

VARIANT NUCLEIC ACID LIBRARIES FOR SINGLE DOMAIN ANTIBODIES**CROSS-REFERENCE**

[0001] This application claims the benefit of U.S. Provisional Patent Application No. 62/904,620 filed on September 23, 2019; U.S. Provisional Patent Application No. 62/935,603 filed on November 14, 2019; and U.S. Provisional Patent Application No. 62/945,761 filed on December 9, 2019, each of which is incorporated by reference in its entirety.

BACKGROUND

[0002] Antibodies possess the capability to bind with high specificity and affinity to biological targets. However, the design of therapeutic antibodies is challenging due to balancing of immunological effects with efficacy. Single domain antibodies such as VHH antibodies have several beneficial characteristics. Thus, there is a need to develop compositions and methods for generation of antibodies such as VHH antibodies for use in therapeutics.

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 antibodies or antibody fragments comprising a CDRH1 comprising an amino acid sequence at least about 90% identical to that set forth in SEQ ID NOs: 152 or 155, a CDRH2 comprising an amino acid sequence at least about 90% identical to that set forth in SEQ ID NOs: 153 or 156, and a CDRH3 comprising an amino acid sequence at least about 90% identical to that set forth in SEQ ID NOs: 154 or 157. Further provided herein are antibodies or antibody fragments, further comprising a CDRL1 comprising an amino acid sequence at least about 90% identical to that set forth in SEQ ID NOs: 158 or 161, a CDRL2 comprising an amino acid sequence at least about 90% identical to that set forth in SEQ ID NOs: 159 or 162, and a CDRL3 comprising an amino acid sequence at least about 90% identical to that set forth in SEQ ID NOs: 160 or 163.

[0005] Provided herein are methods of treating cancer comprising administering the antibody or antibody fragment described herein.

[0006] Provided herein are methods of treating a viral infection comprising administering the antibody or antibody fragment described herein.

[0007] Provided herein are nucleic acid libraries comprising: a plurality of sequences comprising nucleic acids that when translated encode for an antibody or antibody fragment, 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) or a CDR1, CDR2, or CDR3 on a variable region of a light chain (VL); wherein the library comprises at least 30,000 variant sequences; and wherein the antibody or antibody fragments bind to its antigen with a K_D of less than 100 nM. Further provided herein are nucleic acid libraries, wherein the antibody is a single domain antibody. Further provided herein are nucleic acid libraries, wherein the single domain antibody is a VHH antibody. Further provided herein are nucleic acid libraries, wherein the antibody binds to TIGIT. Further provided herein are nucleic acid libraries, wherein the variable region of the heavy chain when translated comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NOs: 84-100. Further provided herein are nucleic acid libraries, wherein the variable region of the light chain when translated comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NOs: 101-117. Further provided herein are nucleic acid libraries, wherein the CDR1, CDR2, or CDR3 on the variable region of the heavy chain comprises an amino acid sequence at least about 90% identical to that set forth in any one of SEQ ID NOs: 67-83 or 118-128. Further provided herein are nucleic acid libraries, wherein the CDR1, CDR2, or CDR3 on the variable region of the light chain comprises an amino acid sequence at least about 90% identical to that set forth in any one of SEQ ID NOs: 129-137. Further provided herein are nucleic acid libraries, wherein the antibody binds to CD47. Further provided herein are nucleic acid libraries, wherein the antibody binds to CD3 epsilon. Further provided herein are nucleic acid libraries, wherein the variable region of the heavy chain when translated comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NOs: 138-141. Further provided herein are nucleic acid libraries, wherein the variable region of the light chain when translated comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NOs: 142-145. Further provided herein are nucleic acid libraries, wherein the nucleic acid library comprises at least 50,000 variant sequences. Further provided herein are nucleic acid libraries, wherein the nucleic acid library comprises at least 100,000 variant sequences. 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 the nucleic acid library has a theoretical diversity of at least 10^9 sequences.

[0008] Provided herein are nucleic acid libraries comprising: a plurality of sequences comprising nucleic acids that when translated encode for a single domain antibody, wherein each sequence of the plurality of sequences comprises a variant sequence encoding for 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 antibody or antibody fragments bind 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 single domain antibody is a VHH antibody. Further provided herein are nucleic acid libraries, wherein the antibody binds to TIGIT. Further provided herein are nucleic acid libraries, wherein the CDR1, CDR2, or CDR3 on the variable region of the heavy chain comprises an amino acid sequence at least about 90% identical to that set forth in any one of SEQ ID NOs: 67-83 or 118-128. Further provided herein are nucleic acid libraries, wherein the variable region of the heavy chain when translated comprises an amino acid sequence at least about 90% identical to that set forth in any one of SEQ ID NOs: 84-100. Further provided herein are nucleic acid libraries, wherein the CDR3 on the variable region of the heavy chain comprises an amino acid sequence at least about 90% identical to that set forth in any one of SEQ ID NOs: 101-117. Further provided herein are nucleic acid libraries, wherein the antibody binds to CD47. Further provided herein are nucleic acid libraries, wherein the antibody binds to CD3 epsilon. Further provided herein are nucleic acid libraries, wherein the variable region of the heavy chain when translated comprises an amino acid sequence at least about 90% identical to that set forth in SEQ ID NOs: 138-141. Further provided herein are nucleic acid libraries, wherein the nucleic acid library comprises at least 50,000 variant sequences. Further provided herein are nucleic acid libraries, wherein the nucleic acid library comprises at least 100,000 variant sequences. 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 the nucleic acid library has a theoretical diversity of at least 10^9 sequences.

[0009] Provided herein are methods for generating a nucleic acid library encoding for a single domain antibody 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 sequences 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 sequences 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 sequences 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 single domain antibody, and wherein at least about 70% of the variant nucleic acids encode for a single domain antibody 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 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 single domain antibody binds to TIGIT. Further provided herein are methods for generating a nucleic acid library, wherein the single domain antibody comprises an amino acid sequence at least about 90% identical to that set forth in any one of SEQ ID NOs: 84-100 or 138-141. Further provided herein are methods for generating a nucleic acid library, wherein the single domain antibody binds to CD47. 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 single domain antibody that binds to an antigen 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 single domain antibody that binds to an antigen 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 single domain antibody that binds to an antigen with a K_D of less than 25 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 single domain antibody that binds to an antigen 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 has a theoretical diversity of at least 10^9 sequences.

BRIEF DESCRIPTION OF THE DRAWINGS

[0010] **Figure 1** presents a diagram of steps demonstrating an exemplary process workflow for gene synthesis as disclosed herein.

[0011] **Figure 2** illustrates an example of a computer system.

- [0012] **Figure 3** is a block diagram illustrating an architecture of a computer system.
- [0013] **Figure 4** 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).
- [0014] **Figure 5** is a block diagram of a multiprocessor computer system using a shared virtual address memory space.
- [0015] **Figures 6-7** depicts a graph of TIGIT affinity distribution for the VHH libraries, depicting either the affinity threshold from 20 to 4000 (**FIG. 6**) or the affinity threshold from 20 to 1000 (**FIG. 7**). Out of 140 VHH binders, 51 variants were < 100 nM and 90 variants were < 200 nM.
- [0016] **Figure 8** depicts graphs of CDR3 counts per length for 'VHH library,' 'VHH shuffle' library, and 'VHH hShuffle library.'
- [0017] **Figure 9** depicts a graph of a TIGIT:CD155 blockade assay for TIGIT VHH Fc binders. Concentration of the TIGIT VHH Fc binders in nanomolar (nM) is on the x-axis and relative HRP signal is on the y-axis.
- [0018] **Figure 10** depicts a graph of CD47 affinity distribution of the CD47 VHH Fc binders. Affinity threshold (monovalent KD) is on the x-axis and count is on the y-axis for 'VHH ratio' library (horizontal bars), 'VHH shuffle' library (black bars), and 'VHH hShuffle' library (dotted bars).
- [0019] **Figure 11** depicts a graph of CD47-SIRPalpha inhibition assay for CD47 VHH Fc binders. Concentration of the CD47 VHH Fc binders in nanomolar (nM) is on the x-axis and relative HRP signal is on the y-axis.
- [0020] **Figures 12A-12B** depict graphs of FACS analysis (**FIG. 12A**) and graphs of a dose curve and specificity (**FIG. 12B**) of GLP1R-43-77.
- [0021] **Figures 13A-13B** depict graphs of FACS analysis (**FIG. 13A**) and graphs of a dose curve and cAMP activity (**FIG. 13B**) of CRTH2-41-51.
- [0022] **Figures 14A-14B** depict graphs of a dose curve (**FIG. 14A**) and FACS analysis (**FIG. 14B**) of CRTH2-44-59.
- [0023] **Figures 15A-15E** 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.
- [0024] **Figure 16A** depicts an example gated dot plot showing CRTH2-27 binding at 100 nM.

- [0025] **Figure 16B** depicts an example APC histogram showing CRTH2-27 binding at 100 nM.
- [0026] **Figure 17A** depicts binding analysis as in previous figures using comparator antibody gPCR-51.
- [0027] **Figure 17B** depicts binding analysis as in previous figures using comparator antibody gPCR-52.
- [0028] **Figures 18A-18B** depict IgG binding curves with CRTH2-9, CRTH2-27, CRTH2-50, CRTH2-32, and CRTH2-42, which have functional effects in cAMP assays.
- [0029] **Figure 19A** depicts results of CRTH2R cAMP assays across all antibodies tested at 300, 100, and 33 nM.
- [0030] **Figure 19B** depicts results of CRTH2R cAMP assays across all antibodies tested at 33 nM.
- [0031] **Figure 20** indicates the negative allosteric effect seen in five of the CRTH2R IgGs (CRTH2-9, CRTH2-27, CRTH2-50, CRTH2-32, and CRTH2-42).
- [0032] **Figures 21A-21C** depict control experiments of allosteric modulators, showing comparator antibody 52 is a positive allosteric modulator.
- [0033] **Figures 22A-22D** depict activity of CRTH2R in β -arrestin recruitment assays of CRTH2R IgGs.
- [0034] **Figure 23** depicts a schema of libraries generated herein.
- [0035] **Figure 24** depicts a schema of design of phage-displayed hyperimmune libraries generated herein.
- [0036] **Figures 25A-25B** depict heavy chain CDR length distribution of the hyperimmune libraries as assessed by next generation sequencing. **Figure 25A** depicts a graph of CDR3 counts per length. **Figure 25B** depicts graphs of CDRH1, CDRH2, and CDRH3 lengths.
- [0037] **Figure 26** depicts a schema of the workflow of selection of soluble protein targets.
- [0038] **Figures 27A-27D** depict graphs of data from hTIGIT ELISA after Round 3 and Round 4 of panning.
- [0039] **Figures 27E-27F** depict schemas of CDRH3 length, yield, and affinity (K_D) for the hTIGIT immunoglobulins.
- [0040] **Figures 28A-28D** depict graphs of data from human CD3 epsilon (hCD3) and cyno CD3 epsilon (cCD3) ELISA after Round 4 and Round 5 of panning.
- [0041] **Figures 28E-28L** depict graphs of cross-reactive human CD3 epsilon (hCD3) and cyno CD3 epsilon (cCD3) immunoglobulins.
- [0042] **Figures 29A-29G** depict graphs of titration of human CD3 on CD8+, CD3+, and CD3- T cells.

