCGGAGCATGGGATACCAGCCTGTGGCTTTTC
1287	CGAWDTSLTVYVF	2111	TGCGGAGCATGGGATACCAGCCTGACTGTTTATGTCTTC
1288	CGAWDTSLTAGVF	2112	TGCGGAGCATGGGACACCAGTCTGACTGCTGGGGTGTTC
1289	CGAWDTSLSSTVVF	2113	TGCGGAGCTTGGGATACCAGCCTGAGTACTGTGGTTTTTC
1290	CGAWDTSLSRYIF	2114	TGCGGAGCATGGGATACCAGCCTGAGTTCTAGATACATATTC
1291	CGAWDTSLSGYVF	2115	TGCGGAGCATGGGATACCAGCCTGAGTGGTTATGTCTTC
1292	CGAWDTSLSGWVF	2116	TGCGGAGCCTGGGATACCAGCCTGAGTGGCTGGGTGTTC
1293	CGAWDTSLSGVLF	2117	TGCGGAGCATGGGATACCAGTCTGAGTGGTGTGCTATTC
1294	CGAWDTSLSGLVF	2118	TGCGGAGCTTGGGATACCAGCTTGAGTGGTCTTGTTTTC
1295	CGAWDTSLSGFVF	2119	TGCGGAGCTTGGGATACCAGCTTGAGTGGTTTTGTTTTC
1296	CGAWDTSLSGEVF	2120	TGCGGAGCATGGGATACCAGCCTGAGTGGTGAGGTCTTT
1297	CGAWDTSLSDFVF	2121	TGCGGAGCTTGGGATACCAGCTTGAGTGATTTTGTTTTC
1298	CGAWDTSRLTAIF	2122	TGCGGAGCATGGGATACCAGCCTGCGAACTGCGATATTC
1299	CGAWDTSRLRF	2123	TGCGGAGCATGGGATACCAGCCTGCGGCTTTTC
1300	CGAWDTSNLNVHVF	2124	TGCGGAGCATGGGATACCAGCCTGAATGTTTCATGTCTTC
1301	CGAWDTSNLKWVF	2125	TGCGGAGCATGGGATACCAGCCTCAATAAATGGGTGTTC
1302	CGAWDTRLRSARLF	2126	TGCGGAGCATGGGATACCCGCCTCAGTGCGCGGCTGTTC
1303	CGAWDTRLRGFIF	2127	TGCGGAGCATGGGATACCAGACTGAGGGGTTTTATTTTC
1304	CGAWDTNLGNVLL	2128	TGCGGAGCATGGGATACTAATTTGGGGAATGTTCTCCTC
1305	CGAWDTNLGKWVF	2129	TGCGGGGCATGGGATACCAACCTGGGTAAATGGGTTTTC
1306	CGAWDTGLEWYVF	2130	TGCGGAGCATGGGATACCGGCCTTGAGTGGTATGTTTTT
1307	CGAWDRSGLWLF	2131	TGCGGAGCATGGGATAGGACTTCTGGATTGTGGCTTTTC
1308	CGAWDRSLVAGLF	2132	TGCGGAGCGTGGGATCGTAGCCTGGTTGCTGGACTCTTC
1309	CGAWDRSLTVYVF	2133	TGCGGAGCGTGGGATAGAAGCCTGACTGTTTATGTCTTC
1310	CGAWDRSLSGYVF	2134	TGCGGAGCATGGGATAGAAGCCTGAGTGGTTATGTCTTC
1311	CGAWDRSLSAYVF	2135	TGCGGAGCATGGGATAGAAGCCTGAGTGCTTATGTCTTC

1312	CGAWDRSLSAVVF	2136	TGCGGAGCATGGGATAGAAGCCTGAGTGCGGTGGTATTC
1313	CGAWDRSLSAGVF	2137	TGCGGAGCATGGGATCGCAGCCTGAGTGCTGGGGTTTTTC
1314	CGAWDRSLRIVVF	2138	TGCGGAGCGTGGGATCGCAGCCTGCGTATTGTGGTATTC
1315	CGAWDRSLRAYVF	2139	TGCGGAGCATGGGATAGAAGTCTGAGGGCTTACGTCTTC
1316	CGAWDRSLNVWLF	2140	TGCGGAGCATGGGATAGAAGTCTGAATGTTTGGCTGTTC
1317	CGAWDRGLNVGWLF	2141	TGCGGCGCCTGGGATAGGGGCCTGAATGTCGGTTGGCTTTTC
1318	CGAWDNRLSILAF	2142	TGCGGCGCATGGGATAATAGACTGAGTATTTGGCCTTC
1319	CGAWDNDLTAYVF	2143	TGCGGAGCTTGGGATAATGACCTGACAGCTTATGTCTTC
1320	CGAWDFSLTPLF	2144	TGCGGGGCATGGGATTTTCAGCCTGACTCCTCTCTTC
1321	CGAWDDYRGVSIYV F	2145	TGCGGAGCCTGGGATGACTATCGGGGTGTGAGTATTTATGTCTT C
1322	CGAWDDRPSSAVVF	2146	TGTGGAGCATGGGATGACCGGCCTTCGAGTGCCGTGGTTTTTC
1323	CGAWDDRLTVVVF	2147	TGCGGAGCATGGGATGACAGACTGACTGTCTGTTGTTTTTC
1324	CGAWDDRLGAVF	2148	TGCGGAGCGTGGGATGACAGGCTGGGTGCTGTGTTC
1325	CGAWDASLNPGRF	2149	TGCGGAGCGTGGGATGCCAGCCTGAATCCTGGCCGGGCATTC
1326	CGAWDAGLREIF	2150	TGCGGAGCATGGGATGCCGGCCTGAGGGAAATTTTC
1327	CGAWAGSPSPWVF	2151	TGCGGAGCTTGGGCTGGCAGTCCGAGTCCTTGGGTTTTTC
1328	CGAFDTTSLAGVF	2152	TGCGGAGCATTCGACACCACCCTGAGTGCTGGCGTTTTTC
1329	CETWESSLSVGVF	2153	TGCGAAACATGGGAGAGCAGCCTGAGTGTTGGGGTCTTC
1330	CETWESSLRVWVF	2154	TGCGAAACATGGGAAAGCAGCCTGAGGGTTTGGGTGTTC
1331	CETWDTSLSGGVF	2155	TGCGAAACGTGGGATACCAGCCTGAGTGGTGGGGTGTTC
1332	CETWDTSLSDFYVF	2156	TGCGAAACATGGGATACCAGCCTGAGTGACTTTTATGTCTTC
1333	CETWDTSLSALF	2157	TGCGAAACATGGGATACCAGCCTGAGTGCCCTCTTC
1334	CETWDTSLRAEVF	2158	TGCGAAACATGGGATACCAGCCTGCGTGCTGAAGTCTTC
1335	CETWDTSLNVVVF	2159	TGCGAAACATGGGATACCAGCCTGAATGTTGTGGTATTC
1336	CETWDTSLGAVVF	2160	TGCGAAACATGGGATACCAGCCTGGGTGCCGTGGTGTTC
1337	CETWDRSLSGVVF	2161	TGCGAAACATGGGATAGAAGCCTGAGTGGTGTGGTATTC
1338	CETWDRSLSAVVF	2162	TGCGAAACATGGGATAGGAGCCTGAGTGCTTGGGTGTTT
1339	CETWDRSLSAVVF	2163	TGCGAAACATGGGATCGCAGCCTGAGTGCTGTGGTCTTC
1340	CETWDRGLSVVVF	2164	TGCGAGACGTGGGATAGAGGCCTGAGTGTTGTGGTTTTTC
1341	CETWDRGLSAVVF	2165	TGCGAAACATGGGATAGGGGCCTGAGTGCACTGGTATTC
1342	CETWDHTLSVVIF	2166	TGCGAAACATGGGATCACACCCTGAGTGTTGTGATATTC
1343	CETWDASLTVVLF	2167	TGCGAAACATGGGATGCCAGCCTGACTGTTGTGTTATTC
1344	CETWDASLSAGVF	2168	TGCGAAACATGGGATGCCAGCCTGAGTGCTGGGGTGTTC
1345	CETWDAGLSEVVF	2169	TGCGAAACGTGGGATGCCGGCCTGAGTGAGGTGGTGTTC
1346	CETFDTSLSVVVF	2170	TGCGAAACATTTGATACCAGCCTGAGTGTTGTAGTCTTC
1347	CETFDTSLNIVVF	2171	TGCGAAACATTTGATACCAGCCTAAATATTGTAGTCTTT
1348	CESWDRSRIGVVF	2172	TGCGAATCATGGGATAGAAGCCGGATTGGTGTGGTCTTC
1349	CESWDRSLSARVY	2173	TGCGAAAGTTGGGACAGGAGTCTGAGTGCCCCGGGTGTAC
1350	CESWDRSLRAVVF	2174	TGCGAATCCTGGGATAGGAGCCTGCGTGCCGTGGTCTTC
1351	CESWDRSLIVVF	2175	TGCGAATCTTGGGATCGTAGTTTGATTGTGGTGTTC
1352	CESWDNNLNEVVF	2176	TGCGAAAGTTGGGATAACAATTTAAATGAGGTGGTTTTTC
1353	CEIWESSPSADDLVF	2177	TGCGAAATATGGGAGAGCAGCCCGAGTGCTGACGATTGGTGT TC
1354	CEAWDTSLSGAVF	2178	TGCGAAGCATGGGATACCAGCCTGAGTGGTGCGGTGTTC
1355	CEAWDTSLSAGVF	2179	TGCGAAGCATGGGATACCAGCCTGAGTGCCGGGGTGTTC

