

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

[0001] This application claims the benefit of U.S. Provisional Patent Application No. 63/115,568 filed on November 18, 2020, U.S. Provisional Patent Application No. 63/104,465 filed on October 22, 2020, U.S. Provisional Patent Application No. 63/073,362 filed on September 1, 2020, U.S. Provisional Patent Application No. 63/069,665 filed on August 24, 2020, U.S. Provisional Patent Application No. 63/034,896 filed on June 4, 2020, U.S. Provisional Patent Application No. 63/016,254 filed on April 27, 2020, each of which is incorporated by reference in its entirety.

BACKGROUND

[0002] Coronaviruses like severe acute respiratory coronavirus 2 (SARS-CoV-2) can cause severe respiratory problems. Therapies are needed for treating and preventing viral infection caused by coronaviruses like SARS-CoV-2. 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. Thus, there is a need to develop compositions and methods for the optimization of antibody properties in order to develop effective therapies for treating coronavirus infections.

INCORPORATION BY REFERENCE

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

BRIEF SUMMARY

[0004] Provided herein are nucleic acid libraries comprising: a plurality of sequences that when translated encode for antibodies or antibody fragments that bind to SARS-CoV-2 or ACE2 protein, wherein each of the sequences comprises a predetermined number of variants within a CDR relative to an input sequence that encodes an antibody, and wherein the library comprises at least 50,000 variant sequences. Further provided herein are nucleic acid libraries, wherein the antibodies or antibody fragments bind to a spike glycoprotein, a membrane protein, an envelope protein, a nucleocapsid protein, or combinations thereof of the SARS-CoV-2. Further provided herein are nucleic acid libraries, wherein the antibodies or antibody fragments bind to a spike glycoprotein. Further provided herein are nucleic acid libraries, wherein the antibodies or antibody fragment bind to a receptor binding domain of the spike glycoprotein. Further provided herein are nucleic acid libraries, wherein the library comprises at least 100,000 variant sequences. Further provided herein are nucleic acid libraries, wherein at least some of the sequences encode for an antibody light chain. Further provided herein are nucleic acid libraries, wherein at least some of the sequences encode for an antibody heavy chain. Further provided herein are nucleic acid libraries, wherein

each sequence of the plurality of sequences comprises at least one variant in the CDR of a heavy chain or light chain relative to the input sequence. Further provided herein are nucleic acid libraries, wherein each sequence of the plurality of sequences comprises at least two variants in the CDR of a heavy chain or light chain relative to the input sequence. Further provided herein are nucleic acid libraries, wherein at least one of the variants is present in at least two individuals. Further provided herein are nucleic acid libraries, wherein at least one of the variants is present in at least three individuals. Further provided herein are nucleic acid libraries, wherein at least one sequence when translated encodes for an antibody or antibody fragment having at least 5X higher binding affinity than a binding affinity of the input sequence. Further provided herein are nucleic acid libraries, wherein at least one sequence when translated encodes for an antibody or antibody fragment having at least 25X higher binding affinity than a binding affinity of the input sequence. Further provided herein are nucleic acid libraries, wherein at least one sequence when translated encodes for an antibody or antibody fragment having at least 50X higher binding affinity than a binding affinity of the input sequence. Further provided herein are nucleic acid libraries, wherein each sequence of the plurality of sequences comprises at least one variant in the CDR of a heavy chain or light chain relative to a germline sequence of the input sequence. Further provided herein are nucleic acid libraries, wherein the CDR is a CDR1, CDR2, and CDR3 on a heavy chain. Further provided herein are nucleic acid libraries, wherein the CDR is a CDR1, CDR2, and CDR3 on a light chain. Further provided herein are nucleic acid libraries, wherein the at least one sequence that when translated encodes for an antibody or antibody fragment having at least 70X higher binding affinity than the input sequence. Further provided herein are nucleic acid libraries, wherein the at least one sequence that when translated encodes for an antibody or antibody fragment having a K_D of less than 50 nM. Further provided herein are nucleic acid libraries, wherein the at least one sequence that when translated encodes for an antibody or antibody fragment having a K_D of less than 25 nM. Further provided herein are nucleic acid libraries, wherein the at least one sequence that when translated encodes for an antibody or antibody fragment having a K_D of less than 10 nM. Further provided herein are nucleic acid libraries, wherein the at least one sequence that when translated encodes for an antibody or antibody fragment having a K_D of less than 5 nM. Further provided herein are nucleic acid libraries, wherein the library encodes a CDR sequence of any one of SEQ ID NOs: 1-591.

[0005] Provided herein are nucleic acid libraries comprising: a plurality of sequences that when translated encode for antibodies or antibody fragments that bind to a coronavirus or a receptor of the coronavirus, wherein each of the sequences comprises a predetermined number of variants within a CDR relative to an input sequence that encodes an antibody, and wherein the library comprises at least 50,000 variant sequences. Further provided herein are nucleic acid libraries, wherein the coronavirus is SARS-CoV, MERS-CoV, CoV-229E, HCoV-NL63, HCoV-OC43, or HCoV-HKU1. Further provided herein are nucleic acid libraries, wherein the receptor of the coronavirus is ACE2 or dipeptidyl peptidase 4 (DPP4). Further provided herein are nucleic acid libraries, wherein the library comprises at least 100,000 variant sequences. Further provided herein are nucleic acid libraries, wherein at least some of the sequences encode for an antibody light chain. Further provided herein are nucleic acid libraries, wherein at least some of the

sequences encode for an antibody heavy chain. Further provided herein are nucleic acid libraries, wherein each sequence of the plurality of sequences comprises at least one variant in the CDR of a heavy chain or light chain relative to the input sequence. Further provided herein are nucleic acid libraries, wherein each sequence of the plurality of sequences comprises at least two variants in the CDR of a heavy chain or light chain relative to the input sequence. Further provided herein are nucleic acid libraries, wherein at least one of the variants is present in at least two individuals. Further provided herein are nucleic acid libraries, wherein at least one of the variants is present in at least three individuals. Further provided herein are nucleic acid libraries, wherein at least one sequence when translated encodes for an antibody or antibody fragment having at least 5X higher binding affinity than a binding affinity of the input sequence. Further provided herein are nucleic acid libraries, wherein at least one sequence when translated encodes for an antibody or antibody fragment having at least 25X higher binding affinity than a binding affinity of the input sequence. Further provided herein are nucleic acid libraries, wherein at least one sequence when translated encodes for an antibody or antibody fragment having at least 50X higher binding affinity than a binding affinity of the input sequence. Further provided herein are nucleic acid libraries, wherein each sequence of the plurality of sequences comprises at least one variant in the CDR of a heavy chain or light chain relative to a germline sequence of the input sequence. Further provided herein are nucleic acid libraries, wherein the CDR is a CDR1, CDR2, and CDR3 on a heavy chain. Further provided herein are nucleic acid libraries, wherein the CDR is a CDR1, CDR2, and CDR3 on a light chain. Further provided herein are nucleic acid libraries, wherein the at least one sequence that when translated encodes for an antibody or antibody fragment having at least 70X higher binding affinity than the input sequence. Further provided herein are nucleic acid libraries, wherein the at least one sequence that when translated encodes for an antibody or antibody fragment having a K_D of less than 50 nM. Further provided herein are nucleic acid libraries, wherein the at least one sequence that when translated encodes for an antibody or antibody fragment having a K_D of less than 25 nM. Further provided herein are nucleic acid libraries, wherein the at least one sequence that when translated encodes for an antibody or antibody fragment having a K_D of less than 10 nM. Further provided herein are nucleic acid libraries, wherein the at least one sequence that when translated encodes for an antibody or antibody fragment having a K_D of less than 5 nM.

[0006] Provided herein are antibodies, wherein the antibody comprises a sequence comprising at least 90% sequence identity to any one of SEQ ID NOs: 1-716.

[0007] Provided herein are antibodies, wherein the antibody comprises a sequence comprising at least 90% sequence identity to SEQ ID NOs: 1-716; and wherein the antibody is a monoclonal antibody, a polyclonal antibody, a bi-specific antibody, a multispecific antibody, a grafted antibody, a human antibody, a humanized antibody, a synthetic antibody, a chimeric antibody, a camelized antibody, a single-chain Fvs (scFv), a single chain antibody, a Fab fragment, a F(ab')₂ fragment, a Fd fragment, a Fv fragment, a single-domain antibody, an isolated complementarity determining region (CDR), a diabody, a fragment comprised of only a single monomeric variable domain, disulfide-linked Fvs (sdFv), an intrabody, an anti-idiotypic (anti-Id) antibody, or ab antigen-binding fragments thereof.

[0008] Provided herein are methods of treating a SARS-CoV-2 infection, comprising administering the antibody as described herein. Further provided herein are methods, wherein the antibody is administered prior to exposure to SARS-CoV-2. Further provided herein are methods, wherein the antibody is administered at least about 1 week prior to exposure to SARS-CoV-2. Further provided herein are methods, wherein the antibody is administered at least about 1 month prior to exposure to SARS-CoV-2. Further provided herein are methods, wherein the antibody is administered at least about 5 months prior to exposure to SARS-CoV-2. Further provided herein are methods, wherein the antibody is administered after exposure to SARS-CoV-2. Further provided herein are methods, wherein the antibody is administered at most about 24 hours after exposure to SARS-CoV-2. Further provided herein are methods, wherein the antibody is administered at most about 1 week after exposure to SARS-CoV-2. Further provided herein are methods, wherein the antibody is administered at most about 1 month after exposure to SARS-CoV-2.

[0009] Provided herein are methods of treating an individual with a SARS-CoV-2 infection with the antibody as described herein comprising: (a) obtaining or having obtained a sample from the individual; (b) performing or having performed an expression level assay on the sample to determine expression levels of SARS-CoV-2 antibodies; and (c) if the sample has an expression level of the SARS-CoV-2 antibodies then administering to the individual the antibody as described herein, thereby treating the SARS-CoV-2 infection.

[0010] Provided herein are methods for optimizing an antibody comprising: (a) providing a plurality of polynucleotide sequences encoding for an antibody or antibody fragment, wherein the antibody or antibody fragment is derived from a subject having SARS-CoV-2; (b) generating a nucleic acid library comprising the plurality of sequences that when translated encode for antibodies or antibody fragments that bind SARS-CoV-2 or ACE2 protein, wherein each of the sequences comprises a predetermined number of variants within a CDR relative to an input sequence that encodes an antibody; wherein the library comprises at least 50,000 variant sequences; and (c) synthesizing the at least 50,000 variant sequences. Further provided herein are methods, wherein the antibody library comprises at least 100,000 sequences. Further provided herein are methods, wherein the method further comprises enriching a subset of the variant sequences. Further provided herein are methods, wherein the method further comprises expressing the antibody or antibody fragments corresponding to the variant sequences. Further provided herein are methods, wherein the polynucleotide sequence is a murine, human, or chimeric antibody sequence. Further provided herein are methods, wherein each sequence of the plurality of variant sequences comprises at least one variant in each CDR of a heavy chain or light chain, relative to the input sequence. Further provided herein are methods, wherein each sequence of the plurality of variant sequences comprises at least two variants in each CDR of a heavy chain or light chain relative to the input sequence. Further provided herein are methods, wherein at least one sequence when translated encodes for an antibody or antibody fragment having at least 5X higher binding affinity than a binding affinity of the input sequence. Further provided herein are methods, wherein at least one sequence when translated encodes for an antibody or antibody fragment having at least 25X higher binding affinity than a binding affinity of the input sequence. Further provided herein are methods, wherein at least one sequence when translated encodes for an antibody or antibody

fragment having at least 50X higher binding affinity than a binding affinity of the input sequence. Further provided herein are methods, wherein each sequence comprises at least one variant in each CDR of a heavy chain or light chain relative to a germline sequence of the input sequence. Further provided herein are methods, wherein the nucleic acid library has a theoretical diversity of at least 10^{10} sequences. Further provided herein are methods, wherein the nucleic acid library has a theoretical diversity of at least 10^{12} sequences.

[0011] Provided herein are antibodies or antibody fragments comprising a variable domain, heavy chain region (VH) and a variable domain, light chain region (VL), wherein VH comprises complementarity determining regions CDRH1, CDRH2, and CDRH3, wherein VL comprises complementarity determining regions CDRL1, CDRL2, and CDRL3, and wherein (a) an amino acid sequence of CDRH1 is as set forth in any one of SEQ ID NOs: 1-36 or 217-282; (b) an amino acid sequence of CDRH2 is as set forth in any one of SEQ ID NOs: 37-72 or 283-348; (c) an amino acid sequence of CDRH3 is as set forth in any one of SEQ ID NOs: 73-108 or 349-414; (d) an amino acid sequence of CDRL1 is as set forth in any one of SEQ ID NOs: 109-144 or 415-473; (e) an amino acid sequence of CDRL2 is as set forth in any one of SEQ ID NOs: 145-180 or 415-473; and (f) an amino acid sequence of CDRL3 is as set forth in any one of SEQ ID NOs: 181-216 or 533-591. Further provided herein are antibodies or antibody fragments, wherein (a) an amino acid sequence of CDRH1 is as set forth in SEQ ID NO: 30; (b) an amino acid sequence of CDRH2 is as set forth in SEQ ID NO: 66; (c) an amino acid sequence of CDRH3 is as set forth in SEQ ID NO: 102; (d) an amino acid sequence of CDRL1 is as set forth in SEQ ID NO: 138; (e) an amino acid sequence of CDRL2 is as set forth in SEQ ID NO: 174; and (f) an amino acid sequence of CDRL3 is as set forth in SEQ ID NO: 210. Further provided herein are antibodies or antibody fragments, wherein (a) an amino acid sequence of CDRH1 is as set forth in SEQ ID NO: 35; (b) an amino acid sequence of CDRH2 is as set forth in SEQ ID NO: 71; (c) an amino acid sequence of CDRH3 is as set forth in SEQ ID NO: 107; (d) an amino acid sequence of CDRL1 is as set forth in SEQ ID NO: 143; (e) an amino acid sequence of CDRL2 is as set forth in SEQ ID NO: 179; and (f) an amino acid sequence of CDRL3 is as set forth in SEQ ID NO: 215. Further provided herein are antibodies or antibody fragments, wherein (a) an amino acid sequence of CDRH1 is as set forth in SEQ ID NO: 12; (b) an amino acid sequence of CDRH2 is as set forth in SEQ ID NO: 48; (c) an amino acid sequence of CDRH3 is as set forth in SEQ ID NO: 84; (d) an amino acid sequence of CDRL1 is as set forth in SEQ ID NO: 120; (e) an amino acid sequence of CDRL2 is as set forth in SEQ ID NO: 156; and (f) an amino acid sequence of CDRL3 is as set forth in SEQ ID NO: 192. Further provided herein are antibodies or antibody fragments, wherein (a) an amino acid sequence of CDRH1 is as set forth in SEQ ID NO: 31; (b) an amino acid sequence of CDRH2 is as set forth in SEQ ID NO: 67; (c) an amino acid sequence of CDRH3 is as set forth in SEQ ID NO: 103; (d) an amino acid sequence of CDRL1 is as set forth in SEQ ID NO: 139; (e) an amino acid sequence of CDRL2 is as set forth in SEQ ID NO: 175; and (f) an amino acid sequence of CDRL3 is as set forth in SEQ ID NO: 211. Further provided herein are antibodies or antibody fragments, wherein (a) an amino acid sequence of CDRH1 is as set forth in SEQ ID NO: 240; (b) an amino acid sequence of CDRH2 is as set forth in SEQ ID NO: 306; (c) an amino acid sequence of

CDRH3 is as set forth in SEQ ID NO: 372; (d) an amino acid sequence of CDRL1 is as set forth in SEQ ID NO: 437; (e) an amino acid sequence of CDRL2 is as set forth in SEQ ID NO: 496; and (f) an amino acid sequence of CDRL3 is as set forth in SEQ ID NO: 555. Further provided herein are antibodies or antibody fragments, wherein (a) an amino acid sequence of CDRH1 is as set forth in SEQ ID NO: 244; (b) an amino acid sequence of CDRH2 is as set forth in SEQ ID NO: 310; (c) an amino acid sequence of CDRH3 is as set forth in SEQ ID NO: 376; (d) an amino acid sequence of CDRL1 is as set forth in SEQ ID NO: 437; (e) an amino acid sequence of CDRL2 is as set forth in SEQ ID NO: 496; and (f) an amino acid sequence of CDRL3 is as set forth in SEQ ID NO: 555. Further provided herein are antibodies or antibody fragments, wherein (a) an amino acid sequence of CDRH1 is as set forth in SEQ ID NO: 270; (b) an amino acid sequence of CDRH2 is as set forth in SEQ ID NO: 336; (c) an amino acid sequence of CDRH3 is as set forth in SEQ ID NO: 402; (d) an amino acid sequence of CDRL1 is as set forth in SEQ ID NO: 461; (e) an amino acid sequence of CDRL2 is as set forth in SEQ ID NO: 520; and (f) an amino acid sequence of CDRL3 is as set forth in SEQ ID NO: 579. Further provided herein are antibodies or antibody fragments, wherein the antibody or antibody fragment binds to a spike glycoprotein. Further provided herein are antibodies or antibody fragments, wherein the antibody or antibody fragment binds to a receptor binding domain of the spike glycoprotein. Further provided herein are antibodies or antibody fragments, wherein the antibody or antibody fragment comprises a K_D of less than 50 nM. Further provided herein are antibodies or antibody fragments, wherein the antibody or antibody fragment comprises a K_D of less than 25 nM. Further provided herein are antibodies or antibody fragments, wherein the antibody or antibody fragment comprises a K_D of less than 10 nM. Further provided herein are antibodies or antibody fragments, wherein the antibody or antibody fragment comprises a K_D of less than 5 nM.

[0012] Provided herein are antibodies or antibody fragments comprising a variable domain, heavy chain region (VH) and a variable domain, light chain region (VL), wherein the VH comprises an amino acid sequence at least about 90% identical to a sequence as set forth in any one of SEQ ID NOs: 592-657, and wherein the VL comprises an amino acid sequence at least about 90% identical to a sequence as set forth in any one of SEQ ID NOs: 658-716. Further provided herein are antibodies or antibody fragments, wherein the VH comprises an amino acid sequence at least about 90% identical to a sequence as set forth in SEQ ID NO: 594, and wherein the VL comprises an amino acid sequence at least about 90% identical to a sequence as set forth in SEQ ID NO: 660. Further provided herein are antibodies or antibody fragments, wherein the VH comprises an amino acid sequence at least about 90% identical to a sequence as set forth in SEQ ID NO: 595, and wherein the VL comprises an amino acid sequence at least about 90% identical to a sequence as set forth in SEQ ID NO: 661. Further provided herein are antibodies or antibody fragments, wherein the VH comprises an amino acid sequence at least about 90% identical to a sequence as set forth in SEQ ID NO: 598, and wherein the VL comprises an amino acid sequence at least about 90% identical to a sequence as set forth in SEQ ID NO: 664. Further provided herein are antibodies or antibody fragments, wherein the VH comprises an amino acid sequence at least about 90% identical to a sequence as set forth in SEQ ID NO: 603, and wherein the VL comprises an amino acid sequence at least about 90% identical to a sequence as set

forth in SEQ ID NO: 669. Further provided herein are antibodies or antibody fragments, wherein the VH comprises an amino acid sequence at least about 90% identical to a sequence as set forth in SEQ ID NO: 615, and wherein the VL comprises an amino acid sequence at least about 90% identical to a sequence as set forth in SEQ ID NO: 680. Further provided herein are antibodies or antibody fragments, wherein the VH comprises an amino acid sequence at least about 90% identical to a sequence as set forth in SEQ ID NO: 631, and wherein the VL comprises an amino acid sequence at least about 90% identical to a sequence as set forth in SEQ ID NO: 691. Further provided herein are antibodies or antibody fragments, wherein the VH comprises an amino acid sequence at least about 90% identical to a sequence as set forth in SEQ ID NO: 645, and wherein the VL comprises an amino acid sequence at least about 90% identical to a sequence as set forth in SEQ ID NO: 704. Further provided herein are antibodies, wherein the antibody comprises a sequence comprising at least 90% sequence identity to any one of SEQ ID NOs: 1-716; and wherein the antibody is a monoclonal antibody, a polyclonal antibody, a bi-specific antibody, a multispecific antibody, a grafted antibody, a human antibody, a humanized antibody, a synthetic antibody, a chimeric antibody, a camelized antibody, a single-chain Fvs (scFv), a single chain antibody, a Fab fragment, a F(ab')₂ fragment, a Fd fragment, a Fv fragment, a single-domain antibody, an isolated complementarity determining region (CDR), a diabody, a fragment comprised of only a single monomeric variable domain, disulfide-linked Fvs (sdFv), an intrabody, an anti-idiotypic (anti-Id) antibody, or ab antigen-binding fragments thereof.