[0043] **Figures 30A-30F** depict graphs of binding affinity for the CRTH2R immunoglobulins CRTH2-48-03 (**FIG. 30A**), CRTH2-48-21 (**FIG. 30B**), and CRTH2-48-27 (**FIG. 30C**) and cAMP assays for CRTH2-48-03 (**FIG. 30D**), CRTH2-48-21 (**FIG. 30E**), and CRTH2-48-27 (**FIG. 30F**).

[0044] **Figures 31A-31B** depict graphs of a dose curve (**FIG. 31A**) and FACS analysis (**FIG. 31B**) of A2AR-90-007.

DETAILED DESCRIPTION

[0045] 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.

[0046] Definitions

[0047] 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.

[0048] 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.

[0049] 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.

[0050] 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.

[0051] Antibody Libraries

[0052] Provided herein are methods, compositions, and systems for generation of antibodies. In some instances, the antibodies are single domain antibodies. Methods, compositions, and

systems described herein for the optimization of antibodies comprise a ratio-variant approach that mirror the natural diversity of antibody sequences. In some instances, libraries of optimized antibodies comprise variant antibody sequences. In some instances, the variant antibody sequences are designed comprising variant CDR regions. In some instances, the variant 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.

[0053] 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.

[0054] Described herein are libraries comprising variation in a CDR region. In some instances, the CDR is CDR1, CDR2, or CDR3 of a variable heavy chain. In some instances, the CDR is CDR1, CDR2, or CDR3 of a variable light chain. In some instances, the libraries comprise multiple variants encoding for CDR1, CDR2, or CDR3. In some instances, the libraries as described herein encode for at least 50, 100, 200, 300, 400, 500, 1000, 1200, 1500, 1700, 2000,

2500, 3000, 3500, 4000, 4500, 5000, or more than 5000 CDR1 sequences. In some instances, the libraries as described herein encode for at least 50, 100, 200, 300, 400, 500, 1000, 1200, 1500, 1700, 2000, 2500, 3000, 3500, 4000, 4500, 5000, or more than 5000 CDR2 sequences. In some instances, the libraries as described herein encode for at least 50, 100, 200, 300, 400, 500, 1000, 1200, 1500, 1700, 2000, 2500, 3000, 3500, 4000, 4500, 5000, or more than 5000 CDR3 sequences. In-silico antibodies libraries are in some instances synthesized, assembled, and enriched for desired sequences.

[0055] Following synthesis of CDR1 variants, CDR2 variants, and CDR3 variants, in some instances, the CDR1 variants, the CDR2 variants, and the CDR3 variants are shuffled to generate a diverse library. In some instances, the diversity of the libraries generated by methods described herein have 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} , or more than 10^{18} 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} , or more than 10^{18} sequences.

[0056] The germline sequences corresponding to a variant sequence may also be modified to generate sequences in a library. For example, sequences generated by methods described herein comprise at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or more than 16 mutations from the germline sequence. In some instances, sequences generated comprise no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or no more than 18 mutations from the germline sequence. In some instances, sequences generated comprise about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or about 18 mutations relative to the germline sequence.

[0057] *Antibody Libraries*

[0058] Provided herein are libraries generated from methods described herein. Antibodies described herein result in improved functional activity, structural stability, expression, specificity, or a combination thereof. In some instances, the antibody is a single domain antibody. In some instances, the single domain antibody comprises one heavy chain variable domain. In some instances, the single domain antibody is a VHH antibody.

[0059] 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 an antibody, wherein the antibody 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 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 an antibody, wherein the antibody 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.

[0060] 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.

[0061] “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.

[0062] “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.

[0063] “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.

[0064] Methods as described herein may be used for generation of libraries encoding a non-immunoglobulin. In some instances, the libraries comprise antibody mimetics. Exemplary antibody mimetics include, but are not limited to, anticalins, affilins, affibody molecules, affimers, affitins, alphabodies, avimers, atrimers, DARPin, fynomers, Kunitz domain-based proteins, monobodies, anticalins, knottins, armadillo repeat protein-based proteins, and bicyclic peptides.

[0065] Libraries described herein comprising nucleic acids encoding for an antibody comprise variations in at least one region of the antibody. 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).

[0066] Provided herein are libraries comprising nucleic acids encoding for an antibody comprising variation in at least one region of the antibody, wherein the region is the CDR region. In some instances, the 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, the VHH libraries described herein comprise at least or about 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. 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.

[0067] Libraries as described herein may comprise varying lengths of a CDRH1, CDRH2, CDRH3, CDRL1, CDRL2, CDRL3, or combinations thereof of amino acids when translated. In some instances, the length of the CDRH1, CDRH2, CDRH3, CDRL1, CDRL2, CDRL3, or combinations thereof of amino acids when translated is 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, or more than 30 amino acids.

[0068] 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. In some instances, the library is a VHH library. In some instances, the library is an antibody library.

[0069] 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} , or more than 10^{18} 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} , or more than 10^{18} sequences.

[0070] Libraries as described herein encoding for an antibody or immunoglobulin 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} , or more than 10^{18} 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} , or more than 10^{18} sequences.

[0071] Methods described herein provide for synthesis of libraries comprising nucleic acids encoding an antibody or immunoglobulin, 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 antibody 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). 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.

[0072] In some instances, the at least one region of the antibody 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).

[0073] Provided herein are libraries comprising nucleic acids encoding for antibodies, 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 antibody 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.

[0074] Libraries comprising nucleic acids encoding for antibodies or immunoglobulins 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.

[0075] A number of variant sequences for the at least one region of the antibody 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.

[0076] Variant sequences for the at least one region of the antibody, 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.

[0077] Following synthesis of antibody libraries, antibody libraries may be used for screening and analysis. For example, antibody 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. 2B**. In some instances, antibody 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. In some instances, antibody libraries are displayed on the surface of a cell or phage. In some instances, antibody libraries are enriched for sequences with a desired activity using phage display.

[0078] In some instances, the antibody 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 antibody libraries are assayed for antibody 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.

[0079] Antibodies or IgGs generated by methods as described herein comprise improved binding affinity. In some instances, the antibody comprises a binding affinity (e.g., K_D) 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 antibody comprises a K_D of less than 400 nM, less than 350 nM, less than 300 nM, less than 250 nM, less than 200 nM, less than 150 nm, less than 100 nM, less than 50 nM, less than 25 nM, less than 15 nM, or less than 10 nM. In some instances, the antibody comprises a K_D of less than 1 nM. In some instances, the antibody comprises a K_D of less than 1.2 nM. In some instances, the antibody comprises a K_D of less than 2 nM. In some instances, the antibody comprises a K_D of less than 5 nM. In some instances, the antibody comprises a K_D of less than 10 nM. In some instances, the antibody comprises a K_D of less than 13.5 nM. In some instances, the antibody comprises a K_D of less than 15 nM. In some instances, the antibody comprises a K_D of less than 20 nM. In some instances, the antibody comprises a K_D of less than 25 nM. In some instances, the antibody comprises a K_D of less than 30 nM.

[0080] In some instances, the affinity of antibodies or IgGs generated by methods as described herein is at least or about 1.5x, 2.0x, 5x, 10x, 20x, 30x, 40x, 50x, 60x, 70x, 80x, 90x, 100x, 200x, or more than 200x improved binding affinity as compared to a comparator antibody. In some instances, the affinity of antibodies or IgGs generated by methods as described herein is at least or about 1.5x, 2.0x, 5x, 10x, 20x, 30x, 40x, 50x, 60x, 70x, 80x, 90x, 100x, 200x, or more than 200x improved function as compared to a comparator antibody. In some instances, the comparator antibody is an antibody with similar structure, sequence, or antigen target.

[0081] Methods as described herein, in some instances, result in increased yield of antibodies or IgGs. In some instances, the yield is at least or about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, or more than 80 micrograms (ug). In some instances, the yield is in a range of about 5 to about 80, about 10 to about 75, about 15 to about 60, about 20 to about 50, or about 30 to about 40 micrograms (ug).

[0082] *Expression Systems*

[0083] Provided herein are libraries comprising nucleic acids encoding for antibody comprising 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.

[0084] Provided herein are libraries comprising nucleic acids encoding for antibody comprising 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.

[0085] Nucleic acid libraries 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.

[0086] 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.

[0087] Described herein are nucleic acid libraries that are expressed in a vector to generate a construct comprising an antibody. 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.

[0088] Provided herein are libraries comprising nucleic acids encoding for antibodies, 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.

[0089] Diseases and Disorders

[0090] Provided herein are libraries comprising nucleic acids encoding for antibodies or immunoglobulins including VHH antibodies that may have therapeutic effects. In some instances, the antibodies or immunoglobulin result in protein when translated that is used to treat a disease or disorder in a subject. 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, 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.

[0091] In some instances, the disease or disorder is associated with TIGIT dysfunction. In some instances, the disease or disorder is associated with aberrant signaling via TIGIT. In some instances, the disease or disorder is associated with CD3 dysfunction. In some instances, the disease or disorder is associated with aberrant signaling via CD3. In some instances, the disease or disorder is cancer. In some instances, the disease or disorder is a viral infection.

[0092] *Protein Targets*

[0093] Provided herein are libraries comprising nucleic acids encoding for antibodies or immunoglobulins including VHH antibodies that may be designed for various protein targets. In some instances, the protein is an ion channel, G protein-coupled receptor, tyrosine kinase receptor, an immune receptor, a membrane protein, or combinations thereof. In some instances, the protein is a receptor. In some instances, the protein is Glucagon-like peptide 1 (GLP1) receptor. In some instances, the protein is Prostaglandin D2 receptor 2 (DP2 or CRTH2) receptor. In some instances, the protein is an adenosine A2A receptor. In some instances, the protein is T cell immunoreceptor with Ig and ITIM domains (TIGIT). In some instances, the protein is Cluster of Differentiation 47 (CD47). In some instances, the protein is Cluster of Differentiation 3 epsilon (CD3ε).

[0094] Provided herein are antibodies or immunoglobulins, wherein the antibody or immunoglobulin comprises a sequence 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: 1-151. In some instances, the antibody or immunoglobulin sequence comprises at least or about 95% sequence identity to any one of SEQ ID NOs: 1-151. In some instances, the antibody or immunoglobulin sequence comprises at least or about 97% sequence identity to any one of SEQ ID NOs: 1-151. In some instances, the antibody or immunoglobulin sequence comprises at least or

about 99% sequence identity to any one of SEQ ID NOs: 1-151. In some instances, the antibody or immunoglobulin sequence comprises at least or about 100% sequence identity to any one SEQ ID NOs: 1-151. In some instances, the antibody or immunoglobulin sequence comprises 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 NOs: 1-151.