1356	CEAWDTSLGGVF	2180	TGCGAAGCATGGGATACCAGCCTGGGTGGTGGGGTGTTTC
1357	CEAWDRSLTGSF	2181	TGCGAAGCATGGGATCGCAGCCTGACTGGTAGCCTGTTC
1358	CEAWDRGLSAVVF	2182	TGCGAAGCGTGGGATAGGGGCCTGAGTGCAGTGGTATTC
1359	CEAWDNILSTVVF	2183	TGCGAAGCCTGGGATAACATCCTGAGTACTGTGGTGTTC
1360	CEAWDISLSAGVF	2184	TGCGAAGCATGGGACATCAGCCTGAGTGCTGGGGTGTTTC
1361	CEAWDADLSGAVF	2185	TGCGAAGCATGGGATGCCGACCTGAGTGGTGCGGTGTTC
1362	CATWTGSFRTGHYV F	2186	TGCGCAACATGGGACTGGTAGTTTCAGAACTGGCCATTATGTCTT C
1363	CATWSSSPRGWVF	2187	TGCGCAACATGGGAGTAGCAGTCCCAGGGGGTGGGTGTTC
1364	CATWHYSLSAGRVF	2188	TGCGCAACATGGCATTACAGCCTGAGTGCTGGCCGAGTGTTC
1365	CATWHTLSIVQF	2189	TGCGCAACATGGCATACCAGCCTGAGTATTGTGCAGTTC
1366	CATWHSTLSADVLF	2190	TGCGCAACATGGCATAGCACCTGAGTGCTGATGTGCTTTTC
1367	CATWHSSLSAGRLF	2191	TGCGCAACATGGCATAGCAGCCTGAGTGCTGGCCGACTCTTC
1368	CATWHIARSAWVF	2192	TGCGCAACATGGCATATCGCTCGGAGTGCCTGGGTGTTC
1369	CATWGSSQSAVVF	2193	TGCGCAACATGGGGTAGTAGTCAGAGTGCCGTGGTATTC
1370	CATWGSSLSAGGVF	2194	TGCGCAACATGGGGTAGCAGCCTGAGTGCTGGGGGTGTTTTTC
1371	CATWEYSLSVVLVF	2195	TGTGCAACATGGGAATACAGCCTGAGTGTTGTGCTGTTC
1372	CATWETTRRASVVF	2196	TGCGCAACATGGGAGACCACCCGACGTGCCTCTTTGTCTTC
1373	CATWETSLNVYVF	2197	TGCGCAACATGGGAGACCAGCCTGAATGTTTATGTCTTC
1374	CATWETSLNVVVF	2198	TGCGCAACATGGGAAACTAGCCTGAATGTTGTGGTCTTC
1375	CATWETSLNLYVF	2199	TGCGCAACATGGGAGACCAGCCTGAATCTTTATGTCTTC
1376	CATWETGLSAGEVF	2200	TGCGCAACATGGGAGACTGGCCTAAGTGCTGGAGAGGTGTTC
1377	CATWESTLSVVVF	2201	TGCGCGACGTGGGAGAGTACCCTAAGTGTTGTGGTTTTTC
1378	CATWESSLSIFVF	2202	TGCGCAACGTGGGAGAGCAGCCTGAGTATTTTTGTCTTC
1379	CATWESSLNTFYVF	2203	TGCGCAACATGGGAAAGCAGCCTCAACACTTTTTATGTCTTC
1380	CATWESRVDTRGLLF	2204	TGCGCAACATGGGAGAGTAGGGTGGTACTCGAGGGTTGTTAT TC
1381	CATWESGLSGAGVF	2205	TGCGCAACATGGGAGAGCGGCCTGAGTGGTGCGGGGGTGTTC
1382	CATWEGSLNTFYVF	2206	TGCGCAACATGGGAAGGCAGCCTCAACACTTTTTATGTCTTC
1383	CATWDYSLSAVVF	2207	TGCGCAACTTGGGATTATAGCCTGAGTGCTGTGGTGTTC
1384	CATWDYRLSIVVF	2208	TGCGCAACATGGGATTACAGACTGAGTATTGTGGTATTC
1385	CATWDYNLGAADF	2209	TGCGCAACATGGGATTATAACCTGGGAGCTGCGGTGTTC
1386	CATWDVTLGVLHF	2210	TGCGCCACATGGGATGTCACCCTGGGTGTCTGCATTTC
1387	CATWDTTLSVWVF	2211	TGCGCAACATGGGATACAACACTGAGTGCTGTGGTCTTC
1388	CATWDTTLSVVLVF	2212	TGCGCAACATGGGATACCACCCTGAGTGTAAGTACTTTTC
1389	CATWDTTLSVEVF	2213	TGCGCAACATGGGATACCACCCTGAGTGTTGAGGTCTTC
1390	CATWDTSPSLSGFWV F	2214	TGCGCAACATGGGATACCAGCCCCAGCCTGAGTGGTTTTTGGG TGTTTC
1391	CATWDTSLTGVVF	2215	TGCGCAACATGGGATACCAGCCTGACTGGTGTGGTATTC
1392	CATWDTSLTGAVF	2216	TGCGCAACATGGGATACCAGCCTGACTGGTGCGGTGTTC
1393	CATWDTSLTAWVF	2217	TGCGCAACATGGGATACCAGCCTGACTGCCTGGGTATTC
1394	CATWDTSLTAVVF	2218	TGCGCAACATGGGATACCAGCCTGACTGCTGTGGTTTTTC
1395	CATWDTSLTAKVF	2219	TGCGCAACATGGGATACTAGCCTGACTGCTAAGGTGTTC
1396	CATWDTSLSVVVF	2220	TGCGCAACATGGGACACCAGCCTGAGTGTTGTGGTTTTTC
1397	CATWDTSLSVGVF	2221	TGCGCTACTTGGGATACCAGCCTGAGTGTTGGGGTATTT
1398	CATWDTSLSWVF	2222	TGCGCAACATGGGATACCAGCCTGAGTCTTGGGTGTTC
1399	CATWDTSLSGGVL	2223	TGCGCAACATGGGATACCAGCCTGAGTGGTGGGGTACTC