[0013] Provided herein are nucleic acid compositions comprising: a) a first nucleic acid encoding a variable domain, heavy chain region (VH) comprising an amino acid sequence at least about 90% identical to a sequence as set forth in any one of SEQ ID NOs: 592-657; b) a second nucleic acid encoding a variable domain, light chain region (VL) comprising at least about 90% identical to a sequence as set forth in any one of SEQ ID NOs: 658-716; and an excipient.

[0014] Provided herein are methods of treating a SARS-CoV-2 infection, comprising administering the antibody or antibody fragment described herein. Further provided herein are methods, wherein the antibody is administered prior to exposure to SARS-CoV-2. Further provided herein are methods, wherein the antibody is administered at least about 1 week prior to exposure to SARS-CoV-2. Further provided herein are methods, wherein the antibody is administered at least about 1 month prior to exposure to SARS-CoV-2. Further provided herein are methods, wherein the antibody is administered at least about 5 months prior to exposure to SARS-CoV-2. Further provided herein are methods, wherein the antibody is administered after exposure to SARS-CoV-2. Further provided herein are methods, wherein the antibody is administered at most about 24 hours after exposure to SARS-CoV-2. Further provided herein are methods, wherein the antibody is administered at most about 1 week after exposure to SARS-CoV-2. Further provided herein are methods, wherein the antibody is administered at most about 1 month after exposure to SARS-CoV-2.

[0015] Provided herein are methods of treating an individual with a SARS-CoV-2 infection with the antibody or antibody fragment described herein comprising: a) obtaining or having obtained a sample from the individual; b) performing or having performed an expression level assay on the sample to determine expression levels of SARS-CoV-2 antibodies; and if the sample has an expression level of the SARS-CoV-2

antibodies then administering to the individual the antibody or antibody fragment described herein, thereby treating the SARS-CoV-2 infection.

[0016] Provided herein are methods for optimizing an antibody comprising: a) providing a plurality of polynucleotide sequences encoding for an antibody or antibody fragment, wherein the antibody or antibody fragment is derived from a subject having SARS-CoV-2; b) generating a nucleic acid library comprising the plurality of sequences that when translated encode for antibodies or antibody fragments that bind SARS-CoV-2 or ACE2 protein, wherein each of the sequences comprises a predetermined number of variants within a CDR relative to an input sequence that encodes an antibody; wherein the library comprises at least 50,000 variant sequences; and c) synthesizing the at least 50,000 variant sequences. Further provided herein are methods, wherein the antibody library comprises at least 100,000 sequences. Further provided herein are methods, wherein the method further comprises enriching a subset of the variant sequences. Further provided herein are methods, wherein the method further comprises expressing the antibody or antibody fragments corresponding to the variant sequences. Further provided herein are methods, wherein the polynucleotide sequence is a murine, human, or chimeric antibody sequence. Further provided herein are methods, wherein each sequence of the plurality of variant sequences comprises at least one variant in each CDR of a heavy chain or light chain, relative to the input sequence. Further provided herein are methods, wherein each sequence of the plurality of variant sequences comprises at least two variants in each CDR of a heavy chain or light chain relative to the input sequence. Further provided herein are methods, wherein at least one sequence when translated encodes for an antibody or antibody fragment having at least 5X higher binding affinity than a binding affinity of the input sequence. Further provided herein are methods, wherein at least one sequence when translated encodes for an antibody or antibody fragment having at least 25X higher binding affinity than a binding affinity of the input sequence. Further provided herein are methods, wherein at least one sequence when translated encodes for an antibody or antibody fragment having at least 50X higher binding affinity than a binding affinity of the input sequence. Further provided herein are methods, wherein each sequence comprises at least one variant in each CDR of a heavy chain or light chain relative to a germline sequence of the input sequence. Further provided herein are methods, wherein the nucleic acid library has a theoretical diversity of at least 10^{10} sequences. Further provided herein are methods, wherein the nucleic acid library has a theoretical diversity of at least 10^{12} sequences.

BRIEF DESCRIPTION OF THE DRAWINGS

[0017] **Figure 1** depicts a workflow for antibody optimization.

[0018] **Figure 2** presents a diagram of steps demonstrating an exemplary process workflow for gene synthesis as disclosed herein.

[0019] **Figure 3** illustrates an example of a computer system.

[0020] **Figure 4** is a block diagram illustrating an architecture of a computer system.

- [0021] **Figure 5** 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).
- [0022] **Figure 6** is a block diagram of a multiprocessor computer system using a shared virtual address memory space.
- [0023] **Figure 7** is a schema of a panning workflow.
- [0024] **Figures 8A-8B** are graphs of panning data from round 4 for antibody 1.
- [0025] **Figures 8C-8D** are graphs of panning data from round 4 for antibody 2.
- [0026] **Figures 9A-9B** are graphs of panning data from round 4 for antibody 3.
- [0027] **Figures 9C-9D** are graphs of panning data from round 4 for antibody 4.
- [0028] **Figure 10** shows graphs of ACE2 binding to SARS-CoV-2 variant antibodies.
- [0029] **Figures 11A-11B** are graphs of affinity data for SARS-CoV-2 variant antibodies.
- [0030] **Figure 12** is a graph of binding of SARS-CoV-2 variant antibodies to S1 protein.
- [0031] **Figure 13A** is a graph of SARS-CoV-2 and ACE2 competition ELISAs from a first set.
- [0032] **Figure 13B** is a graph of SARS-CoV-2 and ACE2 competition ELISAs from a first set showing SARS-CoV-2 variant antibodies.
- [0033] **Figures 14A-14B** are graphs from a first set (**FIG. 14A**) and second set (**FIG. 14B**) showing ACE2 variant antibodies.
- [0034] **Figures 14C-14D** are graphs of SARS-CoV-2 variant antibodies in neutralization assays.
- [0035] **Figure 15** is a graph of SARS-CoV-2 and ACE2 inhibition.
- [0036] **Figures 16A-16B** are graphs of SARS-CoV-2 variant antibodies on VERO E6 inhibition measured by FACS.
- [0037] **Figure 17A** is a graph of SARS-CoV-2 variant antibodies on VERO E6 inhibition measured by FACS as compared to CR3022.
- [0038] **Figures 17B-17C** are graphs of affinity of SARS-CoV-2 variant antibodies determined by coating ELISA plates with SARS-CoV-2 Spike Glycoprotein S1 (**FIG. 17B**) or S protein trimer (**Fig. 17C**).
- [0039] **Figure 17D** is a graph of mean fluorescent intensity (MFI) plotted for each SARS-CoV-2 variant antibody dilution.
- [0040] **Figures 18A-18C** are graphs of mean fluorescent intensity (MFI) plotted for each SARS-CoV-2 variant antibody dilution.
- [0041] **Figures 19A-19B** are graphs of antibody kinetics for variants 2-5, 2-2, and 2-6 (**FIG. 19A**) and variants 1-12, 1-42, 1-20, and 1-19 (**FIG. 19B**).
- [0042] **Figure 19C** is a graph of percent neutralization for variants 1-12, 1-42 and 1-20.
- [0043] **Figure 19D** is a graph of percent neutralization for variants 1-12, 1-42 and 1-20 using live virus.
- [0044] **Figures 20A-20D** are graphs of variant antibodies neutralizing live virus.
- [0045] **Figure 20E** is a graph of variant antibodies neutralizing live virus FRNT.
- [0046] **Figure 20F-20I** show data of variant antibodies neutralizing live virus PRNT.

- [0047] **Figure 21A** shows a graph of percent weight change (y-axis) versus day post injection (PI, x-axis) for positive control convalescent plasma and negative control Mab c7d11.
- [0048] **Figure 21B** shows a graph of percent weight change (y-axis) versus day post injection (PI, x-axis) for variant antibody 6-63.
- [0049] **Figure 21C** shows a graph of percent weight change (y-axis) versus day post injection (PI, x-axis) for variant antibody 6-3.
- [0050] **Figure 21D** shows a graph of percent weight change (y-axis) versus day post injection (PI, x-axis) for variant antibody 6-36.
- [0051] **Figure 21E** shows graphs of percent weight change (y-axis) versus day post injection (PI, x-axis) based on dose.
- [0052] **Figure 21F** shows graphs of percent weight change (y-axis) versus day post injection (PI, x-axis) based on dose for variant antibodies 2-3, 2-63, and 1-20.
- [0053] **Figure 21G** shows a graph of data from a plaque assay to detect infectious virus in Day 9 lungs. The indicated antibodies were administered Day -1. Lungs were collected on Day 9, the right lobe was homogenized, clarified and supernatants were quantified by plaque titration. Individual hamster values are shown as symbols. White symbols indicate no infectious virus detected. The geometric mean PFU/gram is shown as bars. Limit of assay shown as dotted line.
- [0054] **Figure 21H** shows a graph of data from *in situ* hybridization (ISH) to detect infected cells in Day 9 lungs. The indicated antibodies were administered Day -1. Three animals per group were analyzed. Individual hamster values are shown as symbols. Median ISH scores are shown as bars.
- [0055] **Figure 21I** shows a graph of data from cumulative inflammation and edema scores for Day 9 lungs. The indicated antibodies were administered Day -1. Three animals per group were analyzed. Individual hamster cumulative pathology scores are shown as symbols. Median scores are shown as bars.
- [0056] **Figure 22A** shows a graph of the positive control pAb in a neutralization assay.
- [0057] **Figure 22B** shows a graph of neutralization of antibodies 6-63, 6-3, and 1-12 in VSV-SARS B.135 strain.
- [0058] **Figure 23A** shows a graph of the positive control in a neutralization assay.
- [0059] **Figures 23B-23C** show graphs of neutralization by antibodies described herein.
- [0060] **Figure 24A** shows graphs weight change. Animals were immunosuppressed and then exposed to SARS-CoV-2 virus, WA1 strain, on Day 0. Top graph indicates data from the control group that received the cocktail on Day -1 (D-1, Group A). A group was immunosuppressed but not exposed to virus (CYP Control, Group I). Negative control is an IgG monoclonal (Group H). The cocktail was administered on the indicated day post-exposure (Groups B-G; middle and bottom graphs). Arrows indicate day of antibody administration. Symbols are mean $\pm$ SEM. Statistical differences in the area under the curve (AOC) are shown to the right of each line. * indicates p value <0.05 , ns= not significant.
- [0061] **Figure 24B** shows a graph of data from **FIG. 24A** plotted on one graph.

[0062] **Figure 24C** shows a graph of infectious virus in lungs on Day 14. Plaque assays were run on Day 14 lung homogenates. Plaque forming units (PFU) per gram of tissue were calculated and plotted. The limit of the assay is shown as a dotted line. Bars are the geometric means for each group. Group ID are in parentheses. White symbols indicated no infectious virus was detected. CYP= cyclophosphamide; Neg= Negative; Cont= control.

[0063] **Figure 24D** shows a graph infectious virus in lungs of untreated control hamsters. CYP-treated animals were exposed to 1,000 pfu of virus by intranasal route on Day 0. Groups of four animals were euthanized and lungs were collected on the indicated days. Lung homogenates were assayed for infectious virus by plaque assay. Plaque forming units (PFU) per gram lung tissue are plotted. Geometric mean titers and SD are shown.

DETAILED DESCRIPTION

[0064] 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.

[0065] Definitions

[0066] 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.

[0067] 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.

[0068] 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.

[0069] 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. cDNA described herein may be generated by de novo synthesis.

[0070] Antibody Optimization Library for Coronavirus

[0071] Provided herein are methods, compositions, and systems for the optimization of antibodies for coronavirus. In some embodiments, the antibodies are optimized for SARS-CoV, MERS-CoV, CoV-229E, HCoV-NL63, HCoV-OC43, or HCoV-HKU1. In some embodiments, the antibodies are optimized for SARS-CoV-2. In some embodiments, the antibodies are optimized for a receptor that binds to the coronavirus. In some embodiments, the receptor of the coronavirus is ACE2 or dipeptidyl peptidase 4 (DPP4). In some embodiments, the antibodies are optimized based on interactions between the coronavirus and the receptor that binds the coronavirus. In some embodiments, the antibodies are optimized for

angiotensin-converting enzyme 2 (ACE2). In some embodiments, the antibodies are optimized based on interactions between SARS-CoV-2 and ACE2.

[0072] Antibodies are in some instances optimized by the design of in-silico libraries comprising variant sequences of an input antibody sequence (**FIG. 1**). Input sequences **100** are in some instances modified in-silico **102** with one or more mutations or variants to generate libraries of optimized sequences **103**. 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. Selection pressures used during enrichment in some instances includes, but is not limited to, binding affinity, toxicity, immunological tolerance, stability, receptor-ligand competition, or developability. 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. Sequencing at one or more rounds is in some instances used to identify which sequences **105** have been enriched in the library.

[0073] Described herein are methods and systems of in-silico library design. For example, an antibody or antibody fragment sequence is used as input. In some instances, the antibody sequence used as input is an antibody or antibody fragment sequence that binds SARS-CoV-2. In some instances, the input is an antibody or antibody fragment sequence that binds a protein of SARS-CoV-2. In some instances, the protein is a spike glycoprotein, a membrane protein, an envelope protein, a nucleocapsid protein, or combinations thereof. In some instances, the protein is a spike glycoprotein of SARS-CoV-2. In some instances, the protein is a receptor binding domain of SARS-CoV-2. In some instances, the input sequence is an antibody or antibody fragment sequence that binds angiotensin-converting enzyme 2 (ACE2). In some instances, the input sequence is an antibody or antibody fragment sequence that binds an extracellular domain of the angiotensin-converting enzyme 2 (ACE2).

[0074] A database **102** comprising known mutations or variants of one or more viruses is queried **101**, and a library **103** of sequences comprising combinations of these mutations or variants are generated. In some instances, the database comprises known mutations or variants of SARS-CoV-like coronaviruses, SARS-CoV-2, SARS-CoV, or combinations thereof. In some instances, the database comprises known mutations or variants of the spike protein of SARS-CoV-like coronaviruses, SARS-CoV-2, SARS-CoV, or combinations thereof. In some instances, the database comprises known mutations or variants of the receptor binding domain of SARS-CoV-like coronaviruses, SARS-CoV-2, SARS-CoV, or combinations thereof. In some instances, the database comprises mutations or variants of a protein of SARS-CoV-like coronaviruses, SARS-CoV-2, SARS-CoV, or combinations thereof that binds to ACE2.

[0075] In some instances, the input sequence is a heavy chain sequence of an antibody or antibody fragment that binds SARS-CoV-like coronaviruses, SARS-CoV-2, SARS-CoV, or combinations thereof. In some instances, the input sequence is a light chain sequence of an antibody or antibody fragment that binds

SARS-CoV-like coronaviruses, SARS-CoV-2, SARS-CoV, or combinations thereof. In some instances, the heavy chain sequence comprises varied CDR regions. In some instances, the light chain sequence comprises varied CDR regions. In some instances, known mutations or variants from CDRs are used to build the sequence library. Filters **104**, or exclusion criteria, are in some instances used to select specific types of variants for members of the sequence library. For example, sequences having a mutation or variant are added if a minimum number of organisms in the database have the mutation or variant. In some instances, additional CDRs are specified for inclusion in the database. In some instances, specific mutations or variants or combinations of mutations or variants are excluded from the library (e.g., known immunogenic sites, structure sites, etc.). In some instances, specific sites in the input sequence are systematically replaced with histidine, aspartic acid, glutamic acid, or combinations thereof. In some instances, the maximum or minimum number of mutations or variants allowed for each region of an antibody are specified. Mutations or variants in some instances are described relative to the input sequence or the input sequence's corresponding germline sequence. For example, sequences generated by the optimization comprise at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or more than 16 mutations or variants from the input sequence. In some instances, sequences generated by the optimization 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 or variants from the input sequence. In some instances, sequences generated by the optimization comprise about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or about 18 mutations or variants relative to the input sequence. In some instances, sequences generated by the optimization comprise about 1, 2, 3, 4, 5, 6, or 7 mutations or variants from the input sequence in a first CDR region. In some instances, sequences generated by the optimization comprise about 1, 2, 3, 4, 5, 6, or 7 mutations or variants from the input sequence in a second CDR region. In some instances, sequences generated by the optimization comprise about 1, 2, 3, 4, 5, 6, or 7 mutations or variants from the input sequence in a third CDR region. In some instances, sequences generated by the optimization comprise about 1, 2, 3, 4, 5, 6, or 7 mutations or variants from the input sequence in a first CDR region of a heavy chain. In some instances, sequences generated by the optimization comprise about 1, 2, 3, 4, 5, 6, or 7 mutations or variants from the input sequence in a second CDR region of a heavy chain. In some instances, sequences generated by the optimization comprise about 1, 2, 3, 4, 5, 6, or 7 mutations or variants from the input sequence in a third CDR region of a heavy chain. In some instances, sequences generated by the optimization comprise about 1, 2, 3, 4, 5, 6, or 7 mutations or variants from the input sequence in a first CDR region of a light chain. In some instances, sequences generated by the optimization comprise about 1, 2, 3, 4, 5, 6, or 7 mutations or variants from the input sequence in a second CDR region of a light chain. In some instances, sequences generated by the optimization comprise about 1, 2, 3, 4, 5, 6, or 7 mutations or variants from the input sequence in a third CDR region of a light chain. In some instances, a first CDR region is CDR1. In some instances, a second CDR region is CDR2. In some instances, a third CDR region is CDR3. In-silico antibodies libraries are in some instances synthesized, assembled, and enriched for desired sequences.

[0076] The germline sequences corresponding to an input sequence may also be modified to generate sequences in a library. For example, sequences generated by the optimization 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 or variants from the germline sequence. In some instances, sequences generated by the optimization 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 or variants from the germline sequence. In some instances, sequences generated by the optimization comprise about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, or about 18 mutations or variants relative to the germline sequence.

[0077] Provided herein are methods, systems, and compositions for antibody optimization, wherein the input sequence comprises mutations or variants in an antibody region. Exemplary regions of the antibody 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 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). In some instances, sequences generated by the optimization comprise about 1, 2, 3, 4, 5, 6, or 7 mutations or variants from the germline sequence in a first CDR region. In some instances, sequences generated by the optimization comprise about 1, 2, 3, 4, 5, 6, or 7 mutations or variants from the germline sequence in a second CDR region. In some instances, sequences generated by the optimization comprise about 1, 2, 3, 4, 5, 6, or 7 mutations or variants from the germline sequence in a third CDR region. In some instances, sequences generated by the optimization comprise about 1, 2, 3, 4, 5, 6, or 7 mutations or variants from the germline sequence in a first CDR region of a heavy chain. In some instances, sequences generated by the optimization comprise about 1, 2, 3, 4, 5, 6, or 7 mutations or variants from the germline sequence in a second CDR region of a heavy chain. In some instances, sequences generated by the optimization comprise about 1, 2, 3, 4, 5, 6, or 7 mutations or variants from the germline sequence in a third CDR region of a heavy chain. In some instances, sequences generated by the optimization comprise about 1, 2, 3, 4, 5, 6, or 7 mutations or variants from the germline sequence in a first CDR region of a light chain. In some instances, sequences generated by the optimization comprise about 1, 2, 3, 4, 5, 6, or 7 mutations or variants from the germline sequence in a second CDR region of a light chain. In some instances, sequences generated by the optimization comprise about 1, 2, 3, 4, 5, 6, or 7 mutations or variants from the germline sequence in a third CDR region of a light chain. In some instances, a first CDR region is CDR1. In some instances, a second CDR region is CDR2. In some instances, a third CDR region is CDR3.