[0095] In some embodiments, the antibody or immunoglobulin sequence comprises complementarity determining regions (CDRs) comprising a sequence as set forth in **Table 1A**, **Table 14B**, **Table 17**, and **Table 20**. In some embodiments, the 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: 46-83, 118-137, or 152-163. In some instances, the antibody or immunoglobulin sequence comprises complementarity determining regions (CDRs) comprising at least or about 95% homology to any one of SEQ ID NOs: 46-83, 118-137, or 152-163. In some instances, the antibody or immunoglobulin sequence comprises complementarity determining regions (CDRs) comprising at least or about 97% homology to any one of SEQ ID NOs: 46-83, 118-137, or 152-163. In some instances, the antibody or immunoglobulin sequence comprises complementarity determining regions (CDRs) comprising at least or about 99% homology to any one of SEQ ID NOs: 46-83, 118-137, or 152-163. In some instances, the antibody or immunoglobulin sequence comprises complementarity determining regions (CDRs) comprising at least or about 100% homology to any one of SEQ ID NOs: 46-83, 118-137, or 152-163. In some instances, the 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: 46-83, 118-137, or 152-163.

Table 1A.

Construct Description	Amino Acid Sequence	SEQ ID NO
IGHV1-69 CDRH1	GGTFSSYA	152
IGHV1-69 CDRH2	IPIFGTA	153
IGHV1-69 CDRH3	CARNNNNNNNNNFDYW	154
IGHV3-23 CDRH1	GFTFSSYA	155
IGHV3-23 CDRH2	ISGSGGST	156
IGHV3-23 CDRH3	CAKNNNNNNNNNNFDYW	157
IGKV1-39 CDRL1	QSISSY	158
IGKV1-39 CDRL2	AAS	159
IGKV1-39 CDRL3	CQQSYSTPNTF	160
IGKV3-20 CDRL1	QSVSSSY	161

IGKV3-20 CDRL2	GAS	162
IGKV3-20 CDRL3	CQQYGSSPNTF	163

[0096] In some embodiments, the 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: 118-120, 129-131, 152, 155, 158, or 161. In some instances, the antibody or immunoglobulin sequence comprises CDR1 comprising at least or about 95% homology of any one of SEQ ID NOs: 118-120, 129-131, 152, 155, 158, or 161. In some instances, the antibody or immunoglobulin sequence comprises CDR1 comprising at least or about 97% homology to any one of SEQ ID NOs: 118-120, 129-131, 152, 155, 158, or 161. In some instances, the antibody or immunoglobulin sequence comprises CDR1 comprising at least or about 99% homology to any one of SEQ ID NOs: 118-120, 152, 155, 158, or 161. In some instances, the antibody or immunoglobulin sequence comprises CDR1 comprising at least or about 100% homology to any one of SEQ ID NOs: 118-120, 129-131, 152, 155, 158, or 161. In some instances, the 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 NOs: 118-120, 129-131, 152, 155, 158, or 161.

[0097] In some embodiments, the 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: 121-123, 132-134, 153, 156, 159, or 162. In some instances, the antibody or immunoglobulin sequence comprises CDR2 comprising at least or about 95% homology to any one of SEQ ID NOs: 121-123, 132-134, 153, 156, 159, or 162. In some instances, the antibody or immunoglobulin sequence comprises CDR2 comprising at least or about 97% homology to any one of SEQ ID NOs: 121-123, 132-134, 153, 156, 159, or 162. In some instances, the antibody or immunoglobulin sequence comprises CDR2 comprising at least or about 99% homology to any one of SEQ ID NOs: 121-123, 132-134, 153, 156, 159, or 162. In some instances, the antibody or immunoglobulin sequence comprises CDR2 comprising at least or about 100% homology to any one of SEQ ID NOs: 121-123, 132-134, 153, 156, 159, or 162. In some instances, the 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 NOs: 121-123, 132-134, 153, 156, 159, or 162.

[0098] In some embodiments, the 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: 46-83, 124-128, 154, 157, 160, or 163. In some instances, the antibody or immunoglobulin sequence comprises CDR3 comprising at

least or about 95% homology to any one of SEQ ID NOs: 46-83, 124-128, 125-137, 154, 157, 160, or 163. In some instances, the antibody or immunoglobulin sequence comprises CDR3 comprising at least or about 97% homology to any one of SEQ ID NOs: 46-83, 124-128, 125-137, 124-128, 154, 157, 160, or 163. In some instances, the antibody or immunoglobulin sequence comprises CDR3 comprising at least or about 99% homology to any one of SEQ ID NOs: 46-83, 124-128, 125-137, 124-128, 154, 157, 160, or 163. In some instances, the antibody or immunoglobulin sequence comprises CDR3 comprising at least or about 100% homology to any one of SEQ ID NOs: 46-83, 124-128, 125-137, 124-128, 154, 157, 160, or 163. In some instances, the 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 NOs: 46-83, 124-128, 125-137, 124-128, 154, 157, 160, or 163.

[0099] In some embodiments, the 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 any one of SEQ ID NOs: 152; 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 any one of SEQ ID NOs: 153; 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 any one of SEQ ID NOs: 154. In some instances, the antibody or immunoglobulin sequence comprises CDRH1 comprising at least or about 95%, 97%, 99%, or 100% homology to any one of SEQ ID NOs: 152; a CDRH2 comprising at least or about 95%, 97%, 99%, or 100% homology to any one of SEQ ID NOs: 153; and a CDRH3 comprising at least or about 95%, 97%, 99%, or 100% homology to any one of SEQ ID NOs: 154. In some instances, the 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: 152; 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: 153; 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: 154.

[00100] In some embodiments, the 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: 155; 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: 156; 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: 157. In some instances, the antibody or immunoglobulin sequence comprises CDRH1 comprising at least or

about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 155; a CDRH2 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 156; and a CDRH3 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 157. In some instances, the 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: 155; 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: 156; 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: 157.

[00101] In some embodiments, the 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: 158; 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: 159; 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: 160. In some instances, the antibody or immunoglobulin sequence comprises CDRL1 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 158; a CDRL2 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 159; and a CDRL3 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 160. In some instances, the 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: 158; 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: 159; 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: 160.

[00102] In some embodiments, the 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: 161; 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: 162; 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: 163. In some instances, the antibody or immunoglobulin sequence comprises CDRL1 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 161; a CDRL2 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 162; and a CDRL3 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 163. In some instances, the 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: 161; 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: 162; 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: 163.

[00103] In some embodiments, the 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: 152; 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: 153; 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: 154, 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: 158; 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: 159; 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: 160. In some instances, the antibody or immunoglobulin sequence comprises CDRH1 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 152; a CDRH2 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 153; a CDRH3 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 154; a CDRL1 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 158; a CDRL2 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 159; and a CDRL3 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 160. In some instances, the 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: 152; 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: 153; 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: 154; 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: 158; 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: 159; 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: 160.

[00104] In some embodiments, the 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: 152; 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: 153; 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: 154, 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: 161; 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: 162; 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: 163. In some instances, the antibody or immunoglobulin sequence comprises CDRH1 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 152; a CDRH2 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 153; a CDRH3 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 154; a CDRL1 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 161; a CDRL2 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 162; and a CDRL3 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 163. In some instances, the 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: 152; 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: 153; 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: 154; 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: 161; 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: 162; 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: 163.

[00105] In some embodiments, the 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: 155; 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: 156; 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: 157, 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: 158; 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: 159; 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: 160. In some instances, the antibody or immunoglobulin sequence comprises CDRH1 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 155; a CDRH2 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 156; a CDRH3 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 157; a CDRL1 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 158; a CDRL2 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 159; and a CDRL3 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 160. In some instances, the 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: 155; 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: 156; 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: 157; 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: 158; 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: 159; 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: 160.

[00106] In some embodiments, the 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: 155; 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: 156; 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: 157, 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: 161; 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: 162; 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: 163. In some instances, the antibody or immunoglobulin sequence comprises CDRH1 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 155; a CDRH2 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 156; a CDRH3 comprising at least or

about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 157; a CDRL1 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 161; a CDRL2 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 162; and a CDRL3 comprising at least or about 95%, 97%, 99%, or 100% homology to SEQ ID NO: 163. In some instances, the 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: 155; 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: 156; 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: 157; 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: 161; 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: 162; 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: 163.

[00107] 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 NOs: 1-23 or 126-148. 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 NOs: 1-23 or 126-148. 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 NOs: 1-23 or 126-148. 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 NOs: 1-23 or 126-148. 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 NOs: 1-23 or 126-148. 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 SEQ ID NOs: 1-23 or 126-148.

[00108] 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 NOs: 24-45 or 149-151. 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 NOs: 24-45 or 149-151. 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 NOs: 24-45 or 149-151. 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 NOs: 24-45 or 149-151. 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 NOs: 24-45 or 149-151. 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 SEQ ID NOs: 24-45 or 149-151.

[00109] Provided herein are antibodies or immunoglobulins for various protein targets. In some instances, the protein is TIGIT. Described herein, in some embodiments, are antibodies or immunoglobulins that bind to the TIGIT. In some instances, the TIGIT 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 NOs: 84-100. In some instances, the TIGIT antibody or immunoglobulin sequence comprises a heavy chain variable domain comprising at least or about 95% sequence identity to any one of SEQ ID NOs: 84-100. In some instances, the TIGIT antibody or immunoglobulin sequence comprises a heavy chain variable domain comprising at least or about 97% sequence identity to any one of SEQ ID NOs: 84-100. In some instances, the TIGIT antibody or immunoglobulin sequence comprises a heavy chain variable domain comprising at least or about 99% sequence identity to any one of SEQ ID NOs: 84-100. In some instances, the TIGIT antibody or immunoglobulin sequence comprises a heavy chain variable domain comprising at least or about 100% sequence identity to any one of SEQ ID NOs: 84-100. In some instances, the TIGIT 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 NOs: 84-100.

[00110] In some instances, the TIGIT 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 NOs: 101-117. In some instances, the TIGIT antibody or immunoglobulin sequence comprises a light chain variable domain comprising at least or about 95% sequence identity to any one of SEQ ID NOs: 101-117. In some instances, the TIGIT antibody or immunoglobulin sequence comprises a light chain variable domain comprising at least or about 97% sequence identity to any one of SEQ ID NOs: 101-117. In some instances, the TIGIT antibody or immunoglobulin sequence comprises a light chain variable domain comprising at least or about 99% sequence identity to any one of SEQ ID NOs: 101-117. In some instances, the TIGIT antibody or immunoglobulin sequence comprises a light chain variable domain comprising at least or about 100% sequence identity to any one of SEQ ID NOs: 101-117. In some instances, the TIGIT 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 NOs: 101-117.