1400	CATWDTSLSGGVF	2224	TGCGCAACATGGGATACCAGCCTGAGTGGTGGGGTGTTTC
1401	CATWDTSLSGGRVF	2225	TGCGCAACATGGGATACCAGCCTGAGTGGTGGCCGAGTGTTTC
1402	CATWDTSLSGDRVVF	2226	TGCGCAACATGGGATACCAGCCTGAGTGGTGACCGAGTGTTTC
1403	CATWDTSLSSEGVF	2227	TGCGCAACGTGGGATACTAGCCTGAGTGAAGGGGTGTTTC
1404	CATWDTSLSAVVL	2228	TGCGCAACCTGGGATACCAGCCTGAGTGCCGTGGTGCTC
1405	CATWDTSLSAVF	2229	TGCGCAACATGGGATACCAGCCTGAGTGCTGTCTTC
1406	CATWDTSLSARVF	2230	TGCGCGACATGGGATACCAGCCTGAGTGCTCGGGTGTTTC
1407	CATWDTSLSALF	2231	TGCGCAACATGGGATACCAGCCTGAGTGCCTTATTC
1408	CATWDTSLSAHVF	2232	TGCGCAACATGGGATACCAGCCTGAGTGCTCATGTCTTC
1409	CATWDTSLSAGRVF	2233	TGCGCAACATGGGATACCAGCCTGAGTGCTGGCCGGGTGTTTC
1410	CATWDTSLSAEVF	2234	TGCGCAACATGGGATACCAGCCTGAGTGCGGAGGTCTTC
1411	CATWDTSLSADAGG GVF	2235	TGCGCAACATGGGATACCAGCCTGAGTGCTGATGCTGGTGGGG GGGTCTTC
1412	CATWDTSLSRVVVF	2236	TGCGCAACATGGGATACCAGCCTGCGTGCTGTGGTATTC
1413	CATWDTSLSRGVF	2237	TGCGCAACATGGGATACCAGCCTGAGAGGGGTGTTTC
1414	CATWDTSLSPAWVF	2238	TGCGCAACATGGGATACCAGCCTGCCTGCGTGGGTGTTTC
1415	CATWDTSLSNVGVF	2239	TGTGCAACATGGGATACCAGCCTGAATGTTGGGGTATTC
1416	CATWDTSLSGIVLF	2240	TGCGCAACATGGGATACCAGCCTGGGTATTGTGTTATTT
1417	CATWDTSLSGARVVVF	2241	TGCGCAACATGGGACACCAGCCTGGGTGCGCGTGTGGTCTTC
1418	CATWDTSLSGALF	2242	TGTGCAACGTGGGATACCAGTCTAGGTGCCTTGTTTC
1419	CATWDTSLSATGLF	2243	TGCGCAACATGGGATACCAGCCTGGCGACTGGACTGTTC
1420	CATWDTSLSAAWVF	2244	TGCGCAACATGGGATACCAGCCTGGCTGCCTGGGTATTC
1421	CATWDTRL SAVVF	2245	TGCGCAACCTGGGATACCAGGCTGAGTGCTGTGGTCTTC
1422	CATWDTRL SAGVF	2246	TGCGCAACATGGGATACCAGGCTGAGTGCTGGGGTGTTTC
1423	CATWDTRL LITVF	2247	TGTGCAACGTGGGACACACGTCTACTTATTACGGTTTTTC
1424	CATWDTRL SVELF	2248	TGCGCAACATGGGACACCCTCCTGAGTGTTGAACCTCTTC
1425	CATWDTRGRNPHVVVF	2249	TGCGCAACATGGGATACTGGCCGCAATCCTCATGTGGTCTTC
1426	CATWDTRL SSVLF	2250	TGCGCAACATGGGATACCGCCTGTCTTCGGTGTTGTTTC
1427	CATWDTRL SAVF	2251	TGCGCAACGTGGGATACCGCCTGAGTGCGGTTTTTC
1428	CATWDTRL SIGVF	2252	TGCGCTACGTGGGATAGGACCCTGAGTATTGGAGTCTTC
1429	CATWDRSVTAVLF	2253	TGCGCAACGTGGGATCGCAGTGTGACTGCTGTGCTCTTC
1430	CATWDRSLSGVVVF	2254	TGCGCAACCTGGGATAGGAGCCTGAGTGGTGTGGTGTTTC
1431	CATWDRSL SAVVF	2255	TGCGCAACATGGGATAGAAGCCTGAGTGCTGTGGTCTTC
1432	CATWDRSL SAVPWV F	2256	TGCGCAACATGGGATAGAAGCCTGAGTGCTGTTCTTGGGTGTT C
1433	CATWDRSL SAGVF	2257	TGCGCAACATGGGATCGCAGCCTGAGTGCTGGGGTGTTTC
1434	CATWDRSL RAGVF	2258	TGCGCAACGTGGGATAGGAGCCTGCGTGCTGGGGTGTTTC
1435	CATWDRSL NVYVL	2259	TGCGCAACATGGGATCGCAGTCTGAATGTTTATGTCCTC
1436	CATWDRILSAEVF	2260	TGCGCAACGTGGGATCGCATCCTGAGCGCTGAGGTGTTTC
1437	CATWDRGLSTGVF	2261	TGCGCAACGTGGGATAGAGGCCTGAGTACTGGGGTGTTTC
1438	CATWDNYLGA AVF	2262	TGCGCAACATGGGATAACTACCTGGGTGCTGCCGTGTTTC
1439	CATWDNTPSNIVVF	2263	TGCGCAACATGGGATAACACGCCTTCGAATATTGTGGTATTC
1440	CATWDNTRL SVWVF	2264	TGCGCAACATGGGATAATACACTGAGTGTGTGGGTCTTC
1441	CATWDNTRL SVNWVF	2265	TGCGCAACATGGGATAACACCCTGAGTGTC AATTGGGTGTTTC
1442	CATWDNTRL NVFYVF	2266	TGCGCAACCTGGGATAACACACTGAATGTCTTTTATGTTTTTC
1443	CATWDNTRL SSVVF	2267	TGTGCGACATGGGATAATCGGCTCAGTTCTGTGGTCTTC

1444	CATWDNRLSAGVL	2268	TGCGCAACATGGGATAACCGCCTGAGTGCTGGGGTGCTC
1445	CATWDNRLSAGVF	2269	TGCGCAACGTGGGATAACAGGCTGAGTGCTGGGGTGTTT
1446	CATWDNRDWVF	2270	TGCGCAACATGGGATAACAGGGATTGGGTCTTC
1447	CATWDNNLGAGVF	2271	TGCGCAACATGGGATAACAACCTGGGTGCTGGGGTGTTT
1448	CATWDNKLTSAGVF	2272	TGCGCAACATGGGATAACAAGCTGACTTCTGGGGTGTTT
1449	CATWDNLSAWVF	2273	TGCGCAACATGGGATAACATCCTGAGTGCCTGGGGTGTTT
1450	CATWDNDIHSGLF	2274	TGCGCAACCTGGGACAACGATATACATTCTGGGGTGTTT
1451	CATWDLSLSALF	2275	TGCGCAACTTGGGATCTCAGCCTGAGTGCCCTGTTT
1452	CATWDITLSAEVF	2276	TGCGCAACATGGGATATCACCTGAGTGCTGAGGTGTTT
1453	CATWDISPSAGGVF	2277	TGCGCAACGTGGGATATCAGCCCGAGTGCTGGCGGGGTGTTT
1454	CATWDISLSTGRAVF	2278	TGCGCAACATGGGATATCAGTCTAAGTACTGGCCGGGTGTTT
1455	CATWDISLSQVF	2279	TGCGCAACATGGGATATCAGTCTGAGTCAGGTATTC
1456	CATWDIRLSSGVF	2280	TGCGCAACATGGGATATCAGGCTGAGTAGTGGAGTGTTT
1457	CATWDIGPSAGGVF	2281	TGCGCAACGTGGGATATCGGCCCGAGTGCTGGCGGGGTGTTT
1458	CATWDHSRAGVLF	2282	TGCGCAACATGGGATCACAGCCGGGTGTTTGCTATTC
1459	CATWDHSPSVGEVF	2283	TGCGCAACATGGGATCACAGTCCGAGTGTTGGAGAAGTCTTC
1460	CATWDHSLRVGVF	2284	TGCGCAACATGGGATCACAGCCTGCGTGTTGGGGTGTTT
1461	CATWDHSLNIGVF	2285	TGCGCAACATGGGATCACAGCCTGAACATTGGGGTGTTT
1462	CATWDHSLGLWAF	2286	TGCGCAACATGGGATCACAGCCTGGGTCTTTGGGCATTC
1463	CATWDHNLRLVF	2287	TGCGCCACATGGGATCACAACTGCGTCTTGTTTTC
1464	CATWDHILASGVF	2288	TGCGCGACTTGGGATCACATCCTGGCTTCTGGGGTGTTT
1465	CATWDFSLSVWVF	2289	TGCGCAACATGGGATTTCAGCCTGAGTGTTTGGGTGTTT
1466	CATWDFSLSAWVF	2290	TGCGCAACATGGGATTTCAGCCTGAGTGCTTGGGTGTTT
1467	CATWDDTLTAGVF	2291	TGCGCAACATGGGATGACACCCTCACTGCTGGTGTTT
1468	CATWDDRLSAVLF	2292	TGCGCAACATGGGACGACAGGCTGAGTGCTGTGCTTTTC
1469	CATWDDRLDAAVF	2293	TGCGCAACATGGGATGACAGGCTGGATGCTGCGGTGTTT
1470	CATWDATLNTGVF	2294	TGCGCAACATGGGATGCGACCCTGAATACTGGGGTGTTT
1471	CATWDASLSVWLL	2295	TGCGCAACATGGGATGCCAGCCTGAGTGTTTGGCTGCTC
1472	CATWDASLSGGVF	2296	TGCGCGACATGGGATGCCAGCCTGAGTGGTGGGGTGTTT
1473	CATRDTTLSAVLF	2297	TGCGCAACACGGGATACCACCCTCAGCGCCGTTCTGTTT
1474	CATLGSSLSLWVF	2298	TGCGCTACATTGGGTAGTAGCCTGAGTCTCTGGGTGTTT
1475	CATIETSLPAWVF	2299	TGCGCAACAATCGAACTAGCCTGCCTGCCTGGGTATTC
1476	CATGDRSLTVEVF	2300	TGCGCAACAGGGGACAGAAGCCTGACTGTTGAGGTATTC
1477	CATGDLGLTIVF	2301	TGCGCTACAGGGGATCTCGGCCTGACCATAGTCTTC
1478	CASWDYRGRSGWVF	2302	TGCGCATCATGGGATTACAGGGGGAGATCTGGTTGGGGTGTTT
1479	CASWDTTLNVGVF	2303	TGCGCATCATGGGATACCACCCTGAATGTTGGGGTGTTT
1480	CASWDTTLGFVLF	2304	TGCGCTTCATGGGATACCACCCTGGGTTTTGTGTTATTC
1481	CASWDTSLSGGYVF	2305	TGCGCATCATGGGATACCAGCCTGAGTGGTGGTTATGTCTTC
1482	CASWDTSRLAGVF	2306	TGCGCATCATGGGATACCAGCCTCCGTGCTGGGGTGTTT
1483	CASWDTSLGAGVF	2307	TGCGCATCATGGGATACCAGCCTGGGTGCTGGGGTGTTT
1484	CASWDRGLSAVVF	2308	TGCGCATCATGGGACAGAGGCCTGAGTGCAAGTGGTGTTT
1485	CASWDNVLRGVVF	2309	TGTGCTAGTTGGGATAACGTCCTGCGTGGTGTTGTTT
1486	CASWDNRLTAVVF	2310	TGCGCGTCATGGGATAACAGGCTGACTGCCGTGGTTTTTC
1487	CASWDASLSVAF	2311	TGCGCATCATGGGATGCAAGCCTGTCCGTGCTTTTC
1488	CASWDAGLSSYVF	2312	TGCGCTTCGTGGGATGCCGGCCTGAGTTCTTATGTCTTC