[0078] *VHH Libraries*

[0079] Provided herein are methods, compositions, and systems for generation of antibodies or antibody fragments. In some instances, the antibodies or antibody fragments 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, but is not limited to, binding affinity, toxicity, immunological tolerance, stability, receptor-ligand competition, or developability. 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.

[0080] 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.

[0081] 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.

[0082] 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.

[0083] 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 or variants 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 or variants 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 or variants relative to the germline sequence.

[0084] *Coronavirus Antibody Libraries*

[0085] Provided herein are libraries generated from antibody optimization methods described herein. Antibodies described herein result in improved functional activity, structural stability, expression, specificity, or a combination thereof.

[0086] Provided herein are methods and compositions relating to SARS-CoV-2 binding libraries comprising nucleic acids encoding for a SARS-CoV-2 antibody. Further provided herein are methods and compositions relating to ACE2 binding libraries comprising nucleic acids encoding for an ACE2 antibody. Such methods and compositions in some instances are generated by the antibody optimization methods and systems described herein. Libraries as described herein may be further variegated to provide for variant libraries comprising nucleic acids each encoding for a predetermined variant of at least one predetermined reference nucleic acid sequence. Further described herein are protein libraries that may be generated when the nucleic acid libraries are translated. In some instances, nucleic acid libraries as described herein are transferred into cells to generate a cell library. Also provided herein are downstream applications for the libraries synthesized using methods described herein. Downstream applications include identification of variant nucleic acids or protein sequences with enhanced biologically relevant functions, e.g., improved stability, affinity, binding, functional activity, and for the treatment or prevention of an infection caused by a coronavirus such as SARS-CoV-2.

[0087] In some instances, an antibody or antibody fragment described herein comprises a sequence of any one of SEQ ID NOs: 1-716. In some instances, an antibody or antibody fragment described herein comprises a sequence that is at least 80% identical to a sequence of any one of SEQ ID NOs: 1-716. In some instances, an antibody or antibody fragment described herein comprises a sequence that is at least 85% identical to a sequence of any one of SEQ ID NOs: 1-716. In some instances, an antibody or antibody fragment described herein comprises a sequence that is at least 90% identical to a sequence of any one of

SEQ ID NOs: 1-716. In some instances, an antibody or antibody fragment described herein comprises a sequence that is at least 95% identical to a sequence of any one of SEQ ID NOs: 1-716.

[0088] In some instances, an antibody or antibody fragment described herein comprises a CDRH1 sequence of any one of SEQ ID NOs: 1-36 or 217-282. In some instances, an antibody or antibody fragment described herein comprises a sequence that is at least 80% identical to a CDRH1 sequence of any one of SEQ ID NOs: 1-36 or 217-282. In some instances, an antibody or antibody fragment described herein comprises a sequence that is at least 85% identical to a CDRH1 sequence of any one of SEQ ID NOs: 1-36 or 217-282. In some instances, an antibody or antibody fragment described herein comprises a sequence that is at least 90% identical to a CDRH1 sequence of any one of SEQ ID NOs: 1-36 or 217-282. In some instances, an antibody or antibody fragment described herein comprises a sequence that is at least 95% identical to a CDRH1 sequence of any one of SEQ ID NOs: 1-36 or 217-282. In some instances, an antibody or antibody fragment described herein comprises a CDRH2 sequence of any one of SEQ ID NOs: 37-72 or 283-348. In some instances, an antibody or antibody fragment described herein comprises a sequence that is at least 80% identical to a CDRH2 sequence of any one of SEQ ID NOs: 37-72 or 283-348. In some instances, an antibody or antibody fragment described herein comprises a sequence that is at least 85% identical to a CDRH2 sequence of any one of SEQ ID NOs: 37-72 or 283-348. In some instances, an antibody or antibody fragment described herein comprises a sequence that is at least 90% identical to a CDRH2 sequence of any one of SEQ ID NOs: 37-72 or 283-348. In some instances, an antibody or antibody fragment described herein comprises a sequence that is at least 95% identical to a CDRH2 sequence of any one of SEQ ID NOs: 37-72 or 283-348. In some instances, an antibody or antibody fragment described herein comprises a CDRH3 sequence of any one of SEQ ID NOs: 73-108 or 349-414. In some instances, an antibody or antibody fragment described herein comprises a sequence that is at least 80% identical to a CDRH3 sequence of any one of SEQ ID NOs: 73-108 or 349-414. In some instances, an antibody or antibody fragment described herein comprises a sequence that is at least 85% identical to a CDRH3 sequence of any one of SEQ ID NOs: 73-108 or 349-414. In some instances, an antibody or antibody fragment described herein comprises a sequence that is at least 90% identical to a CDRH3 sequence of any one of SEQ ID NOs: 73-108 or 349-414. In some instances, an antibody or antibody fragment described herein comprises a sequence that is at least 95% identical to a CDRH3 sequence of any one of SEQ ID NOs: 73-108 or 349-414.

[0089] In some instances, an antibody or antibody fragment described herein comprises a CDRL1 sequence of any one of SEQ ID NOs: 109-144 or 415-473. In some instances, an antibody or antibody fragment described herein comprises a sequence that is at least 80% identical to a CDRL1 sequence of any one of SEQ ID NOs: 109-144 or 415-473. In some instances, an antibody or antibody fragment described herein comprises a sequence that is at least 85% identical to a CDRL1 sequence of any one of SEQ ID NOs: 109-144 or 415-473. In some instances, an antibody or antibody fragment described herein comprises a sequence that is at least 90% identical to a CDRL1 sequence of any one of SEQ ID NOs: 109-144 or 415-473. In some instances, an antibody or antibody fragment described herein comprises a sequence that is at least 95% identical to a CDRL1 sequence of any one of SEQ ID NOs: 109-144 or 415-473.

least 95% identical to a CDRL1 sequence of any one of SEQ ID NOs: 109-144 or 415-473. In some instances, an antibody or antibody fragment described herein comprises a CDRL2 sequence of any one of SEQ ID NOs: 145-180 or 415-473. In some instances, an antibody or antibody fragment described herein comprises a sequence that is at least 80% identical to a CDRL2 sequence of any one of SEQ ID NOs: 145-180 or 415-473. In some instances, an antibody or antibody fragment described herein comprises a sequence that is at least 85% identical to a CDRL2 sequence of any one of SEQ ID NOs: 145-180 or 415-473. In some instances, an antibody or antibody fragment described herein comprises a sequence that is at least 90% identical to a CDRL2 sequence of any one of SEQ ID NOs: 145-180 or 415-473. In some instances, an antibody or antibody fragment described herein comprises a sequence that is at least 95% identical to a CDRL2 sequence of any one of SEQ ID NOs: 145-180 or 415-473. In some instances, an antibody or antibody fragment described herein comprises a CDRL3 sequence of any one of SEQ ID NOs: 181-216, 442-456, or 532-546. In some instances, an antibody or antibody fragment described herein comprises a sequence that is at least 80% identical to a CDRL3 sequence of any one of SEQ ID NOs: 181-216 or 533-591. In some instances, an antibody or antibody fragment described herein comprises a sequence that is at least 85% identical to a CDRL3 sequence of any one of SEQ ID NOs: 181-216 or 533-591. In some instances, an antibody or antibody fragment described herein comprises a sequence that is at least 90% identical to a CDRL3 sequence of any one of SEQ ID NOs: 181-216 or 533-591. In some instances, an antibody or antibody fragment described herein comprises a sequence that is at least 95% identical to a CDRL3 sequence of any one of SEQ ID NOs: 181-216 or 533-591.

[0090] In some embodiments, the antibody or antibody fragment comprising a variable domain, heavy chain region (VH) and a variable domain, light chain region (VL), wherein VH comprises complementarity determining regions CDRH1, CDRH2, and CDRH3, wherein VL comprises complementarity determining regions CDRL1, CDRL2, and CDRL3, and wherein (a) an amino acid sequence of CDRH1 is as set forth in any one of SEQ ID NOs: 1-36 or 217-282; (b) an amino acid sequence of CDRH2 is as set forth in any one of SEQ ID NOs: 37-72 or 283-348; (c) an amino acid sequence of CDRH3 is as set forth in any one of SEQ ID NOs: 73-108 or 349-414; (d) an amino acid sequence of CDRL1 is as set forth in any one of SEQ ID NOs: 109-144 or 415-473; (e) an amino acid sequence of CDRL2 is as set forth in any one of SEQ ID NOs: 145-180 or 415-473; and (f) an amino acid sequence of CDRL3 is as set forth in any one of SEQ ID NOs: 181-216 or 533-591. In some embodiments, the antibody or antibody fragment comprising a variable domain, heavy chain region (VH) and a variable domain, light chain region (VL), wherein VH comprises complementarity determining regions CDRH1, CDRH2, and CDRH3, wherein VL comprises complementarity determining regions CDRL1, CDRL2, and CDRL3, and wherein (a) an amino acid sequence of CDRH1 is at least or about 80%, 85%, 90%, or 95% identical to any one of SEQ ID NOs: 1-36 or 217-282; (b) an amino acid sequence of CDRH2 is at least or about 80%, 85%, 90%, or 95% identical to any one of SEQ ID NOs: 37-72 or 283-348; (c) an amino acid sequence of CDRH3 is at least or about 80%, 85%, 90%, or 95% identical to any one of SEQ ID NOs: 73-108 or 349-414; (d) an amino acid sequence of CDRL1 is at least or about 80%, 85%, 90%, or 95% identical to any one of SEQ ID NOs: 109-144 or 415-

473; (e) an amino acid sequence of CDRL2 is at least or about 80%, 85%, 90%, or 95% identical to any one of SEQ ID NOs: 145-180 or 415-473; and (f) an amino acid sequence of CDRL3 is at least or about 80%, 85%, 90%, or 95% identical to any one of SEQ ID NOs: 181-216 or 533-591.

[0091] Described herein, in some embodiments, are antibodies or antibody fragments comprising a variable domain, heavy chain region (VH) and a variable domain, light chain region (VL), wherein the VH comprises an amino acid sequence at least about 90% identical to a sequence as set forth in any one of SEQ ID NOs: 592-657, and wherein the VL comprises an amino acid sequence at least about 90% identical to a sequence as set forth in any one of SEQ ID NOs: 658-716. In some instances, the antibodies or antibody fragments comprise VH 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: 592-657, and VL 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: 658-716.

[0092] The term “sequence identity” means that two polynucleotide sequences are identical (i.e., on a nucleotide-by-nucleotide basis) over the window of comparison. The term “percentage of sequence identity” is calculated by comparing two optimally aligned sequences over the window of comparison, determining the number of positions at which the identical nucleic acid base (e.g., A, T, C, G, U, or I) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison (i.e., the window size), and multiplying the result by 100 to yield the percentage of sequence identity.

[0093] The term “homology” or “similarity” between two proteins is determined by comparing the amino acid sequence and its conserved amino acid substitutes of one protein sequence to the second protein sequence. Similarity may be determined by procedures which are well-known in the art, for example, a BLAST program (Basic Local Alignment Search Tool at the National Center for Biological Information).

[0094] Provided herein are libraries comprising nucleic acids encoding for SARS-CoV-2 antibodies. Antibodies described herein allow for improved stability for a range of SARS-CoV-2 or ACE2 binding domain encoding sequences. In some instances, the binding domain encoding sequences are determined by interactions between SARS-CoV-2 and ACE2.

[0095] Sequences of binding domains based on surface interactions between SARS-CoV-2 and ACE2 are analyzed using various methods. For example, multispecies computational analysis is performed. In some instances, a structure analysis is performed. In some instances, a sequence analysis is performed. Sequence analysis can be performed using a database known in the art. Non-limiting examples of databases include, but are not limited to, NCBI BLAST (blast.ncbi.nlm.nih.gov/Blast.cgi), UCSC Genome Browser (genome.ucsc.edu/), UniProt (www.uniprot.org/), and IUPHAR/BPS Guide to PHARMACOLOGY (guidetopharmacology.org/).

[0096] Described herein are SARS-CoV-2 or ACE2 binding domains designed based on sequence analysis among various organisms. For example, sequence analysis is performed to identify homologous sequences in different organisms. Exemplary organisms include, but are not limited to, mouse, rat, equine,

sheep, cow, primate (*e.g.*, chimpanzee, baboon, gorilla, orangutan, monkey), dog, cat, pig, donkey, rabbit, fish, fly, and human. In some instances, homologous sequences are identified in the same organism, across individuals.

[0097] Following identification of SARS-CoV-2 or ACE2 binding domains, libraries comprising nucleic acids encoding for the SARS-CoV-2 or ACE2 binding domains may be generated. In some instances, libraries of SARS-CoV-2 or ACE2 binding domains comprise sequences of SARS-CoV-2 or ACE2 binding domains designed based on conformational ligand interactions, peptide ligand interactions, small molecule ligand interactions, extracellular domains of SARS-CoV-2 or ACE2, or antibodies that target SARS-CoV-2 or ACE2. Libraries of SARS-CoV-2 or ACE2 binding domains may be translated to generate protein libraries. In some instances, libraries of SARS-CoV-2 or ACE2 binding domains are translated to generate peptide libraries, immunoglobulin libraries, derivatives thereof, or combinations thereof. In some instances, libraries of SARS-CoV-2 or ACE2 binding domains are translated to generate protein libraries that are further modified to generate peptidomimetic libraries. In some instances, libraries of SARS-CoV-2 or ACE2 binding domains are translated to generate protein libraries that are used to generate small molecules.

[0098] Methods described herein provide for synthesis of libraries of SARS-CoV-2 or ACE2 binding domains comprising nucleic acids each encoding for a predetermined variant of at least one predetermined reference nucleic acid sequence. In some cases, the predetermined reference sequence is a nucleic acid sequence encoding for a protein, and the variant library comprises sequences encoding for variation of at least a single codon such that a plurality of different variants of a single residue in the subsequent protein encoded by the synthesized nucleic acid are generated by standard translation processes. In some instances, the libraries of SARS-CoV-2 or ACE2 binding domains comprise varied nucleic acids collectively encoding variations at multiple positions. In some instances, the variant library comprises sequences encoding for variation of at least a single codon in a SARS-CoV-2 or ACE2 binding domain. In some instances, the variant library comprises sequences encoding for variation of multiple codons in a SARS-CoV-2 or ACE2 binding domain. An exemplary number of codons for variation include, but are not limited to, at least or about 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 225, 250, 275, 300, or more than 300 codons.

[0099] Methods described herein provide for synthesis of libraries comprising nucleic acids encoding for the SARS-CoV-2 or ACE2 binding domains, wherein the libraries comprise sequences encoding for variation of length of the SARS-CoV-2 or ACE2 binding domains. In some instances, the library comprises sequences encoding for variation of length of at least or about 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 225, 250, 275, 300, or more than 300 codons less as compared to a predetermined reference sequence. In some instances, the library comprises sequences encoding for variation of length of at least or about 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 225, 250, 275, 300, or more than 300 codons more as compared to a predetermined reference sequence.

[00100] Following identification of SARS-CoV-2 or ACE2 binding domains, antibodies may be designed and synthesized to comprise the SARS-CoV-2 or ACE2 binding domains. Antibodies comprising SARS-CoV-2 or ACE2 binding domains may be designed based on binding, specificity, stability, expression, folding, or downstream activity. In some instances, the antibodies comprising SARS-CoV-2 or ACE2 binding domains enable contact with the SARS-CoV-2 or ACE2. In some instances, the antibodies comprising SARS-CoV-2 or ACE2 binding domains enables high affinity binding with the SARS-CoV-2 or ACE2. Exemplary amino acid sequences of SARS-CoV-2 or ACE2 binding domains comprise any one of SEQ ID NOs: 1-716.

[00101] In some instances, the SARS-CoV-2 antibody comprises a binding affinity (e.g., K_D) to SARS-CoV-2 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 SARS-CoV-2 antibody comprises a K_D of less than 1 nM. In some instances, the SARS-CoV-2 antibody comprises a K_D of less than 1.2 nM. In some instances, the SARS-CoV-2 antibody comprises a K_D of less than 2 nM. In some instances, the SARS-CoV-2 antibody comprises a K_D of less than 5 nM. In some instances, the SARS-CoV-2 antibody comprises a K_D of less than 10 nM. In some instances, the SARS-CoV-2 antibody comprises a K_D of less than 13.5 nM. In some instances, the SARS-CoV-2 antibody comprises a K_D of less than 15 nM. In some instances, the SARS-CoV-2 antibody comprises a K_D of less than 20 nM. In some instances, the SARS-CoV-2 antibody comprises a K_D of less than 25 nM. In some instances, the SARS-CoV-2 antibody comprises a K_D of less than 30 nM.

[00102] In some instances, the ACE2 antibody comprises a binding affinity (e.g., K_D) to ACE2 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 ACE2 antibody comprises a K_D of less than 1 nM. In some instances, the ACE2 antibody comprises a K_D of less than 1.2 nM. In some instances, the ACE2 antibody comprises a K_D of less than 2 nM. In some instances, the ACE2 antibody comprises a K_D of less than 5 nM. In some instances, the ACE2 antibody comprises a K_D of less than 10 nM. In some instances, the ACE2 antibody comprises a K_D of less than 13.5 nM. In some instances, the ACE2 antibody comprises a K_D of less than 15 nM. In some instances, the ACE2 antibody comprises a K_D of less than 20 nM. In some instances, the ACE2 antibody comprises a K_D of less than 25 nM. In some instances, the ACE2 antibody comprises a K_D of less than 30 nM.

[00103] In some instances, the SARS-CoV-2 or ACE2 immunoglobulin is an agonist. In some instances, the SARS-CoV-2 or ACE2 immunoglobulin is an antagonist. In some instances, the SARS-CoV-2 or ACE2 immunoglobulin is an allosteric modulator. In some instances, the allosteric modulator is a negative allosteric modulator. In some instances, the allosteric modulator is a positive allosteric modulator. In some instances, the SARS-CoV-2 or ACE2 immunoglobulin results in agonistic, antagonistic, or allosteric effects at a concentration of at least or about 1 nM, 2 nM, 4 nM, 6 nM, 8 nM, 10 nM, 20 nM, 30 nM, 40 nM, 50 nM, 60 nM, 70 nM, 80 nM, 90 nM, 100 nM, 120 nM, 140 nM, 160 nM, 180 nM, 200 nM, 300 nM, 400 nM, 500 nM, 600 nM, 700 nM, 800 nM, 900 nM, 1000 nM, or more than 1000 nM. In some

instances, the SARS-CoV-2 or ACE2 immunoglobulin is a negative allosteric modulator. In some instances, the SARS-CoV-2 or ACE2 immunoglobulin is a negative allosteric modulator at a concentration of at least or about 0.001, 0.005, 0.01, 0.05, 0.1, 0.5, 1 nM, 2 nM, 4 nM, 6 nM, 8 nM, 10 nM, 20 nM, 30 nM, 40 nM, 50 nM, 60 nM, 70 nM, 80 nM, 90 nM, 100 nM, or more than 100 nM. In some instances, the SARS-CoV-2 or ACE2 immunoglobulin is a negative allosteric modulator at a concentration in a range of about 0.001 to about 100, 0.01 to about 90, about 0.1 to about 80, 1 to about 50, about 10 to about 40 nM, or about 1 to about 10 nM. In some instances, the SARS-CoV-2 or ACE2 immunoglobulin comprises an EC₅₀ or IC₅₀ of at least or about 0.001, 0.0025, 0.005, 0.01, 0.025, 0.05, 0.06, 0.07, 0.08, 0.9, 0.1, 0.5, 1, 2, 3, 4, 5, 6, or more than 6 nM. In some instances, the SARS-CoV-2 or ACE2 immunoglobulin comprises an EC₅₀ or IC₅₀ of at least or about 1 nM, 2 nM, 4 nM, 6 nM, 8 nM, 10 nM, 20 nM, 30 nM, 40 nM, 50 nM, 60 nM, 70 nM, 80 nM, 90 nM, 100 nM, or more than 100 nM.