[00111] In some instances, the protein is CD3 epsilon. Described herein, in some embodiments, are antibodies or immunoglobulins that bind to the CD3. In some instances, the CD3 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 NOs: 138-141. In some instances, the CD3 antibody or immunoglobulin sequence comprises a heavy chain variable domain comprising at least or about 95% sequence identity to any one of SEQ ID NOs: 138-141. In some instances, the CD3 antibody or immunoglobulin sequence comprises a heavy chain variable domain comprising at least or about 97% sequence identity to any one of SEQ ID NOs: 138-141. In some instances, the CD3 antibody or immunoglobulin sequence comprises a heavy chain variable domain comprising at least or about 99% sequence identity to any one of SEQ ID NOs: 138-141. In some instances, the CD3 antibody or immunoglobulin sequence comprises a heavy chain variable domain comprising at least or about 100% sequence identity to any one of SEQ ID NOs: 138-141. In some instances, the CD3 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 NOs: 138-141.

[00112] In some instances, the CD3 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 NOs: 142-145. In some instances, the CD3 antibody or immunoglobulin sequence comprises a light chain variable domain comprising at least or about 95% sequence identity to any one of SEQ ID NOs: 142-145. In some instances, the CD3 antibody or immunoglobulin sequence comprises a light chain variable domain comprising at least or about 97% sequence identity to any one of SEQ ID NOs: 142-145. In some instances, the CD3 antibody or immunoglobulin sequence comprises a light chain variable domain comprising at least or about 99% sequence identity to any one of SEQ ID NOs: 142-145. In some instances, the CD3 antibody or immunoglobulin sequence comprises a light chain variable domain comprising at least or about 100% sequence identity to any one of SEQ ID NOs: 142-145. In some instances, the CD3 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 NOs: 142-145.

[00113] Variant Libraries

[00114] *Codon variation*

[00115] 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 1B** provides a listing of each codon possible (and the representative amino acid) for a variant site.

Table 1B. 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				

[00116] 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.

[00117] Highly Parallel Nucleic Acid Synthesis

[00118] 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.

[00119] 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.

[00120] 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.

[00121] 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.

[00122] 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 .

[00123] 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.

[00124] 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.

[00125] Substrates

[00126] 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.

[00127] 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.

[00128] 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.

[00129] 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 .

[00130] 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.

[00131] 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 mm^2 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.

[00132] *Surface materials*

[00133] 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.

[00134] **Surface Architecture**

[00135] 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.

[00136] 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.

[00137] 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.

[00138] 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.

[00139] 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 .

[00140] 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 .

[00141] 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 .

[00142] Surface Modifications

[00143] 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.

[00144] 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.

[00145] 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.

[00146] 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.

[00147] 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 alkoxysilane 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 alkoxysilanes are classified according to their organic functions.

[00148] Polynucleotide Synthesis

[00149] 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.

[00150] 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).

[00151] 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.

[00152] 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).

[00153] 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.

[00154] 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.

[00155] 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.

[00156] 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.

[00157] 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.

[00158] 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.

[00159] 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.

[00160] 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.

[00161] Referring to the Figures, **FIG. 1** 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.

[00162] 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.

[00163] In situ preparation of polynucleotide arrays is generated on a solid support 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 **102**. In some instances, polynucleotides are cleaved from the surface at this stage. Cleavage includes gas cleavage, *e.g.*, with ammonia or methylamine.

[00164] 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 **103**. Prior to or after the sealing **104** 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 **105**. Once released, fragments of single stranded polynucleotides hybridize in order to span an entire long range sequence of DNA. Partial hybridization **105** is possible because each

synthesized polynucleotide is designed to have a small portion overlapping with at least one other polynucleotide in the pool.

[00165] 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 **106**.

[00166] After PCA is complete, the nanoreactor is separated from the device **107** and positioned for interaction with a device having primers for PCR **108**. After sealing, the nanoreactor is subject to PCR **109** and the larger nucleic acids are amplified. After PCR **110**, the nanochamber is opened **111**, error correction reagents are added **112**, the chamber is sealed **113** and an error correction reaction occurs to remove mismatched base pairs and/or strands with poor complementarity from the double stranded PCR amplification products **114**. The nanoreactor is opened and separated **115**. Error corrected product is next subject to additional processing steps, such as PCR and molecular bar coding, and then packaged **122** for shipment **123**.

[00167] 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 **116**, sealing the wafer to a chamber containing error corrected amplification product **117**, and performing an additional round of amplification **118**. The nanoreactor is opened **119** and the products are pooled **120** and sequenced **121**. After an acceptable quality control determination is made, the packaged product **122** is approved for shipment **123**.

[00168] In some instances, a nucleic acid generated by a workflow such as that in **FIG. 1** 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 **102**.

[00169] Computer systems

[00170] 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.

[00171] The computer system **200** illustrated in **FIG. 2** may be understood as a logical apparatus that can read instructions from media **211** and/or a network port **205**, which can optionally be connected to server **209** having fixed media **212**. The system, such as shown in **FIG. 2** can include a CPU **201**, disk drives **203**, optional input devices such as keyboard **215** and/or mouse **216** and optional monitor **207**. 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 **222** as illustrated in **FIG. 2**.

[00172] As illustrated in **FIG. 3**, a high speed cache **304** can be connected to, or incorporated in, the processor **302** to provide a high speed memory for instructions or data that have been recently, or are frequently, used by processor **302**. The processor **302** is connected to a north bridge **306** by a processor bus **308**. The north bridge **306** is connected to random access memory (RAM) **310** by a memory bus **312** and manages access to the RAM **310** by the processor **302**. The north bridge **306** is also connected to a south bridge **314** by a chipset bus **316**. The south bridge **314** is, in turn, connected to a peripheral bus **318**. 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 **318**. 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 **300** can include an accelerator card **322** attached to the peripheral bus **318**. 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.

[00173] Software and data are stored in external storage **324** and can be loaded into RAM **310** and/or cache **304** for use by the processor. The system **300** includes an operating system for managing system resources; non-limiting examples of operating systems include: Linux, Windows™, MACOSTM, BlackBerry OSTM, iOSTM, 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 **300** also includes network interface cards (NICs) **320** and **321** 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.

[00174] **FIG. 4** is a diagram showing a network **400** with a plurality of computer systems **402a**, and **402b**, a plurality of cell phones and personal data assistants **402c**, and Network Attached Storage (NAS) **404a**, and **404b**. In example instances, systems **402a**, **402b**, and **402c** can manage data storage and optimize data access for data stored in Network Attached Storage (NAS) **404a** and **404b**. A mathematical model can be used for the data and be evaluated using distributed parallel processing across computer systems **402a**, and **402b**, and cell phone and personal data assistant systems **402c**. Computer systems **402a**, and **402b**, and cell phone and personal data assistant systems **402c** can also provide parallel processing for adaptive data restructuring of the data stored in Network Attached Storage (NAS) **404a** and **404b**. **FIG. 4** 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.

[00175] **FIG. 5** is a block diagram of a multiprocessor computer system **500** using a shared virtual address memory space in accordance with an example instance. The system includes a plurality of processors **502a-f** that can access a shared memory subsystem **504**. The system incorporates a plurality of programmable hardware memory algorithm processors (MAPs) **506a-f** in the memory subsystem **504**. Each MAP **506a-f** can comprise a memory **508a-f** and one or more field programmable gate arrays (FPGAs) **510a-f**. The MAP provides a configurable functional unit and particular algorithms or portions of algorithms can be provided to the FPGAs **510a-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 **508a-f**, allowing it to execute tasks independently of, and

asynchronously from the respective microprocessor **502a-f**. In this configuration, a MAP can feed results directly to another MAP for pipelining and parallel execution of algorithms.

[00176] 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.

[00177] 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. 3**, 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 **322** illustrated in **FIG. 3**.

[00178] 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

[00179] 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.

[00180] Example 1: Functionalization of a device surface

[00181] 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.

[00182] 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. SPR™ 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.

[00183] 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.

[00184] Example 2: Synthesis of a 50-mer sequence on an oligonucleotide synthesis device

[00185] 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-TRIETHOXSILYLPROPYL)-4-HYDROXYBUTYRAMIDE (Gelest) was used to synthesize an exemplary polynucleotide of 50 bp ("50-mer polynucleotide") using polynucleotide synthesis methods described herein.

[00186] The sequence of the 50-mer was as described in SEQ ID NO.: 104.

5'AGACAATCAACCATTTGGGGTGGACAGCCTTGACCTCTAGACTTCGGCAT##TTTTTTT TTT3' (SEQ ID NO.: 104), 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.

[00187] The synthesis was done using standard DNA synthesis chemistry (coupling, capping, oxidation, and deblocking) according to the protocol in **Table 2** and an ABI synthesizer.

Table 2: Synthesis protocols

General DNA Synthesis Process Name	Table 2	
	Process Step	Time (sec)
WASH (Acetonitrile Wash Flow)	Acetonitrile System Flush	4
	Acetonitrile to Flowcell	23
	N2 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
	Incubate for 25sec	25
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

Table 2		
General DNA Synthesis Process Name	Process Step	Time (sec)
CAPPING (CapA+B, 1:1, Flow)	Acetonitrile System Flush	4
	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

[00188] 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.

[00189] 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.

[00190] Example 3: Synthesis of a 100-mer sequence on an oligonucleotide synthesis device

[00191] 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.: 105) 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-acetoxyundecyltriethoxysilane and n-decyltriethoxysilane, and the polynucleotides extracted from the surface were analyzed on a BioAnalyzer instrument.

[00192] All ten samples from the two chips were further PCR amplified using a forward (5'ATGCGGGGTTCTCATCATC3'; SEQ ID NO.: 106) and a reverse (5'CGGGATCCTTATCGTCATCG3'; SEQ ID NO.: 107) 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

[00193] 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 3** 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 3: 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%

Spot	Error rate	Cycle efficiency
7	1/531 bp	99.81%
8	1/1769 bp	99.94%
9	1/854 bp	99.88%
10	1/1451 bp	99.93%

[00194] 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.

[00195] **Table 4** summarizes error characteristics for the sequences obtained from the polynucleotide samples from spots 1-10.

Table 4: 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 Single Base Deletion	0	0	0	0	0	0	0	0	0	0
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

[00196] Example 4: VHH Libraries

[00197] 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.

[00198] 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 (data not shown). **Table 5A** provides specific values for the VHH-Fc clones for ELISA, Protein A (mg/ml), and K_D (nM). **FIG. 7A** and **FIG. 7B** depict TIGIT affinity distribution for the VHH libraries, over the 20 – 4000 affinity threshold (**FIG. 7A**; monovalent K_D) and the 20 – 1000 affinity threshold (**FIG. 7B**; monovalent K_D). Out of the 140 VHH binders tested, 51 variants had affinity < 100 nM, and 90 variants had affinity < 200 nM. **FIG. 8** shows data of CDR3 counts per length for the ‘VHH ratio’ library, the ‘VHH shuffle library,’ and the ‘VHH hShuffle library.’ **Table 5B** shows number of TIGIT unique clones and

TIGIT binders for the 'VHH ratio' library, the 'VHH shuffle library,' and the 'VHH hShuffle library.'