1489	CASGDTSLSGVIF	2313	TGCGCATCCGGGGATACCAGCCTGAGTGGTGTGATATTC
1490	CARWHTSLSIWVF	2314	TGCGCAAGATGGCATACGAGCCTAAGTATTTGGGTCTTC
1491	CAIWDGTGLSPGQVAF	2315	TGCGCAATATGGGATACCGGCCTGAGTCCTGGCCAAGTTGCCTTC
1492	CAAWHSLGLPVPF	2316	TGCGCAGCATGGCATAGCGGCCTGGGTCTCCCGGTCTTC
1493	CAAWDYSLSAGVF	2317	TGCGCAGCATGGGATTACAGCCTGAGTGCTGGGGTGTTC
1494	CAAWDTTLRVRLF	2318	TGCGCAGCCTGGGATACTACCCTGCGTGTTAGGCTGTTC
1495	CAAWDTSLTAWVF	2319	TGCGCAGCATGGGATACCAGCCTGACTGCCTGGGTTTTTC
1496	CAAWDTSLSGGVF	2320	TGCGCAGCATGGGATACCAGCTTGAGTGGTGGGGTGTTC
1497	CAAWDTSLSGEAVF	2321	TGCGCAGCATGGGATACCAGCCTGAGTGGCGAGGCTGTGTTC
1498	CAAWDTSLSGAVF	2322	TGCGCAGCATGGGATACCAGCTTGAGTGGTGGCGTGTTC
1499	CAAWDTSLSAWVF	2323	TGCGCAGCATGGGATACCAGCCTGAGTGCCTGGGTGTTC
1500	CAAWDTSLSAGVF	2324	TGCGCAGCATGGGATACCAGCCTGAGTGCTGGGGTATTC
1501	CAAWDTSLDTYVF	2325	TGCGCAGCATGGGATACCAGCCTGGATACTTATGTCTTC
1502	CAAWDTRLSGVLF	2326	TGCGCTGCATGGGATACCCGTCTGAGTGGTGTGTTATTC
1503	CAAWDTRLSAGVF	2327	TGCGCAGCATGGGATACCAGGCTGAGTGCTGGGGTGTTC
1504	CAAWDRSLSTGVF	2328	TGCGCAGCATGGGATCGCAGTCTGAGTACTGGAGTTTTTC
1505	CAAWDIRRSVLF	2329	TGCGCAGCGTGGGATATCCGCCGGTCTGTCTTTTC
1506	CAAWDHTQRLSF	2330	TGCGCTGCGTGGGATCACACTCAGCGTCTTTCTTC
1507	CAAWDHLSAGQVF	2331	TGCGCAGCATGGGATCACAGCCTGAGTGCTGGCCAGGTGTTC
1508	CAAVDTGLKEWVF	2332	TGCGCAGCAGTCGATACTGGTCTGAAAGAATGGGTGTTC

[00249] The CDRs were prescreened to contain no amino acid liabilities, cryptic splice sites or nucleotide restriction sites. The CDR variation was observed in at least two individuals and comprises the near-germline space of single, double and triple mutations. The order of assembly is seen in **FIG. 12C**.

[00250] The VH domains that were designed include IGHV1-69 and IGHV3-30. Each of two heavy chain VH domains are assembled with their respective invariant 4 framework elements (FW1, FW2, FW3, FW4) and variable 3 CDR (H1, H2, H3) elements. For IGHV1-69, 417 variants were designed for H1 and 258 variants were designed for H2. For IGHV3-30, 535 variants were designed for H1 and 165 variants were designed for H2. For the CDR H3, the same cassette was used in both IGHV1-69 and IGHV3-30 since both designed use an identical FW4, and because the edge of FW3 is also identical for both IGHV1-69 and IGHV3-30. The CDR H3 comprises an N-terminus and C-terminus element that are combinatorially joined to a central middle element to generate 1×10^{10} diversity. The N-terminal and middle element overlap with a “GGG” glycine codon. The middle and C-terminal element overlap with a “GGT” glycine codon. The CDR H3 comprises 5 subpools that were assembled separately. The various N-terminus and C-terminus elements comprise sequences as seen in **Table 10**.

Table 10. Sequences for N-terminus and C-terminus elements

Element	SEQ ID NO	Sequence
Stem A	2333	CARDLRELECEEWT XXX SRGPCVDPRGVAGSFDVW
Stem B	2334	CARDMYYYDF XXX EVVPADDAFDIW
Stem C	2335	CARDGRGSLPRPKGGP XXX YDSSEDSGGAFDIW
Stem D	2336	CARANQHF XXX GYHYYGMDVW
Stem E	2337	CAKHMSMQ XXX RADLVGDAFDVW

[00251] Example 8. GPCR Binding Protein Functionality

[00252] For a GPCR binding protein, the top 100 - 200 scFvs from phage-selections were converted to full-length immunoglobulins. After immunoglobulin conversion, the clones were transiently transfected in ExpiCHO to produce immunoglobulins. Kingfisher and Hamilton were used for batch IgG purifications followed by lab-chip to collect purity data for all purified immunoglobulins. High yields and purities were obtained from 10 mL cultures as seen in **FIG. 13** and **Table 11**.

Table 11. Immunoglobulin Purity Percentage

Name	IgG % Purity
mAb1	100
mAb2	100
mAb3	100
mAb4	100
mAb5	98
mAb6	100
mAb7	97
mAb8	100
mAb9	100
mAb10	100
mAb11	100
mAb12	100
mAb13	100
mAb14	100
mAb15	100

[00253] Stable cell lines expressing GPCR targets were then generated. **FIG. 14** shows target expression was confirmed by FACS. Cells expressing > 80% of the target were then directly used for cell-based selections. Five rounds of selections were carried out against cells overexpressing target of interest. 10^8 cells were used for each round of selection. Before selection on target expressing cells, phage from each round was first depleted on 10^8 CHO background cells.

Stringency of selections was increased by increasing the number of washes in subsequent rounds of selection. Enrichment ratios were monitored for each round of selection.

[00254] Purified IgGs were tested for cell-binding affinity using FACS (**FIGS. 15A-15C**) and cAMP activity (**FIG. 15D**). Allosteric inhibition was observed.

[00255] Purified IgGs were tested using BVP ELISA. As seen in **FIG. 16** BVP ELISA showed some clones comprising BVP scores comparable to comparator antibodies.

[00256] Example 9: VHH Libraries

[00257] Synthetic VHH libraries were developed. For the 'VHH Ratio' library with tailored CDR diversity, 2391 VHH sequences (iCAN database) were aligned using Clustal Omega to determine the consensus at each position and the framework was derived from the consensus at each position. The CDRs of all of the 2391 sequences were analyzed for position-specific variation, and this diversity was introduced in the library design. For the 'VHH Shuffle' library with shuffled CDR diversity, the iCAN database was scanned for unique CDRs in the nanobody sequences. 1239 unique CDR1's, 1600 unique CDR2's, and 1608 unique CDR3's were identified and the framework was derived from the consensus at each framework position amongst the 2391 sequences in the iCAN database. Each of the unique CDR's was individually synthesized and shuffled in the consensus framework to generate a library with theoretical diversity of 3.2×10^9 . The library was then cloned in the phagemid vector using restriction enzyme digest. For the 'VHH hShuffle' library (a synthetic "human" VHH library with shuffled CDR diversity), the iCAN database was scanned for unique CDRs in the nanobody sequences. 1239 unique CDR1's, 1600 unique CDR2's, and 1608 unique CDR3's were identified and framework 1, 3, and 4 was derived from the human germline DP-47 framework. Framework 2 was derived from the consensus at each framework position amongst the 2391 sequences in the iCAN database. Each of the unique CDR's was individually synthesized and shuffled in the partially humanized framework using the NUGE tool to generate a library with theoretical diversity of 3.2×10^9 . The library was then cloned in the phagemid vector using the NUGE tool.