[00104] In some instances, the affinity of the SARS-CoV-2 or ACE2 antibody 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 SARS-CoV-2 or ACE2 antibody 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.

[00105] Provided herein are SARS-CoV-2 or ACE2 binding libraries comprising nucleic acids encoding for antibodies comprising SARS-CoV-2 or ACE2 binding domains comprise variation in domain type, domain length, or residue variation. In some instances, the domain is a region in the antibody comprising the SARS-CoV-2 or ACE2 binding domains. For example, the region is the VH, CDRH3, or VL domain. In some instances, the domain is the SARS-CoV-2 or ACE2 binding domain.

[00106] Methods described herein provide for synthesis of a SARS-CoV-2 or ACE2 binding library of nucleic acids each encoding for a predetermined variant of at least one predetermined reference nucleic acid sequence. In some cases, the predetermined reference sequence is a nucleic acid sequence encoding for a protein, and the variant library comprises sequences encoding for variation of at least a single codon such that a plurality of different variants of a single residue in the subsequent protein encoded by the synthesized nucleic acid are generated by standard translation processes. In some instances, the SARS-CoV-2 or ACE2 binding library comprises varied nucleic acids collectively encoding variations at multiple positions. In some instances, the variant library comprises sequences encoding for variation of at least a single codon of a VH or VL domain. In some instances, the variant library comprises sequences encoding for variation of at least a single codon in a SARS-CoV-2 or ACE2 binding domain. For example, at least one single codon of a SARS-CoV-2 or ACE2 binding domain is varied. In some instances, the variant library comprises sequences encoding for variation of multiple codons of a VH or VL domain. In some instances, the variant library comprises sequences encoding for variation of multiple codons in a SARS-CoV-2 or ACE2 binding domain. An exemplary number of codons for variation include, but are not limited to, at least or about 1, 5,

10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 225, 250, 275, 300, or more than 300 codons.

[00107] Methods described herein provide for synthesis of a SARS-CoV-2 or ACE2 binding library of nucleic acids each encoding for a predetermined variant of at least one predetermined reference nucleic acid sequence, wherein the SARS-CoV-2 or ACE2 binding library comprises sequences encoding for variation of length of a domain. In some instances, the domain is VH or VL domain. In some instances, the domain is the SARS-CoV-2 or ACE2 binding domain. In some instances, the library comprises sequences encoding for variation of length of at least or about 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 225, 250, 275, 300, or more than 300 codons less as compared to a predetermined reference sequence. In some instances, the library comprises sequences encoding for variation of length of at least or about 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 225, 250, 275, 300, or more than 300 codons more as compared to a predetermined reference sequence.

[00108] Provided herein are SARS-CoV-2 or ACE2 binding libraries comprising nucleic acids encoding for antibodies comprising SARS-CoV-2 or ACE2 binding domains, wherein the SARS-CoV-2 or ACE2 binding libraries are synthesized with various numbers of fragments. In some instances, the fragments comprise the VH or VL domain. In some instances, the SARS-CoV-2 or ACE2 binding libraries are synthesized with at least or about 2 fragments, 3 fragments, 4 fragments, 5 fragments, or more than 5 fragments. The length of each of the nucleic acid fragments or average length of the nucleic acids synthesized may be at least or about 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, or more than 600 base pairs. In some instances, the length is about 50 to 600, 75 to 575, 100 to 550, 125 to 525, 150 to 500, 175 to 475, 200 to 450, 225 to 425, 250 to 400, 275 to 375, or 300 to 350 base pairs.

[00109] SARS-CoV-2 or ACE2 binding libraries comprising nucleic acids encoding for antibodies comprising SARS-CoV-2 or ACE2 binding domains as described herein comprise various lengths of amino acids when translated. In some instances, the length of each of the amino acid fragments or average length of the amino acid synthesized may be at least or about 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, or more than 150 amino acids. In some instances, the length of the amino acid is about 15 to 150, 20 to 145, 25 to 140, 30 to 135, 35 to 130, 40 to 125, 45 to 120, 50 to 115, 55 to 110, 60 to 110, 65 to 105, 70 to 100, or 75 to 95 amino acids. In some instances, the length of the amino acid is about 22 to about 75 amino acids.

[00110] SARS-CoV-2 or ACE2 binding libraries comprising de novo synthesized variant sequences encoding for antibodies comprising SARS-CoV-2 or ACE2 binding domains comprise a number of variant sequences. In some instances, a number of variant sequences is de novo synthesized for a CDRH1, CDRH2, CDRH3, CDRL1, CDRL2, CDRL3, VL, VH, or a combination thereof. In some instances, a number of variant sequences is de novo synthesized for framework element 1 (FW1), framework element 2 (FW2), framework element 3 (FW3), or framework element 4 (FW4). In some instances, a number of variant sequences are de novo synthesized for a SARS-CoV-2 or ACE2 binding domain. The number of variant

sequences may be at least or about 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, or more than 500 sequences. In some instances, the number of variant sequences is about 10 to 300, 25 to 275, 50 to 250, 75 to 225, 100 to 200, or 125 to 150 sequences.

[00111] SARS-CoV-2 or ACE2 binding libraries comprising de novo synthesized variant sequences encoding for antibodies comprising SARS-CoV-2 or ACE2 binding domains comprise improved diversity. In some instances, variants include affinity maturation variants. Alternatively or in combination, variants include variants in other regions of the antibody including, but not limited to, CDRH1, CDRH2, CDRL1, CDRL2, and CDRL3. In some instances, the number of variants of the SARS-CoV-2 or ACE2 binding libraries is least or about 10^4 , 10^5 , 10^6 , 10^7 , 10^8 , 10^9 , 10^{10} , 10^{11} , 10^{12} , 10^{13} , 10^{14} or more than 10^{14} non-identical sequences.

[00112] Following synthesis of SARS-CoV-2 or ACE2 binding libraries comprising nucleic acids encoding antibodies comprising SARS-CoV-2 or ACE2 binding domains, libraries may be used for screening and analysis. For example, libraries are assayed for library displayability and panning. In some instances, displayability is assayed using a selectable tag. Exemplary tags include, but are not limited to, a radioactive label, a fluorescent label, an enzyme, a chemiluminescent tag, a colorimetric tag, an affinity tag or other labels or tags that are known in the art. In some instances, the tag is histidine, polyhistidine, myc, hemagglutinin (HA), or FLAG. For example, SARS-CoV-2 or ACE2 binding libraries comprise nucleic acids encoding antibodies comprising SARS-CoV-2 or ACE2 binding domains with multiple tags such as GFP, FLAG, and Lucy as well as a DNA barcode. In some instances, libraries are assayed by sequencing using various methods including, but not limited to, single-molecule real-time (SMRT) sequencing, Polony sequencing, sequencing by ligation, reversible terminator sequencing, proton detection sequencing, ion semiconductor sequencing, nanopore sequencing, electronic sequencing, pyrosequencing, Maxam-Gilbert sequencing, chain termination (e.g., Sanger) sequencing, +S sequencing, or sequencing by synthesis.

[00113] 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.

[00114] In some embodiments, the antibody is a multivalent antibody. In some embodiments, the antibody is a monovalent, bivalent, or multivalent antibody. In some instances, the antibody is monospecific, bispecific, or multispecific. In some embodiments, the antibody is monovalent monospecific, monovalent bispecific, monovalent multispecific, bivalent monospecific, bivalent bispecific, bivalent multispecific, multivalent monospecific, multivalent bispecific, multivalent multispecific. In some instances, the antibody is homodimeric, heterodimeric, or heterotrimeric.

[00115] 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.

[00116] “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.

[00117] “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.

[00118] “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 felinize antibody.

[00119] Methods as described herein may be used for optimization 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, DARPs, fynomers, Kunitz domain-based proteins, monobodies, anticalins, knottins, armadillo repeat protein-based proteins, and bicyclic peptides.

[00120] 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 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).

[00121] Methods described herein provide for synthesis of libraries comprising nucleic acids encoding an antibody, 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.

[00122] 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). Exemplary regions of the antibody for variation include, but are not limited to, IGHV1-18, IGHV1-69, IGHV1-8, IGHV3-21, IGHV3-23, IGHV3-30/33rn, IGHV3-28, IGHV1-69, IGHV3-74, IGHV4-39, IGHV4-59/61, IGKV1-39, IGKV1-9, IGKV2-28, IGKV3-11, IGKV3-15, IGKV3-20, IGKV4-1, IGLV1-51, IGLV2-14, IGLV1-40, and IGLV3-1. In some instances, the gene is IGHV1-69, IGHV3-30, IGHV3-23, IGHV3, IGHV1-46, IGHV3-7, IGHV1, or IGHV1-8. In some instances, the gene is IGHV1-69 and IGHV3-30. In some instances, the region of the antibody for variation is IGHJ3, IGHJ6, IGHJ, IGHJ4, IGHJ5, IGHJ2, or IGHJ1. In some instances, the region of the antibody for variation is IGHJ3, IGHJ6, IGHJ, or IGHJ4. In some instances, the at least one region of the antibody for

variation is IGHV1-69, IGHV3-23, IGKV3-20, IGKV1-39, or combinations thereof. In some instances, the at least one region of the antibody for variation is IGHV1-69 and IGKV3-20. In some instances, the at least one region of the antibody for variation is IGHV1-69 and IGKV1-39. In some instances, the at least one region of the antibody for variation is IGHV3-23 and IGKV3-20. In some instances, the at least one region of the antibody for variation is IGHV3-23 and IGKV1-39.

[00123] 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.

[00124] Libraries comprising nucleic acids encoding for antibodies 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.

[00125] 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.

[00126] 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.

[00127] 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. 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.

[00128] 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.

[00129] 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.

[00130] *Expression Systems*

[00131] 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.

[00132] 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.

[00133] 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.

[00134] 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.

[00135] 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, 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.

[00136] 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.

[00137] Diseases and Disorders

[00138] Provided herein are SARS-CoV-2 or ACE2 binding libraries comprising nucleic acids encoding for antibodies comprising SARS-CoV-2 or ACE2 binding domains may have therapeutic effects. In some instances, the SARS-CoV-2 or ACE2 binding libraries result in protein when translated that is used to treat a disease or disorder. In some instances, the protein is an immunoglobulin. In some instances, the protein is a peptidomimetic. In some instances, the disease or disorder is a viral infection caused by SARS-CoV-2. In some instances, the disease or disorder is a respiratory disease or disorder caused by SARS-CoV-2.

[00139] SARS-CoV-2 or ACE2 variant antibody libraries as described herein may be used to treat SARS-CoV-2. In some embodiments, the SARS-CoV-2 or ACE2 variant antibody libraries are used to treat or prevent symptoms of SARS-CoV-2. These symptoms include, but are not limited to, fever, chills, cough, fatigue, headaches, loss of taste, loss of smell, nausea, vomiting, muscle weakness, sleep difficulties, anxiety, and depression. In some embodiments, the SARS-CoV-2 or ACE2 variant antibody libraries are used to treat a subject who has symptoms for an extended period of time. In some embodiments, the subject has symptoms for an extended period of time after testing negative for SARS-CoV-2. In some embodiments, the subject has symptoms for an extended period of time including at least 1 week, 2 weeks, 3 weeks, 1 month, 2 months, 3 months, 4 months, 5 months, 6 months, 7 months, 8 months, 9 months, 10 months, 11 months, 1 year, or more than 1 year.

[00140] 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. In some instances, a pharmaceutical composition comprises an antibody or antibody fragment described herein comprising a CDRH1 sequence of any one of SEQ ID NOs: 217-266, 1495-1635, or 2059-2238. In some instances, a pharmaceutical composition comprises an antibody or antibody fragment described herein comprising a CDRH2 sequence of any one of SEQ ID NOs: 267-316, 1636-1776, or 2239-2418. In some instances, a pharmaceutical composition comprises an antibody or antibody fragment described herein comprising a CDRH3 sequence of any one of SEQ ID NOs: 317-366, 1777-1917, or 2419-2598. In some instances, a pharmaceutical composition comprises an antibody or antibody fragment described herein comprising a variable domain, heavy chain region (VH) and a variable domain, light chain region (VL), wherein VH comprises complementarity determining regions CDRH1, CDRH2, and CDRH3, wherein VL comprises complementarity determining regions CDRL1, CDRL2, and CDRL3, and wherein (a) an amino acid sequence of CDRH1 is as set forth in any one of SEQ ID NOs: 1-36 or 217-282; (b) an amino acid sequence of CDRH2 is as set forth in any one of SEQ ID NOs: 37-72 or 283-348; (c) an amino acid sequence of CDRH3 is as set forth in any one of SEQ ID NOs: 73-108 or 349-414; (d) an amino acid sequence of CDRL1 is as set forth in any one of SEQ ID NOs: 109-144 or 415-473; (e) an amino acid sequence of CDRL2 is as set forth in any one of SEQ ID NOs: 145-180 or 415-473; and (f) an amino acid sequence of CDRL3 is as set forth in any one of SEQ ID NOs: 181-216 or 533-591. In some instances, a pharmaceutical composition comprises an antibody or antibody fragment described herein comprising a VH 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: 592-657, and VL 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: 658-716. In some instances, a pharmaceutical composition comprises an antibody or antibody fragment described herein comprising 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: 1918-2058, 2599-2778, and 3095-3173.

[00141] SARS-CoV-2 or ACE2 antibodies as described herein may confer immunity after exposure to SARS-CoV-2 or ACE2 antibodies. In some embodiments, the SARS-CoV-2 or ACE2 antibodies described herein are used for passive immunization of a subject. In some instances, the subject is actively immunized after exposure to SARS-CoV-2 or ACE2 antibodies followed by exposure to SARS-CoV-2. In some embodiments, SARS-CoV-2 or ACE2 antibodies are derived from a subject who has recovered from SARS-CoV-2.

[00142] In some embodiments, the immunity occurs at least about 30 minutes, 1 hour, 5 hours, 10 hours, 16 hours, 20 hours, 24 hours, 2 days, 3 days, 4 days, 5 days, 6 days, 1 week, 2 weeks, or more than 2 weeks after exposure to SARS-CoV-2 or ACE2 antibodies. In some instances, the immunity lasts for at least about 1 day, 2 days, 3 days, 4 days, 5 days, 6 days, 1 week, 2 weeks, 3 weeks, 1 month, 2 months, 3 months, 4

months, 5 months, 6 months, 7 months, 8 months, 9 months, 10 months, 11 months, 1 year, 2 years, 3 years, 4 years, 5 years, or more than 5 years after exposure to SARS-CoV-2 or ACE2 antibodies.

[00143] In some embodiments, the subject receives the SARS-CoV-2 or ACE2 antibodies prior to exposure to SARS-CoV-2. In some embodiments, the subject receives the SARS-CoV-2 or ACE2 antibodies at least about 30 minutes, 1 hour, 4 hours, 8 hours, 12 hours, 16 hours, 20 hours, 1 day, 2 days, 3 days, 4 days, 5 days, 6 days, 1 week, 2 weeks, 3 weeks, 1 month, 2 months, 3 months, 4 months, 5 months, 6 months, 7 months, 8 months, 9 months, 10 months, 11 months, 1 year, 2 years, 3 years, 4 years, 5 years, or more than 5 years prior to exposure to SARS-CoV-2. In some embodiments, the subject receives the SARS-CoV-2 or ACE2 antibodies after exposure to SARS-CoV-2. In some embodiments, the subject receives the SARS-CoV-2 or ACE2 antibodies at least about 30 minutes, 1 hour, 4 hours, 8 hours, 12 hours, 16 hours, 20 hours, 1 day, 2 days, 3 days, 4 days, 5 days, 6 days, 1 week, 2 weeks, 3 weeks, 1 month, 2 months, 3 months, 4 months, 5 months, 6 months, 7 months, 8 months, 9 months, 10 months, 11 months, 1 year, 2 years, 3 years, 4 years, 5 years, or more than 5 years after exposure to SARS-CoV-2.

[00144] SARS-CoV-2 or ACE2 antibodies described herein may be administered through various routes. The administration may, depending on the composition being administered, for example, be oral, pulmonary, intravenous, intraperitoneal, intramuscular, intracavity, subcutaneous, or transdermal.

[00145] Described herein are compositions or pharmaceutical compositions comprising SARS-CoV-2 or ACE2 antibodies or antibody fragment thereof that comprise various dosages of the antibodies or antibody fragment. In some instances, the dosage is ranging from about 1 to 25 mg/kg, from about 1 to 50 mg/kg, from about 1 to 80 mg/kg, from about 1 to about 100 mg/kg, from about 5 to about 100 mg/kg, from about 5 to about 80 mg/kg, from about 5 to about 60 mg/kg, from about 5 to about 50 mg/kg or from about 5 to about 500 mg/kg which can be administered in single or multiple doses. In some instances, the dosage is administered in an amount of about 0.01 mg/kg, about 0.05 mg/kg, about 0.10 mg/kg, about 0.25 mg/kg, about 0.5 mg/kg, about 1 mg/kg, about 5 mg/kg, about 10 mg/kg, about 15 mg/kg, about 20 mg/kg, about 25 mg/kg, about 30 mg/kg, about 35 mg/kg, about 40 mg/kg, about 45 mg/kg, about 50 mg/kg, about 55 mg/kg, about 60 mg/kg, about 65 mg/kg, about 70 mg/kg, about 75 mg/kg, about 80 mg/kg, about 85 mg/kg, about 90 mg/kg, about 95 mg/kg, about 100 mg/kg, about 105 mg/kg, about 110 mg/kg, about 115 mg/kg, about 120, about 125, about 130, about 135, about 140, about 145, about 150, about 155, about 160, about 165, about 170, about 175, about 180, about 185, about 190, about 195, about 200, about 205, about 210, about 215, about 220, about 225, about 230, about 240, about 250, about 260, about 270, about 275, about 280, about 290, about 300, about 310, about 320, about 330, about 340, about 350, about 360 mg/kg, about 370 mg/kg, about 380 mg/kg, about 390 mg/kg, about 400 mg/kg, 410 mg/kg, about 420 mg/kg, about 430 mg/kg, about 440 mg/kg, about 450 mg/kg, about 460 mg/kg, about 470 mg/kg, about 480 mg/kg, about 490 mg/kg, or about 500 mg/kg.