Table 5A.

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

Table 5B. TIGIT unique clones and TIGIT binders

Library	Unique Phage	VHH-Fc binders
VHH Ratio	47	36
VHH Shuffle	58	45
VHH hShuffle	56	53

[00199] Thermostability and competition analysis of the VHH-Fc TIGIT clones is seen in **FIG. 9** and **Table 6**. For the competition assays, 4 ug/mL TIGIT was immobilized and incubated with 0.05 – 100 nM VHH-Fc followed by incubation with 2 ug/mL biotin-CD155 and 1:5000 streptavidin-HRP.

Table 6. Thermostability of VHH-Fc TIGIT clones

Variant	Library	K _D (nM)	T _{m1}	T _{m2}	IC50 (nM)
TIGIT-29-10	Ratio	32	72	87	17.65
TIGIT-29-7	Ratio	41	82	90	9.24
TIGIT-30-30	Shuffle	23	76	87	5.67
TIGIT-30-43	Shuffle	44	82	90	2.32
TIGIT-31-1	hShuffle	12	79	89	17.89
TIGIT-31-6	hShuffle	14	77	87	4.00
TIGIT-31-26	hShuffle	19	79	89	8.20
TIGIT-31-32	hShuffle	27	80	86	2.85
TIGIT-31-8	hShuffle	45	76	84	3.92
TIGIT-31-56	hShuffle	46	74	83	1.52

[00200] CD47 VHH variants were also generated and analyzed. **FIG. 10** shows the CD47 affinity distribution. **Table 7** shows number of CD47 unique clones and TIGIT binders for the ‘VHH ratio’ library, the ‘VHH shuffle library,’ and the ‘VHH hShuffle library.’ **Table 8** shows the binding affinity of the CD47 VHH variants. As seen in **Table 8**, 8 CD47 VHH binders had an affinity less than 100 nM to hCD47 and 6 CD47 VHH binders had an affinity less than 100 nM to cCD47.

Table 7. VHH-Fc CD47 clones

Library	Unique Phage	VHH-Fc binders
VHH Ratio	3	2
VHH Shuffle	1	1
VHH hShuffle	11	6

Table 8. VHH-Fc CD47 binding affinities

Variant	hCD47 K _D (nM)	cCD47 K _D (nM)
CD47-19-2	35	93
CD47-19-3	200	-
CD47-20-1	49	105
CD47-21-1	28	80
CD47-21-2	31	80
CD47-21-3	19	43
CD47-21-4	62	265
CD47-21-6	71	69
CD47-21-10	38	35

[00201] Inhibition and thermostability analysis of the VHH-Fc CD47 clones is seen in **FIG. 11** and **Table 9**. For the inhibition assays, 3 ug/mL of CD47 was immobilized and incubated with 0.3-132 nM of VHH-Fc followed by incubation with 0.25 ug/mL biotin-SIRP alpha and 1:5000 streptavidin-HRP.

Table 9. Thermostability of VHH-Fc CD47 clones

Variant	K _D (nM)	T _{m1}	T _{m2}	SIRPalph IC50 (nM)
CD47-19-2	35	75	88	1.13
CD47-19-3	200	76	87	1.08
CD47-20-1	49	78	89	1.79
CD47-21-1	28	80	88	1.68
CD47-21-2	31	69	88	-
CD47-21-3	19	80	88	1.35
CD47-21-4	62	81	89	-
CD47-21-6	71	79	88	-
CD47-21-10	38	71	85	-

[00202] Example 5. VHH Libraries for GLP1R

[00203] A VHH library for GLP1R was developed similar to methods described in Example 4. Briefly, stable cell lines expressing GLP1R 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.

[00204] 53 out of the 156 unique GLP1R VHH Fc binders had a target cell mean fluorescence intensity (MFI) value that was 2-fold over parental cells. The data for variant GLP1R-43-77 is seen in **FIGS. 12A-12B** and **Tables 10-11**. **Table 11** shows flow cytometry data as detected with the RL1-A channel.

Table 10. Panning summary

Library	Unique Phage	VHH-Fc FACS binders (MFI values 2-fold over parental cells)
VHH hShuffle	58	6
VHH Ratio/ Shuffle	98	47

Table 11. GLP1R-43-77 data

Subset Name with Gating Path	Count	Median:RL1-A
Sample E10.fcs/CHO-parent	11261	237
Sample E10.fcs/CHO-GLP1R	13684	23439

[00205] Example 6. VHH Libraries for CRTH2R

[00206] A VHH library for CRTH2R was developed similar to methods described in Example 4. 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.

[00207] 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. 13A-13B** and **Tables 12-13**. **Table 13** shows flow cytometry data as detected with the RL1-A channel. Data for variant CRTH2-44-59 is seen in **FIGS. 14A-14D**.

Table 12. 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 13. 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

[00208] Example 7. Identification of IgGs for CRTH2R

[00209] 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 100 nM (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. 15A-15E**. An exemplary gated dot plot and APC histogram at 100 nM with CRTH2-27 is shown in **FIGS. 16A-6B**. Two antibodies (gPCR-51 and gPCR-52) were used as a positive control. Binding profiles of the two positive controls are shown in **FIGS. 17A-17B**.

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

SEQ ID NO	CRTH2R Antibody	Heavy Chain
1	CRTH2-74	QVQLVESGGGVVQPGRSLRLSCAASGFSFSEYGIHWVRQAPGKGLEWVAVISYE GSNEYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCARANQHFGPVA GGATPSEEPGSQLTRAELGWDAPPGQESLADELLQLGTEHGYHYYGMDVWGQG TLVTVSS
2	CRTH2-24	QVQLVQSGAEVKKPGSSVKVSCASGGSFSNYGISWVRQAPGQGLEWMGGIPLI GTANYAQKFQGRVTITADESTSTAYMELSSLRSEDVAVYYCARDMYDYFTLGPQ SIGPLGEVVPADDAFDIWGQGLTVTVSS
3	CRTH2-28	QVQLVQSGAEVKKPGSSVNVSCASGGTFSDYAFSWVRQAPGQGLEWMGAIIPF FGTVNYAQKFQGRVTITADESTSTAYMELSSLRSEDVAVYYCARDMYDYFATGT GGPEDDLYPQGEINDGYRIEVPADDAFDIWGQGLTVTVSS
4	CRTH2-39	QVQLVQSGAEVKKPGSSVKVSCASVDTFSRYSISWVRQAPGQGLEWMGGIIPV FDTTNYAQKFQGRVTITADESTSTAYMELSSLRSEDVAVYYCARDMYDYDFGVIL GGTAVGTNNGSANEVVPADDAFDIWGQGLTVTVSS
5	CRTH2-19	QVQLVQSGAEVKKPGSSVKVSCASGGTFSSHAINWVRQAPGQGLEWMGRIIPV GTTTYAQKFQGRVTITADESTSTAYMELSSLRSEDVAVYYCARDMYDYDFYFGL

		TLTGDRNDDEVVPADDAFDIWGQGLTVTVSS
6	CRTH2-9	QVQLVQSGAEVKKPGSSVKVSKASGGTFSSYAISWVRQAPGQGLEWMGGIIPF GTANYAQKFQGRVTITADESTSTAYMELSSLRSEDTAVYYCARDMYDYDFWLGD QSTGSLIGAEVVPADDAFDIWGQGLTVTVSS
7	CRTH2-8	QVQLVQSGAEVKKPGSSVKVSKASGGTFSSYAISWVRQAPGQGLEWMGGIIPF FGSPNYAQKFQGRVTITADESTSTAYMELSSLRSEDTAVYYCARDMYDYDFAAGL EGTITEVFDEEGHQGGTEVVPADDAFDIWGQGLTVTVSS
8	CRTH2-27	QVQLVESGGGVVQPGRSLRLSCAASGFTFDNYGMHWVRQAPGKGLEWVAVISY EGSNKKYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCARDMYDYDFGSI YGEDVVGELPEVVPADDAFDIWGQGLTVTVSS
9	CRTH2-45	QVQLVESGGGVVQPGRSLRLSCAASGFTFSHYAMHWVRQAPGKGLEWVADISH EGSNKYYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCARDGRGSLPRP KGGPTSGGGFSTNIGYGFVVSQSYDSSGGAFFDIWGQGLTVTVSS
10	CRTH2-35	QVQLVQSGAEVKKPGSSVKVSKASGGTFRSYAISWVRQAPGQGLEWMGGIIPIS GTTNYAQKFQGRVTITADESTSTAYMELSSLRSEDTAVYYCARANQHFTRIFGNY QIYFGHFGYHYGMDVWGQGLTVTVSS
11	CRTH2-50	QVQLVQSGAEVKKPGSSVKVSKASGGTFSSYALSWVRKAPGQGLEWMGGTIPI FGTVNYAQKFQGRVTITADESTSTAYMELSSLRSEDTAVYYCARANQHFTRVIGQ PSPAVPSRGYIYHGYHYGMDVWGQGLTVTVSS
12	CRTH2-66	QVQLVESGGGVVQPGRSLRLSCAASGFDGSGYGMHWVRQAPGKGLEWVAVISY EGSNKFYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCARDLRELECEE WTIEVHGQEFVHQRGGVFSRGPVDPGRVAGSFDVWGQGLTVTVSS
13	CRTH2-57	QVQLVQSGAEVKKPGSSVKVSKASGGTFSSYAMSWVRQAPGQGLEWMGGIIP FGTTDYAQKFQGRVTITADESTSTAYMELSSLRSEDTAVYYCARANQHFVKIQGA PVSTPVPFGTTGYHYGMDVWGQGLTVTVSS
14	CRTH2-32	QVQLVESGGGVVQPGRSLRLSCAASGFTFSKHGMHWVRQAPGKGLEWVAFISY EGSEKYYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCARDMYDYDFHY STVGATYYYYLGSETEVVPADDAFDIWGQGLTVTVSS
15	CRTH2-15	QVQLVQSGAEVKKPGSSVKVSKASGGTFSTY AIDWVRQAPGQGLEWMGGIIP FGSPNYAQKFQGRVTITADESTSTAYMELSSLRSEDTAVYYCARANQHFFLYEGT SSSWLHVGHARYGYHYGMDVWGQGLTVTVSS
16	CRTH2-25	QVQLVQSGAEVKKPGSSVKVSKASGGTFRSYGISWVRQAPGQGLEWMGRIIP GTPDYAQKFQGRVTITADESTSTAYMELSSLRSEDTAVYYCARDMYDYDFEDVDE GSLYLDMGRTFEVVPADDAFDIWGQGLTVTVSS
17	CRTH2-42	QVQLVESGGGVVQPGRSLRLSCAASGFAFSSYAMHWVRQAPGKGLEWVAVISY EGSNEYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCARDLRELECEE WTVLQYQKGFHMRWAESGEGSLRGPVDPGRVAGSFDVWGQGLTVTVSS
18	CRTH2-55	QVQLVESGGGVVQPGRSLRLSCAASGTFSSYDMHWVRQAPGKGLEWVAVISY EGSEKYYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAKHMMSMQAST EGDFGLEEVTEGEGVDDRDLVGDAFDVWGQGLTVTVSS
19	CRTH2-60	QVQLVQSGAEVKKPGSSVKVSKASGGTFKNYAINWVRQAPGQGLEWMGAIIP KFGAANYAQKFQGRVTITADESTSTAYMELSSLRSEDTAVYYCARANQHFSVR GLAFGYGYRIGGYHYGMDVWGQGLTVTVSS
20	CRTH2-70	QVQLVQSGAEVKKPGSSVKVSKASGGTFSTNHAIWVRQAPGQGLEWMGGIIP GTPSYAQKFQGRVTITADESTSTAYMELSSLRSEDTAVYYCARDMYDYDFDVISAG VVGAGNPEVVPADDAFDIWGQGLTVTVSS
21	CRTH2-48-9	EVQLLESGGGLVQPGGSLRLSCAASGFSFSTHMSWVRQAPGKGLEWVSTIGGS GGSTYYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCARAHGDSSSWYF SYYMDVWGQGLTVTVSS
22	CRTH2-41-51	EVQLVESGGGLVQPGGSLRLSCAASGGIFRFNAMGWFRQAPGKERELVAGISGSG GDTYYADSVKGRFTISADNSKNTAYLQMNSLKPEDTAVYYCAAFRGIMRPDWG QGLTVTVSS
23	CRTH2-44-6	EVQLVESGGGLVQPGGSLRLSCAASGPTFDTYVMGWFRQAPGKEREFVAAISMS GDDTAYADSVKGRFTISADNSKNTAYLQMNSLKPEDTAVYYCATDLRGRGDVS EYEDVWGQGLTVTVSS
SEQ ID NO	CRTH2R Antibody	Light Chain
24	CRTH2-74	QSVLTQPPSVSAAPGQKVTISCSGSTSNIGKNYVSWYQQLPGTAPKLLIYDDDERP SGIPDRFSGMSGTSATLGITGLQTGDEADYYCEAWDADLSGAVFGGGTKLTVL