[00258] The Catterra SPR system was used to assess binding affinity and affinity distribution for VHH-Fc variants. VHH-Fc demonstrate a range of affinities for TIGIT, with a low end of 12 nM K_D and a high end of 1685 nM K_D (highlighted in top row; sixth from left-hand side and third from right-hand side, respectively). **Table 12** provides specific values for the VHH-Fc clones for ELISA, Protein A (mg/ml), and K_D (nM).

Table 12.

Clone	ELISA	Library	ProA (mg/ml)	K_D (nM)
31-1	5.7	VHH hShuffle	0.29	12
31-6	9.6	VHH hShuffle	0.29	14

31-26	5.1	VHH hShuffle	0.31	19
30-30	8.0	VHH Shuffle	0.11	23
31-32	8.0	VHH hShuffle	0.25	27
29-10	5.0	VHH Ratio	0.19	32
29-7	7.3	VHH Ratio	0.28	41
30-43	13.5	VHH Shuffle	0.18	44
31-8	12.7	VHH hShuffle	0.29	45
31-56	11.7	VHH hShuffle	0.26	46
30-52	4.2	VHH Shuffle	0.22	49
31-47	8.8	VHH hShuffle	0.23	53
30-15	9.3	VHH Shuffle	0.26	55
30-54	5.5	VHH Shuffle	0.30	58
30-49	10.3	VHH Shuffle	0.26	62
29-22	3.4	VHH Ratio	0.27	65
29-30	9.2	VHH Ratio	0.28	65
31-35	5.7	VHH hShuffle	0.24	66
29-1	10.4	VHH Ratio	0.09	68
29-6	6.8	VHH Ratio	0.29	69
31-34	6.0	VHH hShuffle	0.32	70
29-12	6.2	VHH Ratio	0.23	70
30-1	5.4	VHH Shuffle	0.39	71
29-33	3.9	VHH Ratio	0.15	74
30-20	4.6	VHH Shuffle	0.19	74
31-20	6.6	VHH hShuffle	0.37	74
31-24	3.1	VHH hShuffle	0.15	75
30-14	9.9	VHH Shuffle	0.19	75
30-53	7.6	VHH Shuffle	0.24	78
31-39	9.9	VHH hShuffle	0.32	78
29-18	10.9	VHH Ratio	0.19	78
30-9	8.0	VHH Shuffle	0.40	79
29-34	8.6	VHH Ratio	0.21	80
-29-27	8.6	VHH Ratio	0.18	82
29-20	5.9	VHH Ratio	0.26	83
30-55	6.0	VHH Shuffle	0.41	85
30-39	6.1	VHH Shuffle	0.07	88
31-15	6.2	VHH hShuffle	0.32	88
29-21	4.3	VHH Ratio	0.23	88
29-37	5.3	VHH Ratio	0.26	89
29-40	6.6	VHH Ratio	0.31	90
31-30	3.2	VHH hShuffle	0.33	93
31-10	12.3	VHH hShuffle	0.31	94
29-3	13.6	VHH Ratio	0.11	94
30-57	5.2	VHH Shuffle	0.24	95
29-31	4.4	VHH Ratio	0.18	96
31-27	8.1	VHH hShuffle	0.31	96
31-33	6.0	VHH hShuffle	0.32	96
30-40	7.1	VHH Shuffle	0.21	99
31-18	4.1	VHH hShuffle	0.36	99
30-5	9.3	VHH Shuffle	0.05	100

[00259] Example 10. VHH Libraries for CRTH2R

[00260] A VHH library for CRTH2R was developed similar to methods described in Example 9. Briefly, stable cell lines expressing CRTH2R were generated, and target expression was confirmed by FACS. Cells expressing >80% of the target were then used for cell-based selections. Five rounds of cell-based selections were carried out against cells stably overexpressing the target of

interest. 10^8 cells were used for each round of selection. Before selection on target expressing cells, phage from each round was first depleted on 10^8 CHO background cells. Stringency of selections was increased by increasing the number of washes in subsequent rounds of selections. The cells were then eluted from phage using trypsin, and the phage was amplified for the next round of panning. A total of 1000 clones from round 4 and round 5 are sequenced by NGS to identify unique clones for reformatting as VHH-Fc.

[00261] 26 binders out of the 175 unique CRTH2R VHH Fc binders had a target cell mean fluorescence intensity (MFI) value that was 2-fold over parental cells. The data for variant CRTH2-41-51 is seen in **FIGS. 17A-17B** and **Tables 13A-13B**. **Tables 13A-13B** show flow cytometry data as detected with the RL1-A channel. Data for variant CRTH2-44-59 is seen in **FIG. 18** and **FIG. 19**.

Table 13A. Panning Summary

Library	Unique Phage	VHH-Fc FACS binders (MFI values 2 fold over parental cells)
VHH hShuffle	99	16
VHH Ratio/ Shuffle	76	10

Table 13B. CRTH2-41-51 Data

Sample Name	Subset Name	Count	Median:RL1-A
Sample C7.fcs	CRTH2R cells	8663	7441
Sample E10.fcs	Parent Cells	11589	2120

[00262] Example 11. Identification of IgGs for CRTH2R

[00263] Cell binding of anti-CRTH2R antibodies was determined by testing on CHO CRTH2R-positive cells (GFP+) and parental CHO cells (GFP-), comparing parental negative and target positive cells to rule out false-positives. Antibodies as listed in **Table 14A** were titrated starting at 100nM (15ug/mL) with 3-fold titrations, for a total of 8 points. Heavy and light chain sequences for CRTH2R IgG antibodies are shown in **Table 14B**. Binding as detected by mean fluorescence intensity (MFI) by concentration is shown in **FIGS. 20A-20E**. An exemplary gated dot plot and APC histogram at 100 nM with CRTH2-27 is shown in **FIGS. 21A-21B**. Two antibodies (gPCR-51 and gPCR-52) were used as a positive control. Binding profiles of the two positive controls are shown in **FIGS. 22A-22B**.