[00146] SARS-CoV-2 or ACE2 antibodies or antibody fragment thereof described herein, in some embodiments, improve disease severity. In some embodiments, the SARS-CoV-2 or ACE2 antibodies or antibody fragment thereof improve disease severity at a dose level of about 0.01 mg/kg, about 0.05 mg/kg,

about 0.10 mg/kg, about 0.25 mg/kg, about 0.5 mg/kg, about 1 mg/kg, about 5 mg/kg, about 10 mg/kg, about 15 mg/kg, or about 20 mg/kg. In some embodiments, the SARS-CoV-2 or ACE2 antibodies or antibody fragment thereof improve disease severity at a dose level of about 1 mg/kg, about 5 mg/kg, or about 10 mg/kg. In some embodiments, disease severity is determined by percent weight loss. In some embodiments, the SARS-CoV-2 or ACE2 antibodies or antibody fragment thereof protects against weight loss at a dose level of about 0.01 mg/kg, about 0.05 mg/kg, about 0.10 mg/kg, about 0.25 mg/kg, about 0.5 mg/kg, about 1 mg/kg, about 5 mg/kg, about 10 mg/kg, about 15 mg/kg, or about 20 mg/kg. In some embodiments, the SARS-CoV-2 or ACE2 antibodies or antibody fragment thereof protects against weight loss at a dose level of about 1 mg/kg, about 5 mg/kg, or about 10 mg/kg. In some embodiments, SARS-CoV-2 or ACE2 antibodies or antibody fragment thereof

[00147] Variant Libraries

[00148] Codon variation

[00149] 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 1** provides a listing of each codon possible (and the representative amino acid) for a variant site.

Table 1. 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				

[00150] 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.

[00151] Highly Parallel Nucleic Acid Synthesis

[00152] 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.

[00153] 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.

[00154] 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 variants 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 or variants 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.

[00155] 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.

[00156] 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 .

[00157] 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.

[00158] 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.

[00159] Substrates

[00160] 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.

[00161] 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.

[00162] 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.

[00163] 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 mm². 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 mm². 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 μ m. 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 μ m. 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 μ m. 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 μ m. In some cases, each locus has a width of about 0.5-100, 0.5-50, 10-75, or 0.5-50 μ m.

[00164] In some instances, the density of clusters within a substrate is at least or about 1 cluster per 100 mm², 1 cluster per 10 mm², 1 cluster per 5 mm², 1 cluster per 4 mm², 1 cluster per 3 mm², 1 cluster per 2

mm², 1 cluster per 1 mm², 2 clusters per 1 mm², 3 clusters per 1 mm², 4 clusters per 1 mm², 5 clusters per 1 mm², 10 clusters per 1 mm², 50 clusters per 1 mm² or more. In some instances, a substrate comprises from about 1 cluster per 10 mm² to about 10 clusters per 1 mm². 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 μ m. 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 μ m. 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.

[00165] In some instances, a substrate is about the size of a standard 96 well plate, for example between about 100 and about 200 mm by between about 50 and about 150 mm. In some instances, a substrate has a diameter less than or equal to about 1000, 500, 450, 400, 300, 250, 200, 150, 100 or 50 mm. In some instances, the diameter of a substrate is between about 25-1000, 25-800, 25-600, 25-500, 25-400, 25-300, or 25-200 mm. In some instances, a substrate has a planar surface area of at least about 100; 200; 500; 1,000; 2,000; 5,000; 10,000; 12,000; 15,000; 20,000; 30,000; 40,000; 50,000 mm² or more. In some instances, the thickness of a substrate is between about 50- 2000, 50- 1000, 100-1000, 200-1000, or 250-1000 mm.

[00166] *Surface materials*

[00167] 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; fuse 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.

[00168] Surface Architecture

[00169] 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.

[00170] 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.

[00171] 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.

[00172] 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.

[00173] 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 μ m. In some cases, the height of a well is less than about 1000, 900, 800, 700, or 600 μ m.

[00174] 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 μ m. In some cases, the height of a channel is less than 100, 80, 60, 40, or 20 μ m.

[00175] 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 μ m, for example, about 90, 80, 70, 60, 50, 40, 30, 20 or 10 μ m. 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 μ m. 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 μ m.

[00176] Surface Modifications

[00177] 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.

[00178] 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.

[00179] 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.

[00180] 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.

[00181] 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.

[00182] Polynucleotide Synthesis

[00183] 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.

[00184] 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).

[00185] 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.

[00186] 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).

[00187] 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.

[00188] 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.

[00189] 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.

[00190] 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.

[00191] 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.

[00192] 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.

[00193] 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.

[00194] 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.

[00195] Referring to the Figures, **FIG. 2** 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.

[00196] 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.

[00197] 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 **201**, 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 **202**. In some instances, polynucleotides are cleaved from the surface at this stage. Cleavage includes gas cleavage, e.g., with ammonia or methylamine.

[00198] 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 203. Prior to or after the sealing 204 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 205. Once released, fragments of single stranded polynucleotides hybridize in order to span an entire long range sequence of DNA. Partial hybridization 205 is possible because each synthesized polynucleotide is designed to have a small portion overlapping with at least one other polynucleotide in the pool.

[00199] 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 206.

[00200] After PCA is complete, the nanoreactor is separated from the device 207 and positioned for interaction with a device having primers for PCR 208. After sealing, the nanoreactor is subject to PCR 209 and the larger nucleic acids are amplified. After PCR 210, the nanochamber is opened 211, error correction reagents are added 212, the chamber is sealed 213 and an error correction reaction occurs to remove mismatched base pairs and/or strands with poor complementarity from the double stranded PCR amplification products 214. The nanoreactor is opened and separated 215. Error corrected product is next subject to additional processing steps, such as PCR and molecular bar coding, and then packaged 222 for shipment 223.

[00201] 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 216, sealing the wafer to a chamber containing error corrected amplification product 217, and performing an additional round of amplification 218. The nanoreactor is opened 219 and the products are pooled 220 and sequenced 221. After an acceptable quality control determination is made, the packaged product 222 is approved for shipment 223.

[00202] In some instances, a nucleic acid generate by a workflow such as that in FIG. 2 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 202.

[00203] Computer systems

[00204] 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.

[00205] The computer system **300** illustrated in **FIG. 3** may be understood as a logical apparatus that can read instructions from media **311** and/or a network port **305**, which can optionally be connected to server **309** having fixed media **312**. The system, such as shown in **FIG. 3** can include a CPU **301**, disk drives **303**, optional input devices such as keyboard **315** and/or mouse **316** and optional monitor **307**. 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 **322** as illustrated in **FIG. 3**.

[00206] **FIG. 4** is a block diagram illustrating a first example architecture of a computer system **400** that can be used in connection with example instances of the present disclosure. As depicted in **FIG. 4**, the example computer system can include a processor **402** for processing instructions. Non-limiting examples of processors include: Intel Xeon™ processor, AMD Opteron™ processor, Samsung 32-bit RISC ARM 1176JZ(F)-S v1.0™ processor, ARM Cortex-A8 Samsung S5PC100™ processor, ARM Cortex-A8 Apple A4™ processor, Marvell PXA 930™ processor, or a functionally-equivalent processor. Multiple threads of execution can be used for parallel processing. In some instances, multiple processors or processors with multiple cores can also be used, whether in a single computer system, in a cluster, or distributed across systems over a network comprising a plurality of computers, cell phones, and/or personal data assistant devices.

[00207] As illustrated in **FIG. 4**, a high speed cache **404** can be connected to, or incorporated in, the processor **402** to provide a high speed memory for instructions or data that have been recently, or are frequently, used by processor **402**. The processor **402** is connected to a north bridge **406** by a processor bus **408**. The north bridge **406** is connected to random access memory (RAM) **410** by a memory bus **412** and manages access to the RAM **410** by the processor **402**. The north bridge **406** is also connected to a south bridge **414** by a chipset bus **416**. The south bridge **414** is, in turn, connected to a peripheral bus **418**. 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 **418**. 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 **400** can include an accelerator card **422** attached to the peripheral bus **418**. 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.

[00208] Software and data are stored in external storage **424** and can be loaded into RAM **410** and/or cache **404** for use by the processor. The system **400** includes an operating system for managing system resources; non-limiting examples of operating systems include: Linux, WindowsTM, 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 **400** also includes network interface cards (NICs) **420** and **421** 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.

[00209] FIG. 5 is a diagram showing a network **500** with a plurality of computer systems **502a**, and **502b**, a plurality of cell phones and personal data assistants **502c**, and Network Attached Storage (NAS) **504a**, and **504b**. In example instances, systems **502a**, **502b**, and **502c** can manage data storage and optimize data access for data stored in Network Attached Storage (NAS) **504a** and **504b**. A mathematical model can be used for the data and be evaluated using distributed parallel processing across computer systems **502a**, and **502b**, and cell phone and personal data assistant systems **502c**. Computer systems **502a**, and **502b**, and cell phone and personal data assistant systems **502c** can also provide parallel processing for adaptive data restructuring of the data stored in Network Attached Storage (NAS) **504a** and **504b**. FIG. 5 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.

[00210] FIG. 6 is a block diagram of a multiprocessor computer system using a shared virtual address memory space in accordance with an example instance. The system includes a plurality of processors **602a-f** that can access a shared memory subsystem **604**. The system incorporates a plurality of programmable hardware memory algorithm processors (MAPs) **606a-f** in the memory subsystem **604**. Each MAP **606a-f** can comprise a memory **608a-f** and one or more field programmable gate arrays (FPGAs) **610a-f**. The MAP provides a configurable functional unit and particular algorithms or portions of algorithms can be provided to the FPGAs **610a-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 **608a-f**, allowing it to execute tasks independently of, and asynchronously

from the respective microprocessor **602a-f**. In this configuration, a MAP can feed results directly to another MAP for pipelining and parallel execution of algorithms.

[00211] 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.

[00212] 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. 4**, 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 **422** illustrated in **FIG. 4**.

[00213] 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

[00214] 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.

[00215] Example 1: Functionalization of a device surface

[00216] 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.

[00217] 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.

[00218] 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.

[00219] **Example 2: Synthesis of a 50-mer sequence on an oligonucleotide synthesis device**

[00220] 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-TRIETHOXYSIYLPROPYL)-4-HYDROXYBUTYRAMIDE (Gelest) was used to synthesize an exemplary polynucleotide of 50 bp ("50-mer polynucleotide") using polynucleotide synthesis methods described herein.

[00221] The sequence of the 50-mer was as described.
5'AGACAATCAACCATTTGGGGTGGACAGCCTTGACCTCTAGACTTCGGCAT##TTTTTTTTTTT3',
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.

[00222] 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

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

Table 2		
General DNA Synthesis Process Name	Process Step	Time (sec)
Flow)	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
	Acetonitrile System Flush	4
CAPPING (CapA+B, 1:1, Flow)	CapA+B to Flowcell	15
WASH (Acetonitrile Wash Flow)	Acetonitrile System Flush	4
	Acetonitrile to Flowcell	15
	Acetonitrile System Flush	4
OXIDATION (Oxidizer Flow)	Oxidizer to Flowcell	18
WASH (Acetonitrile Wash Flow)	Acetonitrile System Flush	4
	N2 System Flush	4
	Acetonitrile System Flush	4
	Acetonitrile to Flowcell	15
	Acetonitrile System Flush	4
	Acetonitrile to Flowcell	15
	N2 System Flush	4
	Acetonitrile System Flush	4
	Acetonitrile to Flowcell	23
	N2 System Flush	4
	Acetonitrile System Flush	4
DEBLOCKING (Deblock Flow)	Deblock to Flowcell	36
WASH (Acetonitrile Wash Flow)	Acetonitrile System Flush	4
	N2 System Flush	4
	Acetonitrile System Flush	4
	Acetonitrile to Flowcell	18
	N2 System Flush	4.13
	Acetonitrile System Flush	4.13

General DNA Synthesis Process Name	Table 2	
	Process Step	Time (sec)
	Acetonitrile to Flowcell	15

[00223] 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.

[00224] 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.

[00225] Example 3: Synthesis of a 100-mer sequence on an oligonucleotide synthesis device

[00226] 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'

CGGGATCCTTATCGTCATCGTCGTACAGATCCCGACCCATTTGCTGTCCACCAGTCATGCTAGC
CATAACCATGATGATGATGATGATGAGAACCCCGCAT##TTTTTTTTTTT3', where # denotes

Thymidine-succinyl hexamide CED phosphoramidite (CLP-2244 from ChemGenes) on two different silicon chips, the first one uniformly functionalized with N-(3-TRIETHOXYSILYLPROPYL)-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.

[00227] All ten samples from the two chips were further PCR amplified using a forward (5'ATGCGGGTTCTCATCATC3') and a reverse (5'CGGGATCCTTATCGTCATCG3') 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

[00228] 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%
7	1/531 bp	99.81%
8	1/1769 bp	99.94%
9	1/854 bp	99.88%
10	1/1451 bp	99.93%

[00229] 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.

[00230] Table 4 summarizes error characteristics for the sequences obtained from the polynucleotides samples from spots 1-10.

Table 4: Error characteristics

Sample ID/Spot no.	OSA_00 46/1	OSA_00 47/2	OSA_00 48/3	OSA_00 49/4	OSA_00 50/5	OSA_00 51/6	OSA_00 52/7	OSA_00 53/8	OSA_00 54/9	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

[00231] Example 4: Panning and screening for identification of antibodies for SARS-CoV-2 and ACE2

[00232] This example describes identification of antibodies for SARS-CoV-2 and ACE2.

[00233] Phage displayed scFv, VHH, and Fab libraries were panned for binding to biotinylated SARS-CoV-2 S1 and human ACE2. **FIG. 7** shows a schema of the panning strategy. Biotinylated antigen was bound to streptavidin coated magnetic beads at a density of 100 pmol antigen per mg of beads (Thermo Fisher #11206D). Phage libraries were blocked with 5% BSA in PBS. Following magnetic bead depletion for 1 hour at room temperature (RT), the beads were removed, and phage supernatant was transferred to 1 mg of antigen-bound beads in 1 ml PBS and incubated at RT with rotation for 1 hour. Non-binding clones were washed away by addition of 1 ml PBST, increasing the number of washes with each panning round. Trypsin was used to elute the phage bound to the antigen-bead complex. Phage were amplified in TG1 *E. coli* for the next round of selection. This selection strategy was repeated for four rounds, with successively lower amounts of antigen per round. Following all four selection rounds, 400 clones from each of round 2, 3, and 4 were selected for phage expression and phage ELISA screening. Data from the panning is seen in **Table 5**.

Table 5. Panning Data

Antibody	Titer	Round 1	Round 2	Round 3	Round 4	Round 5
1	Input titer	1.5×10^{12}	1.2×10^{13}	4.4×10^{13}	1.8×10^{13}	-
	Output titer	1.2×10^6	1.5×10^6	2.0×10^6	1.4×10^8	-
2	Input titer	1.4×10^{12}	2.6×10^{13}	3.0×10^{13}	1.0×10^{13}	-
	Output titer	9.5×10^5	1.2×10^6	2.2×10^6	1.2×10^8	-
3	Input titer	1.7×10^{12}	2.0×10^{13}	2.8×10^{13}	3.2×10^{13}	-
	Output titer	1.5×10^5	1.7×10^6	1.5×10^6	1.1×10^8	-
4	Input titer	1.2×10^{12}	1.6×10^{13}	3.6×10^{13}	1.5×10^{13}	-
	Output titer	1.3×10^5	2.2×10^7	2.6×10^7	2.5×10^8	-

[00234] To test for binding to SARS-CoV-2 S1 and ACE2, phage were expressed from each picked colony by KO7 superinfection in 384 well microtiter plates. Phage containing supernatant was blocked by 1:1

addition of 4% non-fat milk (NFM). Assay plates were prepared by passive immobilization of 0.4 µg antigen in 384-well Maxisorp plates (Thermo Fisher #464718) and then blocked with 4% NFM. Following 3x wash in PBST, blocked phage supernatant was incubated for 1 hour at RT. After 3x wash in PBST, 0.3 µg/ml anti-M13-HRP (Sino Biological #11973-MM05T-H) was aliquoted for 1 hour incubation at room temperature. Binding of phage-displayed antibody was determined by absorbance at 450 nm with 3,3',5,5'-tetramethylbenzidine (Thermo Fisher #34029). Phage that bound to antigen with 3x over background of human Fc protein were identified as potential binders for sequencing analysis. DNA was amplified by rolling circle amplification from glycerol stocks of each clone and submitted for Sanger sequencing (Genewiz) to capture the VH and VL domains. **FIGS. 8A-8D** shows phage ELISA data from round 4 of SARS-CoV-2 S1 (subunit 1) protein panning for antibody 1 (**FIGS. 8A-8B**) and antibody 2 (**FIGS. 8C-8D**). **FIGS. 9A-9D** shows phage ELISA data from round 4 of ACE2 protein panning for antibody 3 (**FIGS. 9A-9B**) and antibody 4 (**FIGS. 9C-9D**).

[00235] SARS-CoV-2 variants were tested for specificity using a phage ELISA as described above. The antigens used included Acro COVID S1 (S1N-C82E8), COVID S1 RBD Fc fusion (Antigen 1), and COVID S1 RBM Fc fusion (Antigen 2). Data from the phage ELISA is seen in **Table 6A**. **Table 6A** shows screening ELISA mean, fold over background (column A), specificity ELISA, fold over background (column B), and specificity ELISA, percent binding relative to binding to Acro S1 (column C). As seen in **Table 6A**, nearly all receptor binding domain (RBD) specific clones show good binding to full length subunit 1 (S1) and produced S1 RBD Fc. None of the S1 RBD variants were found to bind to S1 RBM Fc.