25	CRTH2-24	QSVLTQPPSVSAAPGQKVTISCSGSSSNIGNNFVSWYQQLPGTAPKLLIYDNIQRPS GIPDRFSGSKSGTSATLGITGLQTGDEADYYCGTWDTSLSAVVFGGGTKLTVLRT VAAPSVFIFPPSDEQLKSGTASVCLLNFFYPREKQVQWKVDNALQSGNSQESVT EQDSKDSTYSLSTLTLTKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC
26	CRTH2-28	QSVLTQPPSVSAAPGQKVTISCSGSSSNIGNNFVSWYQQLPGTAPKLLIYDDHKRP SGIPDRFSGSKSGTSATLGITGLQTGDEADYYCATWDRGLSAAVFGGGTKLTVL
27	CRTH2-39	QSVLTQPPSVSAAPGQKVTISCSGSSSNIGNDNDVSWYQQLPGTAPKLLIYDDDKRP SGIPDRFSGSKSGTSATLGITGLQTGDEADYYCASWDTSLSGGYVFGGGTKLTVL
28	CRTH2-19	QSALTQPASVSGSPGQSITISCTGTSSDVGGYDYVTWYQHPGKAPKLMYDVDVT RPSGVSNRFSGSKSGNTASLTISGLQAEDEADYYCSSYTTSTSYVFGGGTKLTVL
29	CRTH2-9*	QSVLTQPPSVSAAPGQKVTISCSGSSSNIGNNFVSWYQQLPGTAPKLLIYENDERP SGIPDRFSGSKSGTSATLGITGLQTGDEADYYCATWDRLSAAVFGGGTKLTVL
30	CRTH2-8	QSVLTQPPSVSAAPGQKVTISCSGSSSNIGNNFVSWYQQLPGTAPKLLIYDNNQRP SGIPDRFSGSKSGTSATLGITGLQTGDEADYYCGTWDTSLSVFGGGTKLTVL
31	CRTH2-27	QSALTQPASVSGSPGQSITISCTGTSSDVGAYNFVSWYQHPGKAPKLMYDISNR PSGVSNRFSGSKSGNTASLTISGLQAEDEADYYCSSYTRSNTRVFGGGTKLTVL
32	CRTH2-45	QSVLTQPPSVSAAPGQKVTISCSGSSSNIGNNFVSWYQQLPGTAPKLLIYDNNVCRP SGIPDRFSGSKSGTSATLGITGLQTGDEADYYCGTWDNTVSAPWVFGGGTKLTVL
33	CRTH2-35	QSALTQPASVSGSPGQSITISCTGTSSDIGGYEFVSWYQHPGKAPKLMYGVSRP PSGVSNRFSGSKSGNTASLTISGLQAEDEADYYCGSYTSSSTPYVFGGGTKLTVL
34	CRTH2-50	QSALTQPASVSGSPGQSITISCTGTSSDIGGYNFVSWYQHPGKAPKLMYDVSNR PQGVSNRFSGSKSGNTASLTISGLQAEDEADYYCSSYTSSNTYVFGGGTKLTVL
35	CRTH2-66	EIVMTQSPATLSVSPGERATLSCRASQGVGSNLAWYQKPKGAPRLIYRTSIRAT GIPARFSGSGSGTEFTLTISLSQSEDAVYYCQYYSWPPLTFGGGTKVEIK
36	CRTH2-57	QSVLTQPPSVSAAPGQKVTISCSGSSSNIGNNFVSWYQQLPGTAPKLLIYDNFKRP SGIPDRFSGSKSGTSATLGITGLQTGDEADYYCGTWDTSLSAALFGGGTKLTVL
37	CRTH2-32	QSALTQPASVSGSPGQSITISCTGTSSGVGGYDYVSWYQHPGKAPKLMYDDNN RPSGVSNRFSGSKSGNTASLTISGLQAEDEADYYCSSYTSSSTLYVFGGGTKLTVL
38	CRTH2-15	QSVLTQPPSVSAAPGQKVTISCSGSSSNIGNNFVSWYQQLPGTAPKLLIYDNIIRPS GIPDRFSGSKSGTSATLGITGLQTGDEADYYCAAWDTRLSAGVFGGGTKLTVL
39	CRTH2-25	DIQMTQSPSSLSASVGDRTITCRASQGISSTYLNWYQKPKGAPKLLIYATSSLSQ GVPSRFSGSGSGTDFTLTISLQPEDFATYYCQQSYSTP-WTFGGGTKVEIK
40	CRTH2-42	QSALTQPASVSGSPGQSITISCTGTSSDVGGYRYVSWYQHPGKAPKLMYNNVNY RPSGVSNRFSGSKSGNTASLTISGLQAEDEADYYCSSYRSSSTLVFGGGTKLTVL
41	CRTH2-55	QSVLTQPPSVSAAPGQKVTISCSGSSSNIGNNFVSWYQQLPGTAPKLLIYDDDERP SGIPDRFSGSKSGTSATLGITGLQTGDEADYYCGAWDRSLSAVFGGGTKLTVL
42	CRTH2-60	QSVLTQPPSVSAAPGQKVTISCSGSSSNIGNNFVSWYQQLPGTAPKLLIYENRKRPS GIPDRFSGSKSGTSATLGITGLQTGDEADYYCATWDASLKNLVFGGGTKLTVL
43	CRTH2-70	QSVLTQPPSVSAAPGQKVTISCSGSSSNIGNNFVSWYQQLPGTAPKLLIYDNEKRP SGIPDRFSGSKSGTSATLGITGLQTGDEADYYCGTWDERQTDSEYVFGGGTKLTV L
44	CRTH2-9A	QSVLTQPPSVSAAPGQKVTISCSGSSSNIGNNFVSWYQQLPGTAPKLLIYENDERP SGIPDRFSGSKSGTSATLGITGLQTGDEADYYCATWDRLSAAVFGGGETKLT
45	CRTH2-48-9	DIQMTQSPSSLSASVGDRTITCRASQGISDYVNWYQKPKGAPKLLIYGASILQT GVPSRFSGSGSGTDFTLTISLQPEDFATYYCQQSFTTPWTFGGGTKVEIK

Table 14B. Variably Heavy Chain CDR3 Sequences

SEQ ID NO	CRTH2R Antibody	CDRH3
46	CRTH2-74	CARANQHFGPVAGGATPSEEPGSQLTRAELGWDAPPGQESLADELLQLGTEH GYHHYYGMDVW
47	CRTH2-24	CARDMYDYDFTLGPQSIGPLGEVVPADDAFDIW
48	CRTH2-28	CARDMYDYDFATGTGGPEDDLYPQGELNDGYRIEVPADDAFDIW

49	CRTH2-39	CARDMYYDFGVILGGTAVGTNNGSANEVVPADDAFDIW
50	CRTH2-19	CARDMYYDFDYFGLTLTGDRNDDEVVPADDAFDIW
51	CRTH2-9	CARDMYYDFWLGDQSTGSLIGAEVVPADDAFDIW
52	CRTH2-8	CARDMYYDFAAGLEGTTTEVFDEEGHQGGTEVVPADDAFDIW
53	CRTH2-27	CARDMYYDFGSIYGEDVVGELPEVVPADDAFDIW
54	CRTH2-45	CARDGRGSLPRPKGGPTSGGGFSTNIGYGFVVQSYDSSDSSGGAFDIW
55	CRTH2-35	CARANQHFRIFGNYQIYFGHFGYHYYGMDVW
56	CRTH2-50	CARANQHFRVIGQPSAPVPSRGYIYHGYHYYGMDVW
57	CRTH2-66	CARDLRELECEEWTIEVHGQEFVHQRGGVFSRGPVDPGRGVAGGSFVW
58	CRTH2-57	CARANQHFKIQGAPVSTPVPFGTTGYHYYGMDVW
59	CRTH2-32	CARDMYYDFHYSTVGATYYYYLGSETEVVPADDAFDIW
60	CRTH2-15	CARANQHFFLYEGTSSSWLHVGHARYGYHYYGMDVW
61	CRTH2-25	CARDMYYDFEDVDEGSLYLDMGRTFEVVPADDAFDIW
62	CRTH2-42	CARDLRELECEEWTVLQYQKFMHMRWAESGEGSLSRGPVDPGRGVAGSFVW
63	CRTH2-55	CAKHMSMQASTEGDFGLEEVTGEGVDDRADLVGDAFDVW
64	CRTH2-60	CARANQHFSAVRGLAFGYGYRIGGYHYYGMDVW
65	CRTH2-70	CARDMYYDFDVISAGVVGAGNPEVVPADDAFDIW
66	CRTH2-74	CARDMYYDFDVISAGVVGAGNPEVVPADDAFDIW

[00210] 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. 18A-18B**.