Table 14A. CRTH2R Antibody variable heavy and light chain sequences

SEQ ID NO	CRTH2R Antibody	Heavy Chain
2338	CRTH2-74	QVQLVESGGGVVQPGRSRLRLSCAASGFSFSEYGIHWVRQAPGKGLEWVAVISYEG SNEYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARANQHFGPVAGG ATPSEEPGSQLTRAELGWDAPPQQESLADELLQLGTEHGYHYGYMDVWGQGTLLV TVSS
2339	CRTH2-24	QVQLVQSGAEVKKPGSSVKVSCASGGSFSNYGISWVRQAPGQGLEWMGGIIPLI GTANYAQKFQGRVTITADESTSTAYMELSSLRSED TAVYYCARDMYDYFTLGPQS IGPLGEVVPADDAFDIWGQGTLLVTVSS
2340	CRTH2-28	QVQLVQSGAEVKKPGSSVNVSCASGGTFSDYAFSWVRQAPGQGLEWMGAIIPFF GTVNYAQKFQGRVTITADESTSTAYMELSSLRSED TAVYYCARDMYDYFATGTG GPEDDLYPQGE LNDGYRIEVPADDAFDIWGQGTLLVTVSS
2341	CRTH2-39	QVQLVQSGAEVKKPGSSVKVSCASVDTFSRYSISWVRQAPGQGLEWMGGIIPVF DTTNYAQKFQGRVTITADESTSTAYMELSSLRSED TAVYYCARDMYDYDFGVILGG TAVGTNNGSANEVVPADDAFDIWGQGTLLVTVSS
2342	CRTH2-19	QVQLVQSGAEVKKPGSSVKVSCASGGTFSSHAINWVRQAPGQGLEWMGRIIPV GTTTYAQKFQGRVTITADESTSTAYMELSSLRSED TAVYYCARDMYDYDFYFGLT LTGDRNDDEVVPADDAFDIWGQGTLLVTVSS
2343	CRTH2-9	QVQLVQSGAEVKKPGSSVKVSCASGGTFSSY AISWVRQAPGQGLEWMGGIIPFG TANYAQKFQGRVTITADESTSTAYMELSSLRSED TAVYYCARDMYDYFWLGDQS TGLIGAEVVPADDAFDIWGQGTLLVTVSS
2344	CRTH2-8	QVQLVQSGAEVKKPGSSVKVSCASGGTFSTDY AISWVRQAPGQGLEWMGGIIPFF GSPNYAQKFQGRVTITADESTSTAYMELSSLRSED TAVYYCARDMYDYFAAGLEG TITEVFDEEGHQGGTEVVPADDAFDIWGQGTLLVTVSS
2345	CRTH2-27	QVQLVESGGGVVQPGRSRLRLSCAASGFTFDNYGMHWVRQAPGKGLEWVAVISYE GSNKKYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARDMYDYDFGSIY GEDVVGELPEVVPADDAFDIWGQGTLLVTVSS
2346	CRTH2-45	QVQLVESGGGVVQPGRSRLRLSCAASGFTFSHYAMHWVRQAPGKGLEWVADISHE GSNKYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARDGRGSLPRPK GGPTSGGGFSTNIGYGFVVQSYDSSSED SGGAFDIWGQGTLLVTVSS
2347	CRTH2-35	QVQLVQSGAEVKKPGSSVKVSCASGGTFRSY AISWVRQAPGQGLEWMGGIIPIS GTTNYAQKFQGRVTITADESTSTAYMELSSLRSED TAVYYCARANQHFRIFGN Y QIYFGHFGYHYGYMDVWGQGTLLVTVSS
2348	CRTH2-50	QVQLVQSGAEVKKPGSSVKVSCASGGTFSSY ALSWVRKAPGQGLEWMGGTIPIF GTVNYAQKFQGRVTITADESTSTAYMELSSLRSED TAVYYCARANQHFRVIGQP SPAVPSRGYIYHGYHYGYMDVWGQGTLLVTVSS
2349	CRTH2-66	QVQLVESGGGVVQPGRSRLRLSCAASGFD FSGYGMHWVRQAPGKGLEWVAVISYE GSNKFYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARDLRELECEEW T IEVHGQEF AVHQDRGGVFSRGPCVDPRGVAGSFDVWGQGTLLVTVSS
2350	CRTH2-57	QVQLVQSGAEVKKPGSSVKVSCASGGTFSSY AMSWVRQAPGQGLEWMGGIIPL FGTTDYAQKFQGRVTITADESTSTAYMELSSLRSED TAVYYCARANQH FVKIQGA PVSTPVPFGFTTG YHYGYMDVWGQGTLLVTVSS
2351	CRTH2-32	QVQLVESGGGVVQPGRSRLRLSCAASGFTFSKHGMHWVRQAPGKGLEWVAFISYE GSEKYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARDMYDYDFHYST VGATYYYYLGSETEVVPADDAFDIWGQGTLLVTVSS
2352	CRTH2-15	QVQLVQSGAEVKKPGSSVKVSCASGGTFSTY AIDWVRQAPGQGLEWMGGIIPLF GSPNYAQKFQGRVTITADESTSTAYMELSSLRSED TAVYYCARANQHFFLYEGTSS SWLHVGHARYGYHYGYMDVWGQGTLLVTVSS
2353	CRTH2-25	QVQLVQSGAEVKKPGSSVKVSCASGGSFRSYGISWVRQAPGQGLEWMGRIIPLF GTPDYAQKFQGRVTITADESTSTAYMELSSLRSED TAVYYCARDMYDYDFEDVDEG SLYLDMGRTFEVVPADDAFDIWGQGTLLVTVSS
2354	CRTH2-42	QVQLVESGGGVVQPGRSRLRLSCAASGFAFSSYAMHWVRQAPGKGLEWVAVISYE GSNEYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARDLRELECEEW T VLQY GKFHMRWAESGEGSLSRGPCVDPRGVAGSFDVWGQGTLLVTVSS
2355	CRTH2-55	QVQLVESGGGVVQPGRSRLRLSCAASGFTFRSYDMHWVRQAPGKGLEWVAVISYE GSEKYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKHMMSMQASTEG DFGLEEV TGEVDDRADLVGDAFDVWGQGTLLVTVSS

2356	CRTH2-60	QVQLVQSGAEVKKPGSSVKVSCKASGGTFKKNYAINWVRQAPGQGLEWMGAIPK FGAANYAQKFQGRVTITADESTSTAYMELSSLRSEDTAVYYCARANQHFSAVRGL AFGYGYRIGGYHYYGMDVWGQGLTVTVSS
2357	CRTH2-70	QVQLVQSGAEVKKPGSSVKVSCKASGGTFSNHAIWVRQAPGQGLEWMGGIPIFG TPSYAQKFQGRVTITADESTSTAYMELSSLRSEDTAVYYCARDMYDYDFDVISAGV VGAGNPEVVPADDAFDIWGQGLTVTVSS
2358	CRTH2-48-9	EVQLLESGGGLVQPGGSLRLSCAASGFSFSTHAMS WVRQAPGKLEWVSTIGGSG GSTYYADSVKGRFTISRDN SKNTLYLQMNSLR AEDTAVYYCARAHG DSSSWYFS YYYMDVWGQGLTVTVSS
2359	CRTH2-41-51	EVQLVESGGGLVQPGGSLRLSCAASGGIFRFNAMGWFRQAPGKERELVAGISGSG GDTYYADSVKGRFTISADNSKNTAYLQMNSLKPEDTAVYYCAA FRGIMRPDWGQ GTLTVTVSS
2360	CRTH2-44-6	EVQLVESGGGLVQPGGSLRLSCAASGPTFDITYVMGWFRQAPGKEREFVAAISMSG DDTAYADSVKGRFTISADNSKNTAYLQMNSLKPEDTAVYYCATDLRGRGDVSEY EYDWGQGLTVTVSS
SEQ ID NO	CRTH2R Antibody	Light Chain
2361	CRTH2-74	QSVLTQPPSVSAAPGQKVTISCSGSTS NIGKNYVSWYQQLPGTAPKLLIYDDDERPS GIPDRFSGSGSGTSATLGITGLQTGDEADYYCEAWDADLSGAVFGGGTKLTVL
2362	CRTH2-24	QSVLTQPPSVSAAPGQKVTISCSGSSSNIGNNFVSWYQQLPGTAPKLLIYDNIQRPS GIPDRFSGSGSGTSATLGITGLQTGDEADYYCGTWDTSLSAVVFGGGTKLTVLRTV AAPS VFIFPPSDEQLKSGTASVCLN NFYPREK VQWKVDNALQSGNSQESVTE QDSK DSTYLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC
2363	CRTH2-28	QSVLTQPPSVSAAPGQKVTISCSGSSNIGKNYVSWYQQLPGTAPKLLIYDDHKRPS GIPDRFSGSGSGTSATLGITGLQTGDEADYYCATWDRGLSAAVFGGGTKLTVL
2364	CRTH2-39	QSVLTQPPSVSAAPGQKVTISCSGSSSNIGDNDVSWYQQLPGTAPKLLIYDDDKRPS GIPDRFSGSGSGTSATLGITGLQTGDEADYYCASWDTLSGGYVFGGGTKLTVL
2365	CRTH2-19	QSALTQPASVSGSPGQSITISCTGTSSDVGGYDYVTWYQQHPGKAPKLMYD V DTR PSGVSNRFSGSGSGNTASLTISGLQAEDEADYYCSSYTTSTSYVFGGGTKLTVL
2366	CRTH2-9	QSVLTQPPSVSAAPGQKVTISCSGSTS NIGNNYVSWYQQLPGTAPKLLIYENDERPS GIPDRFSGSGSGTSATLGITGLQTGDEADYYCATWDTRLSAAVFGGGTKLTVL
2367	CRTH2-8	QSVLTQPPSVSAAPGQKVTISCSGSSSNIGKNYVSWYQQLPGTAPKLLIYDNNQRPS GIPDRFSGSGSGTSATLGITGLQTGDEADYYCGTWDTSLSVVFVFGGGTKLTVL
2368	CRTH2-27	QSALTQPASVSGSPGQSITISCTGTSSNDVGAYNFVSWYQQHPGKAPKLMYDISNR PSGVSNRFSGSGSGNTASLTISGLQAEDEADYYCSSYTRSNTRVFGGGTKLTVL
2369	CRTH2-45	QSVLTQPPSVSAAPGQKVTISCSGTSSNIENNYVSWYQQLPGTAPKLLIYDNVSRPS GIPDRFSGSGSGTSATLGITGLQTGDEADYYCGTWDNTVSAPWVFGGGTKLTVL
2370	CRTH2-35	QSALTQPASVSGSPGQSITISCTGTSSDIGGYEFVSWYQQHPGKAPKLMYGVSRRP SGVSNRFSGSGSGNTASLTISGLQAEDEADYYCGSYTSSSTPYVFGGGTKLTVL
2371	CRTH2-50	QSALTQPASVSGSPGQSITISCTGTSSDIGGYNFVSWYQQHPGKAPKLMYDVSNRP QGVSNRFSGSGSGNTASLTISGLQAEDEADYYCSSYTSSNTYVVFVFGGGTKLTVL
2372	CRTH2-66	EIVMTQSPATLSVSPGERATLSCRASQGVGSNLAWYQQKPGQAPRLIYRTSIRAT GIPARFSGSGSGTEFTLTISSLQSEDFAVYYCQQYYSWPPLTFGGGKVEIK
2373	CRTH2-57	QSVLTQPPSVSAAPGQKVTISCSGSSSNIEDNYVSWYQQLPGTAPKLLIYDNFKRPS GIPDRFSGSGSGTSATLGITGLQTGDEADYYCGTWDTSLSAALFVFGGGTKLTVL
2374	CRTH2-32	QSALTQPASVSGSPGQSITISCTGTSSGVGGYDYVSWYQQHPGKAPKLMYDDNN RPSGVSNRFSGSGSGNTASLTISGLQAEDEADYYCSSYTGSSSTLYVFGGGTKLTVL
2375	CRTH2-15	QSVLTQPPSVSAAPGQKVTISCSGSGSNIGSNYVSWYQQLPGTAPKLLIYDNIRPS GIPDRFSGSGSGTSATLGITGLQTGDEADYYCAAWDTRLSAGVFGGGTKLTVL
2376	CRTH2-25	DIQMTQSPSSLSASVGRVTITCRASQGISTYLNWYQQKPGKAPKLLIYATSSLQSG VPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQSYSTP-WTFGGGKVEIK
2377	CRTH2-42	QSALTQPASVSGSPGQSITISCTGTSSDVGGYRYVSWYQQHPGKAPKLMYVNVNR PSGVSNRFSGSGSGNTASLTISGLQAEDEADYYCSSYRSSSTLVFGGGTKLTVL
2378	CRTH2-55	QSVLTQPPSVSAAPGQKVTISCSGSSSNIGDNFVSWYQQLPGTAPKLLIYDDDERPS GIPDRFSGSGSGTSATLGITGLQTGDEADYYCGAWDRSLSAVFGGGTKLTVL
2379	CRTH2-60	QSVLTQPPSVSAAPGQKVTISCSGSTS NIGINYVSWYQQLPGTAPKLLIYENRKRPS GIPDRFSGSGSGTSATLGITGLQTGDEADYYCATWDASLKNLVFGGGTKLTVL