Table 6A. SARS-CoV-2 Phage ELISA

	Column A	Column B				Column C		
Variant	ELISA (Avg)	Acro S1	S1-RBD-Fc	S1-Fc	S1-RBM-Fc	S1-RBD-Fc	S1-Fc	S1-RBM-Fc
l-21	47.6	19.5	14.8	1.4	1.4	75.7%	7.2%	7.0%
l-22	40.6	32.7	26.5	2.5	1.5	81.2%	7.6%	4.7%
l-30	29.5	16.8	1.7	5.6	1.4	9.8%	33.3%	8.5%
l-35	28.6	21.6	1.8	1.4	1.6	8.2%	6.5%	7.3%
l-17	27.4	25.1	20.1	1.6	1.3	79.8%	6.2%	5.2%
l-27	27.2	1.5	1.9	1.3	1.6	124.5%	84.8%	108.8%
l-37	27.0	11.6	2.0	5.4	1.4	17.7%	46.7%	12.3%
l-12	25.7	1.7	1.7	1.3	1.7	104.0%	79.6%	102.5%
l-2	24.2	1.8	1.9	1.7	1.7	103.3%	95.4%	93.5%
l-3	23.9	1.6	1.8	1.5	1.5	109.3%	91.2%	91.9%
l-23	23.1	5.1	3.6	1.6	1.6	70.9%	31.1%	32.0%
l-7	22.5	6.8	3.3	1.7	1.6	48.7%	24.8%	23.0%
l-31	20.7	30.4	1.3	15.7	1.2	4.3%	51.6%	3.8%

1-4	20.6	1.0	1.4	0.9	1.1	137.5%	90.8%	105.3%
1-38	20.2	1.1	1.5	0.9	1.2	127.9%	82.8%	100.9%
1-8	19.3	12.0	11.2	1.1	1.2	92.9%	8.8%	10.3%
1-9	18.5	10.7	9.8	0.9	1.0	91.1%	8.7%	9.3%
1-32	17.9	11.5	1.5	4.0	1.1	13.1%	35.3%	9.5%
1-33	17.6	7.1	1.5	3.0	1.0	21.1%	42.1%	14.6%
1-24	32.9	12.9	11.0	1.1	1.3	85.4%	8.9%	10.3%
1-39	24.4	18.6	11.6	2.5	1.3	62.5%	13.2%	6.7%
1-40	22.9	18.7	14.3	1.3	1.1	76.7%	7.1%	6.0%
1-5	20.6	1.5	1.6	1.2	1.4	107.8%	82.9%	92.0%
1-41	18.1	4.5	2.7	1.3	1.2	58.8%	28.9%	26.3%
1-28	17.7	1.0	1.1	1.1	1.2	112.9%	109.7%	124.1%
1-10	16.9	17.4	17.4	1.3	1.1	100.5%	7.3%	6.2%
2-1	45.3	39.3	36.8	20.2	1.5	93.7%	51.3%	3.9%
2-10	43.8	38.9	39.9	9.4	1.2	102.7%	24.1%	3.1%
2-5	30.8	38.3	35.9	24.3	1.1	93.7%	63.3%	3.0%
2-2	23.4	39.4	39.6	4.7	1.1	100.4%	12.0%	2.8%
3-10	17.4	1.2	1.2	1.2	1.1	97.2%	99.6%	93.3%
1-26	19.5	1.4	1.1	1.3	1.0	76.8%	98.0%	74.7%
1-42	34.5	22.7	20.8	1.4	1.0	91.4%	6.4%	4.2%
1-13	28.2	4.8	4.3	0.9	1.2	89.3%	18.4%	24.2%
1-43	21.9	6.7	5.8	3.9	1.0	87.3%	58.3%	14.5%
1-44	24.6	10.4	8.2	1.0	0.8	78.6%	9.6%	7.6%
1-14	21.7	16.8	13.5	1.1	0.9	80.6%	6.7%	5.4%
1-6	20.8	1.8	1.3	1.1	0.8	69.6%	60.5%	45.4%
1-45	24.0	12.1	10.3	1.2	1.0	85.3%	9.6%	8.3%
1-46	21.7	4.6	3.1	1.1	0.9	66.6%	24.5%	19.8%
1-20	26.0	5.7	3.9	1.0	0.8	67.3%	17.0%	14.5%
1-47	22.6	8.1	5.5	1.2	0.9	68.9%	14.6%	11.1%
1-29	23.8	1.1	0.9	1.0	0.9	81.4%	94.0%	78.3%
1-1	32.5	4.5	4.0	1.6	1.1	90.7%	35.3%	24.5%
1-19	22.3	24.2	23.9	1.5	1.2	98.7%	6.4%	4.9%
1-16	22.1	3.4	3.1	1.0	0.9	89.7%	29.5%	27.3%
1-34	28.7	7.5	3.1	1.8	1.0	41.3%	24.5%	13.3%
1-48	22.6	2.5	2.1	0.8	0.8	84.7%	31.5%	30.3%

1-49	27.6	1.0	1.0	0.8	0.8	99.3%	85.6%	78.8%
1-18	22.7	9.8	7.0	1.1	0.8	71.0%	11.0%	8.0%
1-11	33.2	4.8	4.8	0.9	1.0	99.5%	18.7%	20.6%
1-50	21.1	13.9	12.4	1.2	0.8	88.9%	8.8%	5.9%
1-25	27.9	3.6	2.7	1.0	0.8	75.4%	28.1%	22.3%
1-36	23.6	10.1	1.2	1.1	0.8	11.5%	10.6%	8.4%
1-15	24.8	2.5	1.4	1.0	0.8	55.4%	41.5%	32.1%
2-4	15.5	37.7	38.8	7.2	1.2	102.7%	19.0%	3.1%
2-6	22.1	11.1	12.8	1.8	1.1	115.8%	16.4%	9.5%
2-11	28.1	1.1	1.0	1.2	0.9	86.3%	103.4%	76.1%
2-12	18.2	39.8	40.3	14.7	1.1	101.3%	37.0%	2.9%
2-13	19.1	27.3	32.1	3.7	0.9	117.5%	13.4%	3.1%
2-14	17.2	31.7	32.9	4.2	0.9	103.9%	13.3%	2.8%
2-7	25.3	37.2	37.3	7.7	0.9	100.3%	20.6%	2.4%
2-8	32.4	35.9	36.5	5.4	1.1	101.8%	15.0%	3.1%
2-15	13.7	31.0	28.1	3.8	0.9	90.6%	12.4%	3.0%
2-9	14.1	24.1	24.3	3.0	0.8	100.7%	12.2%	3.5%

[00236] Tables 6B-6C show Carterra SPR kinetics for SARS-CoV-2 variant antibodies ranked by off-rate (Table 6B) and by K_D (Table 6C). FIG. 10 shows that ACE2 binds to S1-RBD-Fc and S1-Fc variants.

Table 6B. SARS-CoV-2 Carterra SPR Kinetics Ranked by Off-Rate

IgG	Acro S1	S1-RBD-Fc	S1-Fc	S1-RBM-Fc	k_a (M ⁻¹ s ⁻¹)	K_D (s ⁻¹)
2-10	38.9	39.9	9.4	1.2	2.08E+05	1.15E-04
2-5	38.3	35.9	24.3	1.1	9.39E+04	2.59E-04
1-31	30.4	1.3	15.7	1.2	2.05E+04	7.62E-04
2-2	39.4	39.6	4.7	1.1	7.74E+04	9.06E-04
1-13	4.8	4.3	0.9	1.2	2.17E+05	1.38E-03
1-30	16.8	1.7	5.6	1.4	4.39E+04	1.64E-03
1-29	1.1	0.9	1.0	0.9	5.35E+02	1.97E-03
1-27	1.5	1.9	1.3	1.6	2.27E+05	2.85E-03
1-35	21.6	1.8	1.4	1.6	5.90E+04	5.43E-03
1-36	10.1	1.2	1.1	0.8	1.14E+05	7.85E-03
1-67	18.6	11.6	2.5	1.3	7.36E+04	8.86E-03
1-2	1.8	1.9	1.7	1.7	1.58E+03	2.72E-02
1-5	1.5	1.6	1.2	1.4	2.15E+03	5.40E-02

1-68	6.7	5.8	3.9	1.0	2.46E+06	8.75E-02
1-69	1.1	1.5	0.9	1.2	3.08E+05	1.23E-01
1-4	1.0	1.4	0.9	1.1	4.20E+04	1.24E-01
1-12	1.7	1.7	1.3	1.7	3.59E+05	1.34E-01
1-11	4.8	4.8	0.9	1.0	5.92E+05	1.43E-01
1-10	17.4	17.4	1.3	1.1	1.04E+06	1.47E-01
1-7	4.5	2.7	1.3	1.2	2.09E+05	1.55E-01
1-15	2.5	1.4	1.0	0.8	4.49E+05	1.56E-01
1-25	3.6	2.7	1.0	0.8	1.98E+06	1.56E-01
1-23	5.1	3.6	1.6	1.6	5.06E+05	1.59E-01
1-7	6.8	3.3	1.7	1.6	1.34E+06	1.70E-01
1-47	8.1	5.5	1.2	0.9	6.32E+03	2.40E-01
1-3	1.6	1.8	1.5	1.5	1.42E+06	3.04E-01
1-32	11.5	1.5	4.0	1.1	5.49E+06	3.17E-01
1-20	5.7	3.9	1.0	0.8	5.17E+02	6.95E-01
1-28	1.0	1.1	1.1	1.2	1.23E+08	9.05E+00
1-48	2.5	2.1	0.8	0.8	5.56E+07	9.11E+00
1-24	12.9	11.0	1.1	1.3	1.94E+08	1.46E+01
1-6	1.8	1.3	1.1	0.8	4.49E+08	2.08E+01
1-17	25.1	20.1	1.6	1.3	3.92E+08	2.63E+01
1-49	1.0	1.0	0.8	0.8	6.00E+08	3.22E+01
1-71	11.6	2.0	5.4	1.4	3.69E+08	4.03E+01
1-42	22.7	20.8	1.4	1.0	9.10E+08	1.03E+02

Table 6C. SARS-CoV-2 Carterra SPR Kinetics Ranked by K_D

IgG	Acro S1	S1-RBD-Fc	S1-Fc	S1-RBM-Fc	k_a (M ⁻¹ s ⁻¹)	k_d (s ⁻¹)	K_D (nM)	Rmax (RU)
2-10	38.9	39.9	9.4	1.2	2.08E+05	1.15E-04	0.6	20
2-5	38.3	35.9	24.3	1.1	9.39E+04	2.59E-04	2.8	39
1-13	4.8	4.3	0.9	1.2	2.17E+05	1.38E-03	6.4	23
2-2	39.4	39.6	4.7	1.1	7.74E+04	9.06E-04	11.7	131
1-27	1.5	1.9	1.3	1.6	2.27E+05	2.85E-03	12.5	39
1-68	6.7	5.8	3.9	1.0	2.46E+06	8.75E-02	35.6	73
1-30	16.8	1.7	5.6	1.4	4.39E+04	1.64E-03	37.2	98
1-31	30.4	1.3	15.7	1.2	2.05E+04	7.62E-04	37.2	256
1-6	1.8	1.3	1.1	0.8	4.49E+08	2.08E+01	46.3	87
1-49	1.0	1.0	0.8	0.8	6.00E+08	3.22E+01	53.6	72
1-32	11.5	1.5	4.0	1.1	5.49E+06	3.17E-01	57.8	112
1-17	25.1	20.1	1.6	1.3	3.92E+08	2.63E+01	67.1	46
1-36	10.1	1.2	1.1	0.8	1.14E+05	7.85E-03	68.9	150

1-28	1.0	1.1	1.1	1.2	1.23E+08	9.05E+00	73.5	58
1-24	12.9	11.0	1.1	1.3	1.94E+08	1.46E+01	75.5	48
1-25	3.6	2.7	1.0	0.8	1.98E+06	1.56E-01	78.9	46
1-35	21.6	1.8	1.4	1.6	5.90E+04	5.43E-03	91.9	341
1-71	11.6	2.0	5.4	1.4	3.69E+08	4.03E+01	109.3	113
1-42	22.7	20.8	1.4	1.0	9.10E+08	1.03E+02	113.4	68
1-67	18.6	11.6	2.5	1.3	7.36E+04	8.86E-03	120.4	22
1-7	6.8	3.3	1.7	1.6	1.34E+06	1.70E-01	127.1	31
1-10	17.4	17.4	1.3	1.1	1.04E+06	1.47E-01	141.7	27
1-48	2.5	2.1	0.8	0.8	5.56E+07	9.11E+00	163.8	64
1-3	1.6	1.8	1.5	1.5	1.42E+06	3.04E-01	214.6	87
1-11	4.8	4.8	0.9	1.0	5.92E+05	1.43E-01	240.9	53
1-23	5.1	3.6	1.6	1.6	5.06E+05	1.59E-01	314.6	128
1-15	2.5	1.4	1.0	0.8	4.49E+05	1.56E-01	346.9	79
1-12	1.7	1.7	1.3	1.7	3.59E+05	1.34E-01	372.5	112
1-69	1.1	1.5	0.9	1.2	3.08E+05	1.23E-01	398.4	66
1-7	4.5	2.7	1.3	1.2	2.09E+05	1.55E-01	742.5	160
1-4	1.0	1.4	0.9	1.1	4.20E+04	1.24E-01	2946.4	385
1-29	1.1	0.9	1.0	0.9	5.35E+02	1.97E-03	3684.3	1206
1-2	1.8	1.9	1.7	1.7	1.58E+03	2.72E-02	17228.5	1652
1-5	1.5	1.6	1.2	1.4	2.15E+03	5.40E-02	25170.1	4457
1-47	8.1	5.5	1.2	0.9	6.32E+03	2.40E-01	37971.0	5497
1-20	5.7	3.9	1.0	0.8	5.17E+02	6.95E-01	1344113.2	64406

[00237] Example 5. SARS-CoV-2 and ACE Variants

[00238] SARS-CoV-2 and ACE variant antibodies were tested for specificity and affinity.

[00239] Recombinant S1 Protein (Acros Biosystems Cat. No. S1N-S52H5) was passively immobilized on a 384 well ELISA plate and blocked with BSA. The S1 Panel antibodies were diluted from 50 nM to 0.0076 nM and incubated with the blocked plate. Antibody binding was detected using Goat-anti-Human-HRP secondary and developed with HRP substrate (list here). The absorbance signal was plotted as % of maximal binding and fitted to determine the EC₅₀ of each antibody using GraphPad Prism.

[00240] Exemplary data for affinity of SARS-CoV-2 variant 2-6 is seen in **FIGS 11A-11B**. The binding of SARS-CoV-2 panel of antibodies was measured as seen in **FIG. 12** and **Tables 7A-7F** below.

Table 7A. SARS-CoV-2 Variants EC₅₀

Antibody	EC ₅₀ (nM)
1-31	0.03139
2-6	0.03364
ACRO	0.04831
1-34	0.06522
2-2	0.07992
1-27	0.09283
2-8	0.1029
1-22	0.1248
1-32	0.1406
1-16	0.1435

1-12	0.1585
2-5	0.1615
CR3022	0.1657
1-53	0.1691
1-30	0.2084
1-28	0.2224
1-71	0.2673
1-20	0.3236
1-4	0.4216
1-35	0.4922
1-47	0.5893
1-5	0.774
2-4	0.8792
1-3	0.9724
1-21	1.003
2-19	1.257
1-51	1.465
1-19	1.706
1-42	1.742
2-2	1.789
2-1	1.894
1-33	3.006
2-13	5.139
2-11	6.921
2-15	8.509
2-7	10.09
1-26	11.93
1-24	12.86
1-49	13.04
1-10	18.31
1-1	21.87
1-8	25.09
1-7	26.94
1-72	29.13
1-17	33.17
1-36	34.86
1-73	43.58
2-10	46.43
1-9	46.88
2-17	51.86
1-52	57.88
2-18	74.71

1-29	83.41
2-12	95.94
1-25	107
2-9	118.3
1-23	123.9
1-48	296
2-14	854.7

Table 7B. SARS-CoV-2 Variants Frequency and ELISA Data

IgG	Freq.	ELISA (Avg)
1-21	1	47.6
1-22	3	40.6
1-30	1	29.5
1-35	3	28.6
1-17	79	27.4
1-27	1	27.2
1-12	2	25.7
1-2	14	24.2
1-3	2	23.9
1-23	1	23.1
1-7	4	22.5
1-31	1	20.7
1-4	1	20.6
1-8	2	19.3
1-9	2	18.5
1-32	1	17.9
1-33	1	17.6
1-24	1	32.9
1-5	1	20.6
1-28	2	17.7
1-10	1	16.9
1-26	1	19.5
1-13	1	28.2
1-14	1	21.7
1-6	1	20.8
1-20	3	26.0
1-29	1	23.8
1-1	1	32.5
1-19	1	22.3
1-16	1	22.1
1-34	1	28.7
1-18	1	22.7
1-11	1	33.2
1-25	1	27.9
1-36	1	23.6
1-15	1	24.8
1-51	1	7.1
1-52	1	3.5
1-53	1	21.7

Table 7C. SARS-CoV-2 S1 Variants

			DB/S1	DB/S-T	DC/S1	DC/S-T	Inhibitor	Fc/S1	Fc/S-T
IgG	Freq	ELISA (Avg)	K _D (nM)	K _D (nM)	K _D (nM)	K _D (nM)	IC50 (nM)	K _D (nM)	K _D (nM)f
1-21	1	47.6					6.7		
1-22	3	40.6					73.2		
1-30	1	29.5	37.2		4.7	475861.6	25.3		
1-35	3	28.6	91.9	569.8	47.5	16.7	209.2	139.4	0.5
1-17	79	27.4	67.1		6649.2				10.0
1-27	1	27.2	12.5		5519.6			9.6	9.6
1-12	2	25.7	372.5		10.6		33.0	6.1	14.0
1-2	14	24.2	17228.5	304.3					
1-3	2	23.9	214.6	423.0	252.8	5306.9	15.7		10.0
1-23	1	23.1	314.6						
1-7	4	22.5	127.1						
1-31	1	20.7	37.2		14.4		9.2	17.5	
1-4	1	20.6	2946.4	6.6	659.2	129.2		25.5	12.3
1-8	2	19.3					97.3		
1-9	2	18.5					282.4		
1-32	1	17.9	57.8		14.8			2979.8	
1-33	1	17.6			8.1	1739312.4			
1-24	1	32.9	75.5		2043.7				
1-40	2	22.9							
1-5	1	20.6	25170.1	1155.7					
1-28	2	17.7	73.5	13.4	520.3	96.6		3236.7	8.6
1-10	1	16.9	141.7				17.7		
1-26	1	19.5					8.0		
1-13	1	28.2	6.4						
1-14	1	21.7		8.1					
1-6	1	20.8	46.3	684.2					
1-20	3	26.0	1344113.2		34.4	145.0		10.3	17.3
1-29	1	23.8	3684.3	36.1					19.3
1-1	1	32.5							
1-19	1	22.3			85.4	0.0		14.3	18.6
1-16	1	22.1			2282.0	4487.3	1.7	178.8	2.2
1-34	1	28.7			8.2	623.6		13.4	1.4
1-18	1	22.7							
1-11	1	33.2	240.9	30.2					
1-25	1	27.9	78.9				3.0		6.1
1-36	1	23.6	68.9	3.9			9.3		33.2
1-15	1	24.8	346.9						
1-51	1	7.1							
1-52	1	3.5					739.8		
1-53	1	21.7					1426.0		
2-16	1	7.1					433.4		
2-17	1	3.5							
2-18	1	43.0							

2-19	1	21.7							
2-2	1	12.8							

Table 7D. SARS-CoV-2 S1 Variants Frequency and ELISA Data

IgG	Freq.	ELISA(Avg)
2-1	1	45.3
2-10	1	43.8
2-5	46	30.8
2-2	2	23.4
2-4	1	15.5
2-6	5	22.1
2-11	3	28.1
2-12	1	18.2
2-13	1	19.1
2-14	1	17.2
2-7	1	25.3
2-8	1	32.4
2-15	1	13.7
2-9	1	14.1
2-16	1	7.1
2-17	1	3.5
2-18	1	43.0
2-19	1	21.7
2-2	1	12.8

Table 7E. SARS-CoV-2 S1 Variants

			DB/S1	DB/S-T	DC/S1	DC/S-T	Inhibitor	Fc/S1	Fc/S-T
IgG	Freq	ELISA (Avg)	K_D (nM)	K_D (nM)	K_D (nM)	K_D (nM)	IC50 (nM)	K_D (nM)	K_D (nM)
2-10	1	43.8	0.6	125.6					
2-5	46	30.8	2.8			17.3	4.2	1.9	90.2
2-2	2	23.4	11.7	1.2		1.4	3.6	58.3	0.8
2-4	1	15.5		4337.7					
2-6	5	22.1		3.0	563.8	0.4		32.3	0.01
2-11	3	28.1		3.1				90.0	284.5
2-12	1	18.2		6.4		63.8	1.0		3.2
2-13	1	19.1					45.0		
2-14	1	17.2		2.5					34.8
2-7	1	25.3							252.5
2-8	1	32.4		115.1	33.4	47.4	12.2	52.8	
2-15	1	13.7		3.5			4.7		
2-9	1	14.1					23.2		23582.5

Table 7F. Antibody Panel ELISA Binding Titrations (EC50)

ANTIBODY	EC50 (nM)
2-8	0.08001
1-35	0.09604
1-3	0.133
2-5	0.1332
1-27	0.1479
1-31	0.2035
2-6	0.283
1-2	0.523
1-34	0.5584
2-2	0.612
1-67	0.9402
1-16	1.409
1-12	2.15
1-28	2.284
1-4	2.559
1-1	4.157
1-19	4.413
1-22	6.548
1-5	7.833
1-42	7.92
2-15	7.92
1-49	8.669
1-53	10.25
2-9	12.58
1-33	17.5
1-26	48.46
2-29	63.43
1-7	95.66
1-25	95.66
1-51	98.17
2-17	100
2-2	100

[00241] Data for competition ELISA for a first set of SARS-CoV-2 S1 RBD and ACE2 variant antibodies is seen in FIGS. 13A-13B and FIG. 14A. SARS-CoV-2 variant antibodies with high potency in order of

potency included variant 2-2, Acro mAb (1.5333), variant 2-5, variant 1-12, and variant 2-9. ACE variant antibodies with high potency in order of potency included variant 4-52, variant 4-17, variant 4-39, Acro mAb (1.533), variant 4-54, and variant 3.5. Data for competition ELISA for a second set of ACE2 variant antibodies is seen in **FIG. 14B**. Variant antibodies with high potency in order of potency included variant 4-101, variant 4-140, variant 4-121, variant 4-118, and Acro mAb (2.76 nM). **FIGS. 14C-14D** show the SARS-CoV-2 variant antibodies show potent neutralization.