[00211] Example 8. Antagonist activity using cAMP assay

[00212] 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.

[00213] Effect of antibody on detected signal in relative light units (rlu) was determined (data not shown). 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. 19A**. Antibodies depicted in **FIG. 19B** 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.

[00214] Example 9. Allosteric modulation of PGD₂-induced cAMP signal

[00215] 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.

[00216] Results of the cAMP assays is seen in **FIG. 20**. A right-ward shift the PGD2 dose response curve (and increase in IC₅₀ value) indicates a negative allosteric effect. As shown in **FIG. 20**, 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.

[00217] **Example 10. Agonist activity of PGD2-induced cAMP signal**

[00218] CRTH2R IgG antibodies were assayed for agonist function. Agonist activity was determined by assaying CRTH2R IgG antibodies described in Example 7 in PGD2-induced cAMP signal.

[00219] 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).

[00220] **Example 11. Control experiments showing allosteric modulators**

[00221] 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 9. 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.

[00222] Results are shown in **FIGS. 21A – 21C**. OC000459 causes a strong right-ward shift of the curve and a 459-fold increase in the IC₅₀ value (**FIG. 21A**). Incubation with CRTH2R AB51 caused no change in IC₅₀ value (**FIG. 21B**). 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. 21C**).

[00223] **Example 12. CRTH2R β -arrestin recruitment assay for antagonist modulation**

[00224] 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. 22A-22D**.

[00225] Example 13. CRTH2R β -arrestin recruitment assay for allosteric modulation

[00226] 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.

[00227] 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.

[00228] Example 14. Hyperimmune immunoglobulin library

[00229] A hyperimmune immunoglobulin (IgG) library was created using similar methods as described in Example 4. 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 CDRH3's 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. 24**.

[00230] The heavy chain CDR length distribution of the hyperimmune antibody libraries were assessed by next generation sequencing (NGS). The data of CDR length distribution is shown in **FIGS. 25A-25B**. Generally, selection of soluble protein targets undergo five rounds of selection involving a PBST wash three times in Round 1, a PBST wash five times in Round 2, a PBST wash seven times in Round 3, a PBST wash nine times in Round 4, and a PBST wash twelve times in Round 5. A non-fat milk block was used. *See* **FIG. 26**.

[00231] For human TIGIT (hTIGIT), 1 μ M biotinylated antigen was mixed with 300 μ l Dynabead M-280 at 10 mg/mL to generate a concentration of 100 pmol per 100 μ l. The details of the various rounds of selection are seen in **Table 15**.

Table 15. Protein panning selection

Manual			
Round	Washes	Antigen Amount	Concentration
1	3	100 pmol	1 uM
2	6	20 pmol	200 nM
3	9	10 pmol	100 nM
4	12	5 pmol	50 nM
5	12	5 pmol	50 nM
Kingfisher (KF)			
1	2	100 pmol	1 uM
2	4	20 pmol	200 nM
3	6	10 pmol	100 nM
4	8	5 pmol	50 nM
5	8	5 pmol	50 nM

[00232] After various rounds of selection, hTIGIT IgGs were analyzed. Data is seen in **FIGS. 27A-27F** and **Table 16**. **FIGS. 27A-27D** show ELISA data from Round 3 and Round 4. **FIGS. 27E-27F** show data of CDRH3 length, yield (ug), and K_D (nM) for the hTIGIT IgGs analyzed.

Table 16. Protein panning data

Round	Target	Antigen	Washes	KF Washes	Titer	KF liter
1	hTIGIT	100 pmol	3	-	4.40E+06	-
2	hTIGIT	50 pmol	5	4	4.40E+07	6.80E+06
3	hTIGIT	20 pmol	7	4	6.00E+08	2.80E+09
4	hTIGIT	10 pmol	9	5	5.00E+08	6.00E+08
5	hTIGIT	10 pmol	-	-	-	-

[00233] Seventeen non-identical hTIGIT immunoglobulins were identified with monovalent affinity ranging from 16 nM to over 300 nM. Most of these immunoglobulins expressed well and produced over 20 ug purified protein at 1 ml expression volume. Sequences for hTIGIT immunoglobulins are seen in **Table 17**.

Table 17. TIGIT sequences

CDRH3		
SEQ ID NO:	IgG	Amino Acid Sequence
67	TIGIT-55-01	CARVAGSSGWAFDYW
68	TIGIT-55-02	CATLRLYSSGGGIDYW
69	TIGIT-55-03	CARIVGATTRTYYYYGMDVW
70	TIGIT-55-04	CARVRNRASDIW
71	TIGIT-55-05	CARAPYSSSSWFDYW
72	TIGIT-55-06	CARNSYGPPRSFGMDVW
73	TIGIT-55-07	CARTPYRSGWADYW
74	TIGIT-55-08	CTRSWYYYYGMDVW

75	TIGIT-55-09	CARGYGGYGYW
76	TIGIT-55-10	CAKAGDYDYFDYW
77	TIGIT-55-11	CASVKRWGYFNNWW
78	TIGIT-55-12	CARVRVGAYDAFDIW
79	TIGIT-55-13	CARNSGWFMFPDYW
80	TIGIT-55-14	CARRSGWYIDSW
81	TIGIT-55-15	CARREGDYMGPWFDPW
82	TIGIT-55-16	CASIRERRFDW
83	TIGIT-55-17	CARHSLTPYNFWSGYYSRSFDIW
Variable Heavy Chain		
84	TIGIT-55-01	EVQLLESGGGLVQPGGSLRLSCAASGFTFGSYGMSWVRQAPGKGLEWVSSISG SGSTTTYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARVAGSSGW AFDYWGQGLTVTVSS
85	TIGIT-55-02	EVQLLESGGGLVQPGGSLRLSCAASGLTFSNYAMTWVRQAPGKGLEWVSGISR SGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCATLRLYSSGG GIDYWGQGLTVTVSS
86	TIGIT-55-03	EVQLLESGGGLVQPGGSLRLSCAASGFTFHNYAMTWVRQAPGKGLEWVSAIT GSGTSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARIVGATTR TYYYYGMDVWGQGLTVTVSS
87	TIGIT-55-04	EVQLLESGGGLVQPGGSLRLSCAASGFRFGNYAMSWVRQAPGKGLEWVSAIT GSGGNTFYADSVKGRFTISRDN SKNTLYLQINSLRAEDTAVYYCARVRNRASDI WGQGLTVTVSS
88	TIGIT-55-05	EVQLLESGGGLVQPGGSLRLSCAASGFVFSSYAMNWVRQAPGKGLEWVSTVS GSGGTTYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARAPYSSSS WFDYWGQGLTVTVSS
89	TIGIT-55-06	EVQLLESGGGLVQPGGSLRLSCAASGFTFSRYTMNWVRQAPGKGLEWVSGISG SGGGAYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARN SYGPPRS FGMDVWGQGLTVTVSS
90	TIGIT-55-07	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYGMTWVRQAPGKGLEWVSAISG RGSSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARTPYRSGW ADYWGQGLTVTVSS
91	TIGIT-55-08	EVQLLESGGGLVQPGGSLRLSCAASGFMFSDYAMSWVRQAPGKGLEWVSGIS GSGGYTTYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCTRSWYYYY GMDVWGQGLTVTVSS
92	TIGIT-55-09	EVQLLESGGGLVQPGGSLRLSCAASGF AFRSYAMGWVRQAPGKGLEWVSTIS GGGGNTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARGYGGY GYWGQGLTVTVSS
93	TIGIT-55-10	EVQLLESGGGLVQPGGSLRLSCAASGFTFSK SAMS WVRQAPGKGLEWVSAISG SGGLTTYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKAGDYDY FDYWGQGLTVTVSS
94	TIGIT-55-11	EVQLLESGGGLVQPGGSLRLSCAASGFTFTNYGMSWVRQAPGKGLEWVSSISG SGSTTTYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCASVKRWGY FNWWGQGLTVTVSS
95	TIGIT-55-12	EVQLLESGGGLVQPGGSLRLSCAASGFTLSSYAMAWVRQAPGKGLEWVSTLS GSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARVRVGAY DAFDIWGQGLTVTVSS
96	TIGIT-55-13	EVQLLESGGGLVQPGGSLRLSCAASGFTFSTYGMNWVRQAPGKGLEWVSTISG SGGSTYFADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARN SGWFMPF DYWGQGLTVTVSS
97	TIGIT-55-14	EVQLLESGGGLVQPGGSLRLSCAASGFMFSRYAMSWVRQAPGKGLEWVSSISG SGGTTYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARRSGWYI DSWGQGLTVTVSS
98	TIGIT-55-15	EVQLLESGGGLVQPGGSLRLSCAASGFTFSDYAMGWVRQAPGKGLEWVSTISG SGSRTTYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARREGDYM GNWFDWPWGQGLTVTVSS