2380	CRTH2-70	QSVLTQPPSVSAAPGQKVTISCSGSTSNIGNNFVSWYQQLPGTAPKLLIYDNEKRPS GIPDRFSGSKSGTSATLGLTGLQTGDEADYYCGTWDERQTDES YVFGGGTKLTVL
2381	CRTH2-48-9	DIQMTQSPSSLSASVGDRTITCRASQSDYVNWYQQKPGKAPKLLIYGASILQTG VPSRFSGSGSGTDFTLTISLQPEDFATYYCQQSFTTPWTFGGGTKVEIK

Table 14B. Variably Heavy Chain CDR3 Sequences

SEQ ID NO	CRTH2R Antibody	CDRH3
2382	CRTH2-74	CARANQHFGP VAGGATPSEEPGSQLTRAELGWDAPPGQESLADELLQLGTEH GYHHYYGMDVW
2383	CRTH2-24	CARDMYDFTLGPQSIGPLGEVVPADDAFDIW
2384	CRTH2-28	CARDMYDFATGTGGPEDDLYPQGE LNDGYRIEVVPADDAFDIW
2385	CRTH2-39	CARDMYDFGVILGGTAVGTNNGSANEVVPADDAFDIW
2386	CRTH2-19	CARDMYDFDYFGLTLTGDRNDDEVVPADDAFDIW
2387	CRTH2-9	CARDMYDFWLGDQSTGSLIGAEVVPADDAFDIW
2388	CRTH2-8	CARDMYDFFAAGLEG TITEVFDEEGHQGGTEVVPADDAFDIW
2389	CRTH2-27	CARDMYDFGSIYGEDVVGELPEVVPADDAFDIW
2390	CRTH2-45	CARDGRGSLPRPKGGPTSGGGFSTNIGYGFVVQSYDSSSEDSSGGAFDIW
2391	CRTH2-35	CARANQHFRIFGN YQIYFGHFGYHYYGMDVW
2392	CRTH2-50	CARANQHFRVIGQPSAVPSRGYIYHGYHYYGMDVW
2393	CRTH2-66	CARDLRELECEEW TIEVHGQEF AVHQDRGGVFSRGP CVDPGRGVAGGSFDVW
2394	CRTH2-57	CARANQHFKIQGAPVSTPVPFGTGTGYHYYGMDVW
2395	CRTH2-32	CARDMYDFHYSTVGATYYYYLGSETEVVPADDAFDIW
2396	CRTH2-15	CARANQHFFLYEGTSSSWLHVGHARYGYHYYGMDVW
2397	CRTH2-25	CARDMYDFEDVDEGSLYLD MGRTFEVVPADDAFDIW
2398	CRTH2-42	CARDLRELECEEWTVLQY GKFHMRWAESGEGSLSRGPCVDPRGVAGSFDVW
2399	CRTH2-55	CAKHMSMQASTEGDFGLEEV TGEVDDRADLVGDAFDVW
2400	CRTH2-60	CARANQHFSAVRGLAFGYGYRIGGYHYYGMDVW
2401	CRTH2-70	CARDMYDFDVISAGVVGAGNPEVVPADDAFDIW
2402	CRTH2-74	CARDMYDFDVISAGVVGAGNPEVVPADDAFDIW

[00264] In subsequent examples, five antibodies were shown to have functional effects in cAMP assays; CRTH2-9, CRTH2-27, CRTH2-50, CRTH2-32, and CRTH2-42. The binding curves of these antibodies are compared in **FIGS. 23A-23B**.

[00265] Example 12. Antagonist activity using cAMP assay

[00266] A library of CRTH2R IgG antibodies were assayed to determine antagonist function in PGD₂-induced cAMP signals. Briefly, cells were pre-incubated with IgG (titration 1:3) for 1 hour at room temperature. Subsequently, cells were stimulated with PGD₂ (0.59 nM) for 30 min at 37°C in the presence of forskolin, since CRTH2R is G α_i coupled.

[00267] Results showing effect of antibody on detected signal in relative light units (rlu) are shown in. At the highest concentration tested (300 nM), some of the CRTH2R IgGs caused an upward deflection of the signal, indicating inhibition of the cAMP signal induced by PGD₂ stimulation. For comparison, bar charts showing the ratio of IgG treated versus control treated for the three highest IgG concentrations tested are shown in **FIG. 24A**. Antibodies depicted in **FIG. 24B** show CRTH2R IgG antibodies which resulted in more than a 20% antagonist activity at 33

nM, specifically CRTH2-74, CRTH2-24, CRTH2-28, CRTH2-19, CRTH2-45, CRTH2-9, CRTH2-8, CRTH2-15, CRTH2-42, CRTH2-60, and CRTH2-70.

[00268] Example 13. Allosteric modulation of PGD2-induced cAMP signal

[00269] CRTH2R IgG antibodies were assayed for allosteric activity. Allosteric modulation was determined by assaying CRTH2R IgG antibodies in PGD2-induced cAMP signal. Briefly, cells were re-incubated with no IgG antibody or 100 nM CRTH2R IgG antibody. Subsequently, cells were stimulated with PGD2 at various concentrations in the presence of forskolin followed by assay for cAMP activity.

[00270] Results of the cAMP assays is seen in **FIG. 25**. A right-ward shift the PGD2 dose response curve (and increase in IC₅₀ value) indicates a negative allosteric effect. As shown in **FIG. 25**, five of the CRTH2R IgG (CRTH2-9, CRTH2-27, CRTH2-50, CRTH2-32, and CRTH2-42) caused an IC₅₀ fold difference of >2.0 compared with PGD2 alone, suggesting they are negative allosteric modulators.

[00271] Example 14. Agonist activity of PGD2-induced cAMP signal

[00272] CRTH2R IgG antibodies were assayed for agonist function. Agonist activity was determined by assaying CRTH2R IgG antibodies described in Example 11 in PGD2-induced cAMP signal.

[00273] Briefly, cells were treated with PGD2 or CRTH2R IgG antibodies both in the presence of forskolin. The CRTH2R IgG antibodies included CRTH2-74, CRTH2-24, CRTH2-28, CRTH2-39, CRTH2-19, CRTH2-9, CRTH2-8, CRTH2-27, CRTH2-45, CRTH2-35, CRTH2-50, CRTH2-66, CRTH2-57, CRTH2-32, CRTH2-15, CRTH2-25, CRTH2-42, CRTH2-55, CRTH2-60, and CRTH2-70. Treatment stimulations were performed for 30 min at 37 °C. cAMP assays were then performed (data not shown).