[00242] Inhibition assays were also performed as seen in **FIG. 15**.

[00243] SARS-CoV-2 variant antibodies were assayed for Vero inhibition using FACS. Briefly, Vero cells stripped with Cell Stripper (~20 minutes with 90% viability after removal). Cells were plated at 0.1×10^6 cells per well. Stock solution of the variant antibodies were at 200nM titrated 1:3. SARS-CoV-2 S protein RBD, SPD-C5259 were made up at 6ug/mL. Variant antibody titrations were mixed 1:1 with 6ug/mL S protein (50uL IgG: 50 uL S protein). 100 uL of the mixture were added to cells and then incubated on ice for 1 hour. The cells were washed 1X followed by addition of 50 uL of goat anti-mouse secondary made up at 1:200. The cells were then incubated on ice for 1 hour in the dark, washed three times, and the plates were then read. Data for SARS-CoV-2 variant antibodies is seen in **FIGS. 16A-16D**, **FIGS. 17A-17D**, and **Tables 8 and 9**. As seen in the data, several variant antibodies blocked labeled S1 RBD from binding to ACE2 on the Vero cells including variants 2-8, 2-5, 2-2, 2-4, and 1-63.

Table 8.

Antibody	IC50 (nM)
Acro Anti-S1	2.7
1-30	NC
1-35	NC
1-12	NC
1-31	NC
1-63	106.6
2-5	4.4
2-2	3.0
2-4	46.3
2-6	NC
2-8	19.5

Table 9.

	S1 Monomer (nM)	S Trimer (nM)
1-31	0.22	0.80
1-30	0.67	4.02
1-35	0.15	0.76
1-12	2.08	0.61
1-63	1.40	14.39
2-8	1.52	7.08

2-5	0.17	0.59
2-2	0.13	0.64
2-4	1.58	10.18
2-6	0.07	0.43

[00244] A summary of epitope binning for SARS-CoV-2 variant antibodies is seen in **Table 10** below.

Table 10. SARS-CoV-2 Epitope Binning

ID	Acro mAb	2-2	Abcam CR3022	2-5	2-8	2-11	1-32	1-16	2-6	1-35
*Acro mAb	0	0	2	1	1	1	1	1	1	1
*2-2	0	0	0	0	0	0	2	0	2	1
Abcam CR3022	2	0	0	0	0	1	1	2	2	2
*2-5	1	0	0	0	0	1	1	1	1	1
*2-8	1	0	0	0	0	1	1	1	1	1
2-11	1	0	0	0	0	0	1	2	1	1
**1-32	2	1	3	1	2	1	0	0	1	1
1-16	1	1	2	2	1	1	1	0	0	0
2-6	2	2	3	2	2	2	1	0	0	0
1-35	1	1	2	2	1	1	1	0	0	0

* Anti-S1 inhibiting IgG in FACS (Vero E6)

** Anti-S1 inhibiting IgG in ELISA (soluble ACE2)

[00245] The variant antibodies were also measured in binding against other coronaviruses. Data shows that the variant antibodies do not bind significantly to S1 HCoV-229E (Sino), S1 HCoV-HKU1 (Sino), S1 HCoV-NL63 (Sino), or S1 HCoV-OC43 (Sino) (data not shown).

[00246] The data shows that the SARS-CoV-2 and ACE2 variant antibodies have high specificity and affinity to their antigen targets with affinities in the picomolar to nanomolar range.

[00247] Example 6. SARS-CoV-2 S1 Variants

[00248] VHH-Fc antibodies targeting S1 were titrated 1:3 starting at 200 nM and mixed 1:1 with SARS-COV-2-S1 RBD (mouse IgG2Fc tag). The RBD/VHH-Fc complex was added to Vero E6 cells expressing endogenous ACE2 receptor and incubated. Cells were subsequently washed and an anti-mouse secondary was used to measure binding of S1 RBD to ACE2, thus assessing the inhibition of S1. Over 60 clones demonstrated potent inhibition. Data is seen in **FIGS. 18A-18C** and **Tables 11 and 12**.

Table 11.

Sample	IC ₅₀ [nM]
2-2	0.56
5-56	0.68
5-1	0.75
5-67	0.76
5-47	0.80
5-8	0.94
5-38	0.96
5-37	1.01
5-34	1.21
5-20	1.23

5-55	1.45
5-46	1.52
5-50	1.61
5-5	1.79
2-5	2.146
5-60	2.15
5-15	2.19
5-29	2.19
Acro Anti S1	3.80
5-49	5.61
5-44	11.73

Table 12.

Sample	IC ₅₀ [nM]
6-85	0.2044
2-2	0.56
6-63	0.74
6-3	0.74
6-78	0.7427
6-20	0.76
6-91	0.89
6-44	0.97
6-55	0.97
6-73	1.01
6-26	1.07
6-76	1.11
6-45	1.16
6-60	1.31
6-40	1.36
6-81	1.383
6-10	1.44
6-7	1.53
6-39	1.53
6-109	1.60
6-38	1.94
2-5	2.146
6-30	2.94
6-57	3.13
Acro Anti S1	3.49
6-67	3.80
6-77	4.041
6-100	5.07
6-47	5.86
6-41	6.60
6-88	7.118
6-105	7.82
6-34	8.24
6-54	8.90
6-18	12.29
6-1	15.76
6-65	37.47

[00249] Antibody kinetics were measured for variants 2-5, 2-2, and 2-6 (FIG. 19A) and variants 1-12, 1-42, 1-20, and 1-19 (FIG. 19B). Data is seen in FIGS. 19A-19B. The data shows that the antibodies bind with nanomolar affinities. FIG. 19C shows percent neutralization for variants 1-12, 1-42 and 1-20. FIG. 19D shows percent neutralization for variants 1-12, 1-42 and 1-20 using live virus.

[00250] Example 7. Neutralization of Live Virus

[00251] VHH-Fc antibodies targeting S1 were titrated 1:3 starting at 200 nM and mixed 1:1 with SARS-CoV2-S1 RBD (mouse IgG2Fc tag). The RBD/VHH-Fc complex was added to Vero E6 cells expressing endogenous ACE2 receptor and incubated. Cells were subsequently washed and an anti-mouse secondary was used to measure binding of S1 RBD to ACE2, thus assessing the inhibition of S1. Over 60 clones demonstrated potent inhibition. The data is seen in FIGS. 20A-20B and Table 13.

Table 13.

Antibody	EC50 (ug/mL)
6-63	0.06
6-3	0.06
2165mAb*	0.08
5-1	0.10
6-60	0.15
6-55	0.21
5-20	0.27
1-20	0.37
5-34	0.54
6-85	0.84
6-76	1.08
6-73	1.46
1-42	2.03
1-12	2.10
6-26	2.97
6-20	5.03
6-78	8.26
2-6	11.77
2-5	18.31
2-2	67.57
1-63	106.90

[00252] FIG. 20B shows that variant 6-2 showed higher neutralization versus IgG in live virus. Variants 6-63, 6-3, and 5-1 showed comparable neutralization versus 2165 mAb derived from a COVID-19 subject.

[00253] FIGS. 20C-20E and Table 14 show data from VHH single domain antibodies in VSV-pseudotype SARS-CoV-2 neutralization assays. Variants 5-1, 6-3, and 6-63 showed improved neutralization in pseudovirus testing including in live virus FRNT (FIG. 20E). Variants 6-3, 6-60, 6-63, and 6-76 showed potent neutralization in live virus PRNT as seen in Table 15 and FIG. 20F. Data as seen in Table 16 and FIGS. 20G-20H show that variants 6-3, 6-63, and 1-20 exhibited potent neutralization in live virus PRNT.

Table 14.

Antibody	NC50 (ug/mL)
6-63	0.06
6-3	0.06
5-1	0.10
6-60	0.15
6-55	0.21
5-20	0.27
1-20	0.37
5-34	0.54
6-85	0.84
6-76	1.08
6-73	1.46
1-42	2.03
1-12	2.10
6-26	2.97
6-20	5.03
6-78	8.26
2-6	11.77
2-5	18.31
2-2	67.57
1-63	106.90

Table 15.

Antibody	PRNT90 (ng/mL)
5-1	15.6
5-20	3.9
5-34	15.6
6-26	62.5
6-60	<0.98
6-63	3.9
6-3	<0.98
6-55	62.5
6-76	3.9
6-78	250
6-20	15.6
6-73	250

6-85	62.5
------	------

Table 16.

Antibody	EC80 (ug/mL)
6-63	0.057861
6-3	0.115234
1-20	0.171875
6-60	0.236816
5-20	0.376
6-76	0.425781
6-55	0.445313
5-34	0.570313
6-42	0.734375
5-1	0.810547
1-12	0.855469
6-85	1.0625
5-38	1.5
5-67	1.566406
1-42	1.78125
6-73	2.015625
6-20	2.3125
5-47	2.457031
1-19	2.78125
6-44	3.15625
6-26	3.609375
6-45	4.015625
5-37	12.4375
6-78	13.75
5-63	14.1875
5-8	15.8125
6-6	24.375
6-13	25.6875
6-24	31.375
1-63	50.5
6-32	60.625
2-6	72.34043
6-22	103.25
5-56	104.75
2-5	106.5
6-82	107.75
5-32	180.1418
6-91	240.5
2-2	354.6099

5-51	387.9433
------	----------

[00254] The variants were also tested in pseudovirus neutralization and live virus PRNT studies. Variants 5-20, 6-60, 6-63, and 6-3 showed potent neutralization in live virus PRNT as seen in FIG. 20I.

[00255] Example 8. *In Vivo* Evaluation of Variant Coronavirus Immunoglobulins

[00256] This Example assesses the variant coronavirus immunoglobulins in a Syrian hamster model (immunosuppressed) of COVID-19 disease.

[00257] 8-10 week-old female Syrian hamsters were immunosuppressed using cyclophosphamide (140 mg/kg day 3 days before challenge and then 100 mg/kg every 4 days by i.p. route). Eleven groups of six hamsters per group were injected with antibody on day -1 relative to challenge by the intraperitoneal route (i.p.). On Day 0 all hamster were challenged with 1,000 PFU SARS CoV-2 Washington isolate by the intranasal route and weighed daily. % Weight change relative to starting weight was calculated. Pharyngeal swabs were collected on Days -1, 1, 4, 7, 9. Day 9 lungs were collected and homogenized for viral load. Groups are shown in Table 17.

Table 17. Groups

Group	Diluent/volume injected i.p.
Convalescent plasma	NA/2.5 mL
Negative control MAb c7d11	PBS/2.5 mL
6-63	PBS/2.5 mL
6-3	PBS/2.5 mL
6-36	PBS/2.5 mL

NA= not applicable

[00258] Animals injected intraperitoneally (i.p.) with the Negative Control antibody lost weight starting losing significant amounts of weight between Days 5 and 6 and continued to decline until the end of the experiment on Day 9. The maximum mean weight loss of the group was -11.7%. In contrast, animals injected with positive control human convalescent plasma maintained weight within -3.2% of their weight on Day 0 indicating this plasma protected against disease manifested by weight loss (FIG. 21A).

[00259] Groups of six animals were injected i.p. with 10, 5, or 1 mg/kg of monoclonal antibody 6-63 diluted in PBS. All groups maintained their weight at or above starting weight indicating the antibody protected against disease resulting weight loss (FIG. 21B).

[00260] Groups of six animals were injected i.p. with 10, 5, or 1 mg/kg of monoclonal antibody 6-3 diluted in PBS. The 1 and 10 mg/kg groups maintained their weight at or above starting weight at all time points. The 5 mg/kg group weight dipped slightly below the convalescent control on Days 7-9 but clearly was different from the Negative Control antibody. Together, these data indicate the antibody decreased weight loss associated with disease (FIG. 21C).

[00261] Groups of six animals were injected i.p. with 10, 5, or 1 mg/kg of monoclonal antibody 6-36 diluted in PBS. The 10 and 5 mg/kg groups maintained their weight at levels similar to the positive control at all time points. The 1 mg/kg group weight dropped significantly similar to the Negative Control. These data indicate antibody 6-36 at 1 mg/kg is insufficient to provide benefit, but a 5-fold or greater dose is adequate to reduce disease as determined by weight loss (FIG. 21D).

[00262] FIG. 3421E shows data from the variant antibodies grouped by dose. FIG. 21F shows graphs of percent weight change for antibodies 6-3, 6-63, and 1-20.

[00263] In wild type hamsters, virus is typically cleared by Day 7. However, in the cyclophosphamide model, viral levels are not suppressed unless there is intervention (e.g. protective antibodies administered) or the cyclophosphamide is discontinued to allow immune response and clearance. In this experiment the positive control human convalescent serum eliminated virus from the lungs from all except one hamster. In contrast, all but one of the hamsters injected with negative control antibody still had infectious virus in the lungs. Interestingly, hamsters prophylactically treated (24 hour previous to exposure) with any of the three antibodies at the highest dose (10 mg/kg) had infectious virus in the lungs of at least half the animals assessed 9 days later. Paradoxically, 6-63 and 6-3 at the lower doses (5 and 1 mg/kg) had animals with relatively less infectious virus in the lungs. 4 of 6 animals injected with 1-20 at 5 mg/kg dose animals had no detectable virus in the lungs. When the doses of that antibody was reduced to 1 mg/kg, all but one animal had infectious virus. Data is seen in FIGS. 21G-21H.

[00264] Lung pathology inflammation and edema scores from three animals were added per group and plotted (Fig. 21I). These were the same lungs used to score ISH. The convalescent sera positive control median score was 2 and the negative control was 4. The only groups with a median score lower than the negative control group were 1 and 5 mg/kg 6-63, 10 and 5 mg/kg 6-3, and 5 mg/kg 1-20. The highest median pathology scores were the 1 and 10 mg/kg 1-20 groups. The lowest median pathology score was the 5 mg/kg 6-63 group.

[00265] Example 9. VSV- pseudotype neutralization analysis of antibodies for SARS-CoV-2 B.135 (South African strain)

[00266] Antibodies described herein were tested in a VSV-pseudotype neutralization assay for SARS-CoV-2 B.135 (South African strain).

[00267] Briefly, aerial semi-log dilutions of all test antibodies (TA) and control were prepared and mixed with the VSV-pseudotype virus in a 1:1 ratio for 1 h at RT followed by incubation over Vero cells (ATCC® CCL-81TM) seeded at 60,000 cells per well at 37°C. The cells were lysed the following day and luciferase activity was measured to assess the potency of each TA to block viral entry into the Vero cells. All samples will be run in triplicate. Data analysis is conducted using XLFit and Graphpad Prism. The testing concentrations and plate plan are seen in Table 18A.

Table 18A. Testing Concentrations and Plate Plan

Samples	Stock (mg/mL)	Target conc/dilution	In plate concentration
---------	---------------	----------------------	------------------------

1	6-63	6.39	100 ug/mL	50.00 ug/mL
2	6-3	10.05	100 ug/mL	50.00 ug/mL
3	1-12	2.24	100 ug/mL	50.00 ug/mL
4	mouse pAb	1	1: 25	1: 50

[00268] Data is seen in **FIGS. 22A-22B**. **FIG. 22A** illustrates positive control pAb has an NT₅₀ of 1: 14, 993 dilution as expected. **FIG. 22B** illustrates antibodies 6-63 and 6-3 neutralize VSV-SARS B.135 strain with IC₅₀s of ~3.07 ug/mL and 0.143 ug/mL, respectively. Antibody 1-12 failed to neutralize VSV-SARS B.135 strain.

[00269] Example 10. VSV-pseudotype neutralization analysis of antibodies for SARS-CoV-2 D614G variant

[00270] Antibodies described herein were tested in a VSV-pseudotype neutralization assay for SARS-CoV-2 SARS CoV-2 S D614G variant.

[00271] Briefly, serial semi-log dilutions of all test antibodies (TA) and control were prepared and mixed with the VSV-pseudotype virus in a 1:1 ratio for 1 h at RT followed by incubation over Vero cells (ATCC® CCL-81TM) seeded at 60,000 cells per well at 37°C. The cells were lysed the following day and luciferase activity was measured to assess the potency of each TA to block viral entry into the Vero cells. All samples will be run in triplicate. Data analysis is conducted using XLFit and Graphpad Prism.

[00272] Data is seen in **FIGS. 23A-23C**. **FIG. 23A** shows the positive control.

[00273] Example 11. Antibody cocktails for treating SARS-CoV-2 in Syrian hamsters

[00274] This Example demonstrates pre- and post-exposure efficacy of antibody cocktails in Syrian hamsters.

[00275] Methods

[00276] In this study the hamsters were transiently immunosuppressed using cyclophosphamide. As a strategy to de-risk selecting viruses with neutralization escape mutations a cocktail of a nanobody (nAb) and a monoclonal antibody (MAb) known to bind different spike protein epitopes were combined and used. The combined dose was 20 mg/Kg in this proof-of-concept experiment. The cocktail consisted of 10mg/Kg of VHH nanobody 6-63 and 10mg/Kg of monoclonal antibody 1-20. An equal number of male and female animals were used in each group.

[00277] Seventy-eight hamsters were used for this experiment according to **Table 18B**. On Day 0, animals were exposed via intranasal (IN) instillation to 1,000 pfu of SARS-CoV-2 virus in 50 µL volume. The volume was distributed between both nares. To transiently immunosuppress, all animals were treated with cyclophosphamide starting on Day -3 (140 mg/kg dose) followed by additional doses (100 mg/kg) on Days 1, 5, and 9.

[00278] On the indicated day post exposure, MAb/nAb cocktail or c7D11 was administered via the intraperitoneal (IP) route. On Day 0 blood samples were collected from Group I for hematology to confirm immunosuppression. Group I was also the control for any adverse effects of cyclophosphamide treatment on the hamsters. Clinical scores and individual animal weights were recorded daily. Pharyngeal swabs and

other key events are listed in **Table 18B**. Animals in Groups A-I were euthanized on day 14 and lungs were collected for virology and pathology. Group J animals were used for a serial pathology component of this study. Two animals from Group J (2 male and 2 female) were euthanized starting on Day 1 and then each day up to and including Day 6.

Table 18B. Experimental design

Group	Number of Hamsters (Pain Category)	Virus Exposure ^a (Day 0)	Treatment	Treatment Day
A	6 (3 male, 3 female) (D)	SARS CoV-2	Cocktail ^b 20 mg/Kg	-1
B	6 (3 male, 3 female) (D)		Cocktail ^b 20 mg/Kg	+1
C	6 (3 male, 3 female) (D)		Cocktail ^b 20 mg/Kg	+2
D	6 (3 male, 3 female) (D)		Cocktail ^b 20 mg/Kg	+3
E	6 (3 male, 3 female) (D)		Cocktail ^b 20 mg/Kg	+4
F	6 (3 male, 3 female) (D)		Cocktail ^b 20 mg/Kg	+5
G	6 (3 male, 3 female) (E)		Cocktail ^b 20 mg/Kg	+6
H	6 (3 male, 3 female) (E)		Neg IgG control	+1
I	6 (3 male, 3 female)(C)	No virus	none	Cyclophosphamide control
J	24 (12 male, 12 female) (E)	SARS CoV-2	none	No treatment-pathology control ^d

78 Syrian hamsters

^a challenge with 1,000 pfu of virus in 50 microliter volume

^b cocktail= 6-63 combined with 1-20 1:1 w/v delivered 2.5 mL per animal by i.p. route

^c negative control 20mg/kg

^d Two animals (2 male and 2 female) from Group J were euthanized for pathology/virology on Days 1, 2, 3, 4, 5, 6

[00279] Results

[00280] *Cyclophosphamide treatment in uninfected animals does not result in weight loss.* Control animals (CYP Controls, Group I), that were treated with CYP but not challenged gained weight overtime.