99	TIGIT-55-16	EVQLLES GGGGLVQPGGSLRLSCAASGFAFSSYAMGWVRQAPGKGLEWVSAITS SGGGTYTADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCASIRERRFDF WGQGT LVT VSS
100	TIGIT-55-17	EVQLLES GGGGLVQPGGSLRLSCAASGFTFSNHAMAWVRQAPGKGLEWVSGIS SGGGTYTADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARHSLTPY NFWSGYYRSRFDIWGQGT LVT VSS
Variable Light Chain		
101	TIGIT-55-01	DIQMTQSPSSLSASVGDRVITITCRASQAISNYLNWYQQKPGKAPKLLIYAASRL QSGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQESYSTPFTFGGGTKVEIK
102	TIGIT-55-02	DIQMTQSPSSLSASVGDRVITITCRASQYISTYLNWYQQKPGKAPKLLIYAASRL QSGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQNYITPLTFGGGTKVEIK
103	TIGIT-55-03	DIQMTQSPSSLSASVGDRVITITCRASQYISSYLNWYQQKPGKAPKLLIYGAFSL QSGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQSYITPYTFGGGTKVEIK
104	TIGIT-55-04	DIQMTQSPSSLSASVGDRVITITCRASQTIITLYLNWYQQKPGKAPKLLIYAASRL SGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQSYSLPWTFGGGTKVEIK
105	TIGIT-55-05	DIQMTQSPSSLSASVGDRVITITCRASQSVRSYLNWYQQKPGKAPKLLIYTATSL ESGVP SRFSGSGSGTDFTLTISSLQPEDFATYYCQQSYGLPRTFGGGTKVEIK
106	TIGIT-55-06	DIQMTQSPSSLSASVGDRVITITCRASQSISKYLNWYQQKPGKAPKLLIYGASSLR GGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQSYRPLTFGGGTKVEIK
107	TIGIT-55-07	DIQMTQSPSSLSASVGDRVITITCRASQNIKTYLNWYQQKPGKAPKLLIYAASRL HTGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQTYSIPQTFGGGTKVEIK
108	TIGIT-55-08	DIQMTQSPSSLSASVGDRVITITCRAGQSIRSYLNWYQQKPGKAPKLLIYAASRL QSGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQSYITPYTFGGGTKVEIK
109	TIGIT-55-09	DIQMTQSPSSLSASVGDRVITITCRASQSIRRYLNWYQQKPGKAPKLLIYAASRL QIGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQTYSSPYTFGGGTKVEIK
110	TIGIT-55-10	DIQMTQSPSSLSASVGDRVITITCRTSQSIRRYLNWYQQKPGKAPKLLIYRASRL QSGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQSYNLTFTFGGGTKVEIK
111	TIGIT-55-11	DIQMTQSPSSLSASVGDRVITITCRASQNINYYLNWYQQKPGKAPKLLIYGASSL QNGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQSYITPYTFGGGTKVEIK
112	TIGIT-55-12	DIQMTQSPYLSASVGDRVITITCRASQSIRRYLNWYQQKPGKAPKLLIYRASTL QTGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQTYSSPYTFGGGTKVEIK
113	TIGIT-55-13	DIQMTQSPSSLSASVGDRVITITCRTSQSISTYLNWYQQKPGKAPKLLIYATSRLQ SGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQSYTPTFTFGGGTKVEIK
114	TIGIT-55-14	DIQMTQSPSSLSASVGDRVITITCRASQSVSRYYLNWYQQKPGKAPKLLIYGSSNL QSGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQESYSTPFTFGGGTKVEIK
115	TIGIT-55-15	DIQMTQSPSSLSASVGDRVITITCRASQAISRNLNWYQQKPGKAPKLLIYGASNL QTGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQSHSTPVTFTFGGGTKVEIK
116	TIGIT-55-16	DIQMTQSPSSLSASVGDRVITITCRASQRISTYLNWYQQKPGKAPKLLIYGTSSLQ SGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQSYIIPWTFTFGGGTKVEIK
117	TIGIT-55-17	DIQMTQSPSSLSASVGDRVITITCRASQSISSYVNWYQQKPGKAPKLLIYGASRL QDGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQSYITPYTFGGGTKVEIK

[00234] Identification of human CD3 epsilon (hCD3) and cyno CD3 epsilon (cCD3)

immunoglobulins was performed. The details of the various rounds of selection are seen in **Table 18**.

Table 18. Protein panning selection

Manual		
Round	Washes	Concentration
1	3	500 nM hCD3 and 500 nM cCD3
2	6	100 nM hCD3
3	9	50 nM hCD3
4	12	50 nM hCD3
5	12	10 nM hCD3

Kingfisher (KF)		
1	2	500 nM hCD3 and 500 nM cCD3
2	4	100 nM hCD3
3	6	50 nM hCD3
4	8	50 nM hCD3
5	8	10 nM hCD3

[00235] After various rounds of selection, CD3 epsilon (CD3 ϵ) IgGs were analyzed. Data is seen in **FIGS. 28A-28L** and **Tables 19A-19B**. **FIGS. 28A-28F** show ELISA data from Round 4 and Round 5. **FIGS. 28G-28L** show data of cross-reactivity of human CD3 epsilon and cyno CD3 epsilon immunoglobulins.

Table 19A. Protein panning data

Round	Target	Antigen	Washes	KF Washes	Titer	KF Titer
1	hCD3/cCD3	100 pmol	3	-	2.40E+06	-
2	hCD3	50 pmol	5	4	1.20E+08	1.80E+06
3	cCD3	20 pmol	7	4	8.00E+06	2.40E+07
4	hCD3	10 pmol	9	5	1.00E+07	2.10E+06
5	cCD3	10 pmol	12	7	1.50E+07	1.00E+08

Table 19B.

Round	Target	Antigen	Washes	Output Titer
1	hCD3	5ug	4	9.00E+04
2	cCD3	5ug	5	1.40E+05
3	hCD3	2.5ug	6	3.00E+06
4	cCD3	2.5ug	7	4.00E+06
5	hCD3	2.5ug	8	2.20E+07

[00236] Nineteen non-identical hCD3 epsilon and cyno CD3 epsilon immunoglobulins were identified including five that are human/cyno CD3 epsilon cross-reactive immunoglobulins. One of the human/cyno CD3 epsilon cross-reactive antibody, CD3-56-05 binds to human and cyno CD3 epsilon with 67 and 107 nM affinity, respectively. Sequences for hCD3 epsilon and cCD3 epsilon immunoglobulins are seen in **Table 20**.

Table 20. CD3 epsilon sequences

Variable Heavy Chain CDR1 (CDRH1)		
SEQ ID NO:	IgG	Amino Acid Sequence
118	CD3-138-6	GYTFTSNMH

119	CD3-56-5	FTFSSYAMN
120	CD3-155-03	FTFSSYAIN
Variable Heavy Chain CDR2 (CDRH2)		
SEQ ID NO:	IgG	Amino Acid Sequence
121	CD3-138-6	VASISSYYGYTYA
122	CD3-56-5	VSAVSGSGGRITYA
123	CD3-155-03	VSALSGSGGSTYYA
Variable Heavy Chain CDR3 (CDRH3)		
SEQ ID NO:	IgG	Amino Acid Sequence
124	CD3-138-6	GGNYNLTGYYPLAY
125	CD3-56-5	ARERATTLDY
126	CD3-155-03	ARRSAQLGDY
127	CD3-56-5A	CARERATTLDYW
128	CD3-56-11	CARDSLTTRGYYYMDVW
Variable Light Chain CDR1 (CDRL1)		
SEQ ID NO:	IgG	Amino Acid Sequence
129	CD3-138-6	RASQDISTYLN
130	CD3-56-5	RASQTIYSHLN
131	CD3-155-03	RASQSISSFLN
Variable Light Chain CDR2 (CDRL2)		
SEQ ID NO:	IgG	Amino Acid Sequence
132	CD3-138-6	YTDRLQT
133	CD3-56-5	VASRLQS
134	CD3-155-03	AAPSLQS
Variable Light Chain CDR3 (CDRL3)		
SEQ ID NO:	IgG	Amino Acid Sequence
135	CD3-138-6	QQGGALPFT
136	CD3-56-5	QQSFSTSWT
137	CD3-155-03	QQSFRTPT
Variable Heavy Chain		
138	CD3-56-5	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYAMNWVRQAPGKGLEWVSAVSGSGGRITYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARERATTLDYWGQGLVTVSS
139	CD3-56-11	EVQLLESGGGLVQPGGSLRLSCAASGFRFSTYAMNWVRQAPGKGLEWVSGISGSGGSKYHADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARDSLTTRGYYYMDVWGQGLVTVSS
140	CD3-138-6	EVQLLESGGGLVQPGGSLRLSCAASGGYTFTSNMHWVRQAPGKGLEWVASISSYGYTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCGGNYNLTGYYPLAYWGQGLVTVSS
141	CD3-155-9	EVQLLESGGGLVQPGGSLRLSCAASGFTFSNYALNWVRQAPGKGLEWVSAVTGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARVRATTLDYWGQGLVTVSS
Variable Light Chain		
142	CD3-56-5	DIQMTQSPSSLSASVGDRTITCRASQTIYSHLNWYQQKPGKAPKLLIYVASRLQSGVPSRFGSGSGTDFTLTISLQPEDFATYYCQQSFSTSWTFGGGKVEIK
143	CD3-56-11	DIQMTQSQSSLSASVGDRTITCRASQSIRTSNLYQQPGKAPKLLIYAASRLQSGVPSRFGSGSGTDFTLTISLQPEDFATYYCQQSYSTLYSFGGKVEIK
144	CD3-138-6	DIQMTQSPSSLSASVGDRTITCRASQDISTYLNWYQQKPGKAPKLLIYTDRLQTVPSRFGSGSGTDFTLTISLQPEDFATYYCQQGGALPFTFGGKVEIK
145	CD3-155-9	DIQMTQSPSSLSASVGDRTITCRTSQSISTYLNWYQQKPGKAPKLLIYTASRLQSGVPSRFGSGSGTDFTLTISLQPEDFATYYCQQSYSTPWTFGGKVEIK

[00237] 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.

[00238] CRTH2R immunoglobulins were assessed for binding affinity and allosteric modulator function of PGD2-induced cAMP. As seen in **FIGS. 30A-30F**, 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 21**.

Table 21. CRTH2R sequences

Variable Heavy Chain		
SEQ ID NO:	IgG	Amino Acid Sequence
146	CRTH2-48-3	EVQLVESGGGLVQAGGSLRLSCAASGSIFRINAMGWFRQAPGKEREGVAAINNF GTTKYADSVKGRFTISADNAKNTVYLMNSLKPEDTAVYYCAA VRWGPRNDD RYDWGQGTQVTVSS
147	CRTH2-48-21	EVQLVESGGGLVQAGGSLRLSCAASGSFFSINAMGWFRQAPGKEREFVAGITRS GVSTSYADSVKGRFTISADNAKNTVYLMNSLKPEDTAVYYCAAHRIVVGGTS VGDWRWGQGTQVTVSS
148	CRTH2-48-27	EVQLVESGGGLVQAGGSLRLSCAASGSIFHINAMGWFRQAPGKEREGVAAINNF GTTKYADSVKGRFTISANNAKNTVYLMNSLKPEDTAVYYCAA VRWGPRNDD RYDWGQGTQVTVSS
Variable Light Chain		
149	CRTH2-48-3	DIQMTQSPSSLSASVGDRVTITCRASQSISSDLNHWYQQKPGKAPKLLIYFASGLQS GVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQSYSSPLTFGGGTKVEIKR
150	CRTH2-48-21	DIQMTQSPSSLSASVGDRVTITCRISQSISSNYLNHWYQQKPGKAPKLLIYATSSLES GVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQSYSTLLTFGGGTKVEIKR
151	CRTH2-48-27	DIQMTQSPSSLSASVGDRVTITCRASQSISSRYLHWYQQKPGKAPKLLIYGASRLE SGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCRQSYSTPWTFGGGTKVEIKR

[00239] Example 15. Hyperimmune immunoglobulin library for A2A Receptor

[00240] A hyperimmune immunoglobulin (IgG) library was created using similar methods as described in Examples 4 and 14. 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. The CDRH3 sequences were incorporated into the VHH hShuffle library described in Example 4. The final library diversity was determined to be 1.3×10^{10} .

[00241] 73 out of 88 unique clones had a target cell MFI values 2 fold over parental cells. 15 out of 88 unique Clones with target cell MFI values 20 fold over parental cells. Data for adenosine A2A receptor variant A2AR-90-007 is seen in **FIGS. 31A-31B**.

[00242] This Example shows generation of a VHH library for the A2AR with high affinity and K_D values in the sub-nanomolar range.

[00243] 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.