[00274] Example 15. Control experiments showing allosteric modulators

[00275] Allosteric modulation was determined for a known CRTH2R antagonist (small molecule OC000459) and two control antibodies. Experiments were performed similar to those described in Example 13. Briefly, cells were treated with OC000459, comparator CRTH2R AB51 antibody, or comparator CRTH2R AB52 antibody. Cells were then stimulated with PGD2 in the presence of forskolin.

[00276] Results are shown in **FIGS. 26A – 26C**. OC000459 causes a strong right-ward shift of the curve and a 459-fold increase in the IC₅₀ value (**FIG. 26A**). Incubation with CRTH2R AB51 caused no change in IC₅₀ value (**FIG. 26B**). Incubation with the comparator antibody #52 caused a 3.5-fold decrease in the IC₅₀ value, indicating it is a positive allosteric modulator, i.e. it has agonistic effects (**FIG. 26C**).

[00277] Example 16. CRTH2R β -arrestin recruitment assay for antagonist modulation

[00278] Antagonist modulation by nine CRTH2R IgG antibodies was determined. The nine CRTH2R IgG antibodies included CRTH2-9, CRTH2-27, CRTH2-50, CRTH2-32, CRTH2-42, CRTH2-74, CRTH2-55, CRTH2-28, and CRTH2-39. The antagonist function of these nine antibodies as compared to OC000459 was determined using a PGD2-induced β -arrestin recruitment. Results, including a positive control using small molecule OC000459, are shown in **FIGS. 27A-27D**.

[00279] Example 17. CRTH2R β -arrestin recruitment assay for allosteric modulation

[00280] Allosteric modulation by nine CRTH2R IgGs were determined. The nine CRTH2R IgGs included CRTH2-9, CRTH2-27, CRTH2-50, CRTH2-32, CRTH2-42, CRTH2-74, CRTH2-55, CRTH2-28, and CRTH2-39. The allosteric modulation of these nine antibodies as compared to OC000459 was determined using a PGD2-induced β -arrestin recruitment.

[00281] Briefly, cells were pre-incubated with IgG (100 nM) for 1 hour at room temperature followed by PGD2 stimulation for 90 min at 37°C. Data was normalized against the first data point (lowest PGD2 and zero Ab) in each graph. Results, including a positive control using small molecule OC000459, are shown in **FIGS. 27A-27D**.

[00282] Example 18. CRTH2R hyperimmune immunoglobulin library

[00283] A hyperimmune immunoglobulin (IgG) library was created using similar methods as described in Example 9. Briefly, the hyperimmune IgG library was generated from analysis of databases of human naïve and memory B-cell receptor sequences consisting of more than 37 million unique IgH sequences from each of 3 healthy donors. More than two million CDRH3 sequences were gathered from the analysis and individually constructed using methods similar to Examples 1-3. Any duplicate CDRH3s and potential liability motifs that frequently pose problems in development were removed during the library synthesis step. These CDRH3 sequence diversities were then combinatorially assembled and incorporated onto the DP47 human framework to construct a highly functional antibody Fab library with 1×10^{10} size. A schematic of the design can be seen in **FIG. 28**.

[00284] A CRTH2R hyperimmune immunoglobulin library was generated. Briefly, five rounds of cell-based selections were carried out against cells stably overexpressing the target of interest. 10^8 cells were used for each round of selection. Before selection on target expressing cells, phage from each round was first depleted on 10^8 CHO background cells. Stringency of selections was increased by increasing the number of washes in subsequent rounds of selections. The cells were then eluted from phage using trypsin, and the phage gets amplified for the next round of panning.

[00285] CRTH2R immunoglobulins were assessed for binding affinity and allosteric modulator function of PGD2-induced cAMP. As seen in **FIGS. 29A-29F**, three specific CRTH2R immunoglobulins were identified with sub nanomolar to single digit nanomolar cell binding affinities to hCRTH2R and had inhibitory activities in the allosteric cAMP assay. The sequences for the three CRTH2R immunoglobulins CRTH2-48-3, CRTH2-48-21, and CRTH2-48-27 are seen in **Table 15**.

Table 15. CRTH2R sequences

Variable Heavy Chain		
SEQ ID NO:	IgG	Amino Acid Sequence
2403	CRTH2-48-3	EVQLVESGGGLVQAGGSLRLSCAASGSIFRINAMGWFRQAPGKEREGVAAINNF GTTKYADSVKGRFTISADNAKNTVYLQMNSLKPEDTAVYYCAAVRWGPRNDDR YDWGQGTQVTVSS
2404	CRTH2-48-21	EVQLVESGGGLVQAGGSLRLSCAASGSFFSINAMGWFRQAPGKEREFVAGITRSG VSTSYADSVKGRFTISADNAKNTVYLQMNSLKPEDTAVYYCAHRIVVGGTSVG DWRWGQGTQVTVSS
2405	CRTH2-48-27	EVQLVESGGGLVQAGGSLRLSCAASGSIFHINAMGWFRQAPGKEREGVAAINNF GTTKYADSVKGRFTISANNAKNTVYLQMNSLKPEDTAVYYCAAVRWGPRNDDR YDWGQGTQVTVSS
Variable Light Chain		
2406	CRTH2-48-3	DIQMTQSPSSLSASVGDRVTITCRASQSISSDLNHWYQQKPGKAPKLLIYFASGLQS GVPSRFSGSGSGTDFTLTISLQPEDFATYYCQQSYSSPLTFGGGKVEIKR
2407	CRTH2-48-21	DIQMTQSPSSLSASVGDRVTITCRISQSISSDLNHWYQQKPGKAPKLLIYATSSLES GVPSRFSGSGSGTDFTLTISLQPEDFATYYCQQSYSTLLTFGGGKVEIKR
2408	CRTH2-48-27	DIQMTQSPSSLSASVGDRVTITCRASQSISSDLNHWYQQKPGKAPKLLIYGASRLS GVPSRFSGSGSGTDFTLTISLQPEDFATYYCRQSYSTPWTFGGKVEIKR

[00286] Example 19. GPCR Libraries with Varied CDR's

[00287] A GPCR library was created using a CDR randomization scheme.

[00288] Briefly, GPCR libraries were designed based on GPCR antibody sequences. Over sixty different GPCR antibodies were analyzed and sequences from these GPCRs were modified using a CDR randomization scheme.

[00289] The heavy chain IGHV3-23 design is seen in **FIG. 30A**. As seen in **FIG. 30A**, IGHV3-23 CDRH3's had four distinctive lengths: 23 amino acids, 21 amino acids, 17 amino acids, and 12 amino acids, with each length having its residue diversity. The ratio for the four lengths were the following: 40% for the CDRH3 23 amino acids in length, 30% for the CDRH3 21 amino acids in length, 20% for the CDRH3 17 amino acids in length, and 10% for the CDRH3 12 amino acids in length. The CDRH3 diversity was determined to be 9.3×10^8 , and the full heavy chain IGHV3-23 diversity was 1.9×10^{13} .

[00290] The heavy chain IGHV1-69 design is seen in **FIG. 30B**. As seen in **FIG. 30B**, IGHV1-69 CDRH3's had four distinctive lengths: 20 amino acids, 16 amino acids, 15 amino acids, and 12 amino acids, with each length having its residue diversity. The ratio for the four lengths were the

following: 40% for the CDRH3 20 amino acids in length, 30% for the CDRH3 16 amino acids in length, 20% for the CDRH3 15 amino acids in length, and 10% for the CDRH3 12 amino acids in length. The CDRH3 diversity was determined to be 9×10^7 , and the full heavy chain IGHV-69 diversity is 4.1×10^{12} .

[00291] The light chains IGKV 2-28 and IGLV 1-51 design is seen in **FIG. 30C**. Antibody light chain CDR sequences were analyzed for position-specific variation. Two light chain frameworks were selected with fixed CDR lengths. The theoretical diversities were determined to be 13800 and 5180 for kappa and light chains, respectively.

[00292] The final theoretical diversity was determined to be 4.7×10^{17} and the final, generated Fab library had a diversity of 6×10^9 . See **FIG. 30D**.

[00293] **Example 20. CRTH2R Libraries with Varied CDR's**

[00294] A CRTH2R library is created using a CDR randomization scheme similarly described in Example 19.

[00295] Briefly, CRTH2R libraries are designed based on GPCR antibody sequences. Over sixty different GPCR antibodies are analyzed and sequences from these GPCRs are modified using a CDR randomization scheme. CRTH2R variant IgGs designed using the CDR randomization scheme are purified and are assayed to determine cell-based affinity measurements and for functional analysis.

[00296] While preferred embodiments of the present disclosure have been shown and described herein, it will be obvious to those skilled in the art that such embodiments are provided by way of example only. Numerous variations, changes, and substitutions will now occur to those skilled in the art without departing from the disclosure. It should be understood that various alternatives to the embodiments of the disclosure described herein may be employed in practicing the disclosure. It is intended that the following claims define the scope of the disclosure and that methods and structures within the scope of these claims and their equivalents be covered thereby.