[00281] *Negative control antibody c7D11 at 20 mg/kg, Group H, does not protect against disease associated weight loss.* Hamsters in Group H lost weight starting on Day 6 (**FIG. 24A**). Weight loss continued until Day 10 when it leveled off. Animals were still below 10% of their starting weight on the last day of the experiment (Day 14).

[00282] *The cocktail administered one day prior to exposure protected against weight loss.* Hamsters in Group A maintained their weight and stayed within 1% of starting weight (**FIG. 24A**). This confirmed that

treatment with neutralizing antibodies before exposure was sufficient to protect against significant weight loss.

[00283] *Post-exposure treatment of CYP hamster model of COVID19 produces variable weight loss effects/patterns.* A onetime treatment of a cocktail containing 1-20 and 6-63 at final dose of 20mg/Kg was administered on Days 1 (B), 2 (C), 3 (D), 4 (E), 5 (F), and 6 (G). The percent weight change relative to Day 0 are shown in **FIG. 24A**. Arrows and dotted vertical lines indicate the day of treatment specific for the treatment regimen being compared. The same data is shown collectively in **FIG. 24B**. Note that there were two animals in Group B (Day 1) that dropped weight atypically during the experiment and one animal succumbed on Day 12. That animal had a necrotic/hemorrhaging testicle due to an apparent torsion event and was excluded from analysis. No assignable cause was identified for the second animal so that animal was not excluded from analysis. Statistical analysis was performed to compare both the CYP Control (Group I) and the negative control antibody (c7d11, Group H) to all other groups. The significance between groups at individual timepoints and differences in area under the curve (AOC) were determined. Treatment with the cocktail one day after exposure (Group B) results were confounded by the outlier animal. There was not a significant differences in AOC between Group B and Group H suggesting no protection; however, there was also no significant difference in the AOC between Group A and Group I indicating Group B weight loss was not significantly different from the CYP control group that was not exposed. Treatment with the cocktail two days after exposure (Group C) clearly protected. There was significant differences in AOC between Group C and Group H; and no significant difference in the AOC between Group C and Group I. Treatment after three days (Day 3 Group D) did not result in a significant difference in AOC between Group D and H. However, treatment on day 4 or 5 after exposure (Groups E and F, respectively) did significant reduce the AOC relative to Group H. Treatment on day 6 (Group G) was similar to treatment on day 3 where no significant difference in AOR between Group G and H. Although there was no significance in the AOC for the Day 6 treatment group, the last three timepoints weight loss was significantly less than the negative control group. Interestingly, groups administered antibody on Days 3, 4, 5, or 6 started to gain weight starting on Day 9 whereas the negative control antibody treated animals did not. This suggests that there was a benefit of all cocktail treatment even at as late as 6 days post-exposure.

[00284] *Infectious virus in lungs (Day 14/15).* In wild type hamsters, virus is typically cleared by Day 7. However, in the cyclophosphamide model virus is not suppressed unless there is intervention (e.g. protective antibodies administered) or the cyclophosphamide is discontinued to allow immune response and clearance. Here, our controls demonstrate that unexposed hamsters were negative for virus (Cyp Cont), whereas all hamsters exposed to virus and treated with an off-target monoclonal antibody (Neg Cont) had more than 10,000 pfu of virus per gram of lung tissue. Most of the hamsters treated with the Cocktail one day prior or one day post virus exposure had detectable levels of virus in lung samples collected on Day 14 (Groups A and B). However, almost all of the hamsters treated with antibody ≥ 2 day after exposure had undetectable levels of antibody in their lungs. There was only a single animal exception in the Day 2, 3, 4 and 6 treated groups. All of the hamsters treated on Day 5 had lungs that were free of infectious virus. *See FIG. 24C.*

[00285] *Sequential Sampling.* Hamsters immunosuppressed and exposed to virus on Day 0 were sampled overtime to monitor the infection. Infectious virus was detected in the lungs of 3 of 4 hamsters on Day 1. Levels of infectious virus then increased more than 4 logs by Day 2. Levels of virus then leveled off and stayed between 7-9 log₁₀ through the last sampling timepoint (Day 6) analyzed to-date. Thus, animals administered the cocktail after Day 1 likely had very high levels of infectious virus present in their lungs before treatment with the antibody. See FIG. 24D.

[00286] *Conclusion*

[00287] The data demonstrates that when administered at or before Day 2 relative to virus exposure, the combination of antibody and nanobody at 20 mg/kg was sufficient to provide convincing benefit as determined by the disease parameters analyzed. When treated after Day 2, animals still developed weight loss but recovered in approximately 4 days. Two weeks after virus exposure, more infectious virus was detected in the lungs of hamsters receiving antibody early (Day -1 or 1) than day 2 or later. Day 2 (or later) treatment of exposed animals occurred at a time when viral burden in the lungs was already remarkably high. CYP, by itself, does not induce weight loss or clinical signs of disease.

[00288] **Example 12. Immunity conferred by SARS-CoV-2 and ACE2 antibodies**

[00289] Antibodies described in Examples 4-6 are used to confer immunity in a subject. A subject is passively immunized with a SARS-CoV-2 or ACE2 antibody. The subject is then exposed to SARS-CoV-2 after immunization with the SARS-CoV-2 or ACE2 antibody. Exposure can be within a few days or within a few months. The subject can also receive the SARS-CoV-2 or ACE2 antibody immediately following exposure with SARS-CoV-2. Although the subject is exposed to SARS-CoV-2, the subject has developed an immunity against SARS-CoV-2 and infection is prevented.

[00290] **Example 13. Sequences**

[00291] Tables 19-25 show exemplary sequences for CDRH1-H3 and CDRL1-L3 as well as variant heavy chains and variant light chains for the SARS-CoV-2 and ACE2 variants.

Table 19. SARS-CoV-2 S1 Variable Heavy Chain CDRs

Variant	SEQ ID NO	CDRH1	SEQ ID NO	CDRH2	SEQ ID NO	CDRH3
1-1	1	FTFSSHAMY	37	SAISGSAGSTYYA	73	CAHDTKDFWSGYCIFDPW
1-2	2	FTFSSQAMS	38	SAISGSGGGTYA	74	CAKDRRFGEFDPW
1-3	3	FTFSSQAMS	39	SAISGSGGGTYA	75	CAKDRRFGEFDPW
1-4	4	FTFSSQAMS	40	SAISGSGGGTYA	76	CAKDRRFGEFDPW
1-5	5	FTFSSQAMS	41	SAISGSGGGTYA	77	CAKDRRFGEFDPW
1-6	6	FTFSSQAMS	42	SAISGSGGGTYA	78	CAKDRRFGEFDPW
1-7	7	FTFSSYDMS	43	SVISGSGGSTYYA	79	CAKGPLVGWYFDLW
1-8	8	FTFSSYDMS	44	SVISGSGGSTYYA	80	CAKGPLVGWYFDLW
1-9	9	FTFSSYDMS	45	SVISGSGGSTYYA	81	CAKGPLVGWYFDLW

1-10	10	FTFSSYDMS	46	SVISGSGGSTYYA	82	CAKGPLVGWYFDLW
1-11	11	ITFSSYAMS	47	SGISGSGGSTYYA	83	CAKHGSGTIFGVVIAKYYFDYC
1-12	12	ITFSSYAMS	48	SGISGSGGSTYYA	84	CAKHGSGTIFGVVIAKYYFDYW
1-13	13	VTFSSYAMS	49	SAITGSGGSTYYA	85	CAKHGSGTIFGVVIAKYYFDYW
1-14	14	ITFSSYAMS	50	SGISGSGGSTYYA	86	CANHGSGTIFGVVIAKYYFDYW
1-15	15	FTFSSHAMY	51	SAISGSAGSTYYA	87	CARDTKDFWSGYSIFDPW
1-16	16	FTFSSYAMY	52	SAISGSAGSTYYA	88	CARDTNDFW
1-17	17	FTFSSYAMY	53	SAISGSAGSTYYA	89	CARDTNDFWSGYSIFDPW
1-18	18	FTFSSYAMY	54	SAISGSAGSTYYA	90	CARDTNDFWSGYSIFDPW
1-19	19	FTFTSYAMY	55	SAISGSAGSTYYA	91	CARDTNDFWSGYSIFDPW
1-20	20	FTFSSYAMY	56	SAISGSAGSTYYA	92	CARDTNDFWSGYSIFHPW
1-21	21	FTFSSYTMS	57	SIISGSGGSTYYA	93	CAREGYRDYLWYFDLW
1-22	22	FTFSSYAMN	58	SIISGSGGSTYYA	94	CAREGYRDYLWYFDLW
1-23	23	FTFSSY AIS	59	SIISGSGGSTYYA	95	CAREGYRDYLWYFDLW
1-24	24	FTFSSYAMN	60	SIISGSGGSTYYA	96	CAREGYRDYLWYFDLW
1-25	25	FTFSDYAMN	61	SIISGSGGSTYYA	97	CAREGYRDYLWYFDLW
1-26	26	FTFSSY AIS	62	SAISGSGYSTYYA	98	CARGAPDSSGYFQGEVYFDY W
1-27	27	FTFSSYAMT	63	SAISGSGGGTYA	99	CASSSSWQFDYW
1-28	28	FTFSSYGMS	64	SAISGSGGGTYA	100	CASSSSWQFDYW
1-29	29	FTFSSYAMT	65	SAISGSGGGTYA	101	CASSSSWQFDYW
1-30	30	FTFSSYAMT	66	SAISGSGGSTFYA	102	CTRPPYGDYGDYW
1-31	31	FTFSSYAMN	67	SAISGSGGSTFYA	103	CTRPPYGDYGDYW
1-32	32	FTFSSYAMY	68	SAISGSGGSTFYA	104	CTRPPYGDYGDYW
1-33	33	FTFSSYAMI	69	SAISGSGGSTFYA	105	CTRPPYGDYGDYW
1-34	34	FTFSSYAMY	70	SAISSSGGSTYYA	106	CTRPPYGDYGDYW
1-35	35	FTFSNYAMS	71	SDISGSGGSTYYA	107	CVKGTIPFGVIRSAFDYW
1-36	36	LTFSSYAMS	72	SDISGSGGSTYYA	108	CVKGTIPFGVIRSAFDYW

Table 20. SARS-CoV-2 S1 Variable Light Chain CDRs

Variant	SEQ ID NO	CDRL1	SEQ ID NO	CDRL2	SEQ ID NO	CDRL3
1-1	109	TGTSSDIGSYNLVS	145	EGTKRPS	181	CCSYAGSRITYVF
1-2	110	TGTSSGVGSYNLVS	146	EGIKRPS	182	CCSYAGSSSFVVF
1-3	111	TGISSDVGSYNLVS	147	EGTKRPS	183	CCSYAGSSSFVVF
1-4	112	TGTSSDVGTYNLVS	148	EGSQRPS	184	CCSYAGSSSFVVF

1-5	113	TGTSSGVGSYNLVS	149	EASKRPS	185	CCSYAGSYTFVVF
1-6	114	TGTSSGVGSYNLVS	150	EGIKRPS	186	CCSYAGSSSFVVF
1-7	115	TGTSSDFGSYNLVS	151	EGNKRPS	187	CCSYAGSSTFVVF
1-8	116	TGTSDVGSYNLVS	152	EGSKRPS	188	CCSYAGSRYVVF
1-9	117	TGTSSDVGDYNLVS	153	EGSNRPS	189	CCSYAGSSSFVVF
1-10	118	TGTSSNVGSYNLVS	154	EGTKRPS	190	CCSYAGSSSFVVF
1-11	119	TGTSSDVGHYNLVS	155	EGTKRPS	191	CCSYAGSSSFVVF
1-12	120	TGTSSDVGHYNLVS	156	EGTKRPS	192	CCSYAGSSSFVVF
1-13	121	TGTSSDVGRYNLVS	157	EGGKRPS	193	CCSYASSSTLVF
1-14	122	TGTSSDVGHYNLVS	158	EGTKRPS	194	CCSYAGSSSFVVF
1-15	123	TGTSSDIGSYNLVS	159	EGTNRPS	195	CCSYAGSRTYVF
1-16	124	TGTSSDIGSYNLVS	160	EGTKRPS	196	CCSYAGSRTYVF
1-17	125	TGTSSDIGSYNLVS	161	EGTKRPS	197	CCSYAGSRTYVF
1-18	126	TGTSSDIGSYNLVS	162	EGTKRPS	198	CCSYAGSRTYVF
1-19	127	TGTSSDIGSYNLVS	163	EGTKRPS	199	CCSYAGSRTYVF
1-20	128	TGTSSDIGSYNLVS	164	EGTKRPS	200	CCSYAGSRTYVF
1-21	129	TGTSSDVGSNNLVS	165	EGNKRPS	201	CCSYAGSVVF
1-22	130	TGTSTDVGSYNLVS	166	EGSQRPS	202	CCSYAGSSTVF
1-23	131	TGTSDVGSYNLVS	167	EGNKRPS	203	CCSYAGSYTVF
1-24	132	TGTSSDVGYNLVS	168	EGNKRPS	204	CCSYAGGSVVF
1-25	133	SGTSSDVGSYNLVS	169	EASKRPS	205	CCSYAGSSTVF
1-26	134	TGTSSDVGGYNLVS	170	EGTKRPS	206	CCSFVRSSAHVVF
1-27	135	TGTSSYVGHYNLVS	171	EGSRRPS	207	CCSYAGSYTHYVF
1-28	136	TGTSSGVGSYNLVS	172	EGSQRPS	208	CCSYAGSSTHYVF
1-29	137	TGTSSYVGHYNLVS	173	EGSRRPS	209	CCSYAGSYTHYVF
1-30	138	TGTSRDVGSYNLVS	174	EGTKRPS	210	CCSYAGSRTPVVF
1-31	139	TGTSSDVGKYNLVS	175	EGSQRPS	211	CCSYAGSRTPVVF
1-32	140	TGTSSDVGGYNLVS	176	EGNKRPS	212	CCSYAGSSTFPVVF
1-33	141	TGTSSDVGSYNLLS	177	EASKRPS	213	CCSYAGSRTPVVF
1-34	142	TGTSSDVGGYNLVS	178	EASKRPS	214	CCSYAGSYIPVVF
1-35	143	TGTSSDVGSYSLVS	179	EGTKRPS	215	CCSYAGSYSYVVF
1-36	144	TGTSSDVGSYSLVS	180	EGDKRPS	216	CCSYAGSYSYVVF

Table 21. SARS-CoV-2 S1 Variable Heavy Chain CDRs

Name	SEQ ID	CDRH1	SEQ ID	CDRH2	SEQ ID	CDRH3
------	--------	-------	--------	-------	--------	-------

	NO		NO		NO	
1-21	217	FTFSSYTMS	283	SIISGSGGSTYYA	349	CAREGYRDYLWYFDLW
1-22	218	FTFSSYAMN	284	SIISGSGGSTYYA	350	CAREGYRDYLWYFDLW
1-30	219	FTFSSYAMT	285	SAISGSGGSTFYA	351	CTRPPYGDYGDYW
1-35	220	FTFSSNYAMS	286	SDISGSGGSTYYA	352	CVKGTIPIFGVIRSAFDYW
1-17	221	FTFSSYAMY	287	SAISGSAGSTYYA	353	CARDTNDFWSGYSIFDPW
1-27	222	FTFSSYAMT	288	SAISGSGGGTYA	354	CASSSSWQFDYW
1-12	223	ITFSSYAMS	289	SGISGSGGSTYYA	355	CAKHGSGTIFGVVIAKYYFDYW
1-2	224	FTFSSQAMS	290	SAISGSGGGTYA	356	CAKDRRFGEFDPW
1-3	225	FTFSSQAMS	291	SAISGSGGGTYA	357	CAKDRRFGEFDPW
1-23	226	FTFSSY AIS	292	SIISGSGGSTYYA	358	CAREGYRDYLWYFDLW
1-7	227	FTFSSYDMS	293	SVISGSGGSTYYA	359	CAKGPLVGWYFDLW
1-31	228	FTFSSYAMN	294	SAISGSGGSTFYA	360	CTRPPYGDYGDYW
1-4	229	FTFSSQAMS	295	SAISGSGGGTYA	361	CAKDRRFGEFDPW
1-8	230	FTFSSYDMS	296	SVISGSGGSTYYA	362	CAKGPLVGWYFDLW
1-9	231	FTFSSYDMS	297	SVISGSGGSTYYA	363	CAKGPLVGWYFDLW
1-32	232	FTFSSYAMY	298	SAISGSGGSTFYA	364	CTRPPYGDYGDYW
1-33	233	FTFSSYAMI	299	SAISGSGGSTFYA	365	CTRPPYGDYGDYW
1-24	234	FTFSSYAMN	300	SIISGSGGSTYYA	366	CAREGYRDYLWYFDLW
1-40	235	FTFSSYAMY	301	SAISGSAGSTYYA	367	CARDTNDFWSGYSIFDPW
1-5	236	FTFSSQAMS	302	SAISGSGGGTYA	368	CAKDRRFGEFDPW
1-28	237	FTFSSYGMS	303	SAISGSGGGTYA	369	CASSSSWQFDYW
1-10	238	FTFSSYDMS	304	SVISGSGGSTYYA	370	CAKGPLVGWYFDLW
1-26	239	FTFSSY AIS	305	SAISGSGYSTYYA	371	CARGAPDSSGYFQGEVYFDYW
1-42	240	VTFSSYAMS	306	SAITGSGGSTYYA	372	CAKHGSGTIFGVVIAKYYFDYW
1-13	241	FTFSSHAMS	307	SAISGSAGSTYYA	373	CARDTKDFWSGYCIFDPW
1-14	242	FTFTSYAMS	308	SAISGSGGSTYYA	374	CANDTNDFWFGYWIFDPW
1-6	243	FTFSSQAMS	309	TAISGSGGGTYA	375	CAKDTIFGEFYPW
1-20	244	ITFSSYAMS	310	SGISGSGGSTYYA	376	CANHGSGTIFGVVIAKYYFDYW
1-47	245	FTFSSQAMS	311	SAISGSGGGTYA	377	CAKDRRFGEFDPW
1-29	246	FTFSSYAMY	312	SAISGSAGSTYYA	378	CARDTNDFWSGYSIFDPW
1-1	247	FTFSSHAMY	313	SAISGSAGSTYYA	379	CARDTKDFWSGYCIFDPW
1-19	248	ITFSSYAMS	314	SGISGSGGSTYYA	380	CAKHGSGTIFGVVIAKYYFDYC
1-16	249	LTFSSYAMS	315	SDISGSGGSTYYA	381	CVKGTIPIFGVIRSAFDYW
1-34	250	FTFSSYAMT	316	SAISGSGGGTYA	382	CASSSSWQFDYW
1-48	251	FTFSSHAMY	317	SAISGSAGSTYYA	383	CAHDTKDFWSGYCIFDPW

1-49	252	FTFSSYAMY	318	SAISGSAGSTYYA	384	CARDTNDFW
1-18	253	FTFSSHAMS	319	SAISGSAGSTYYA	385	CAHDTKDFWSGYCIFDPW
1-11	254	FTFSSYDMS	320	SAISGSGGTTYA	386	CARGPLDFW
1-25	255	FTFTSYAMY	321	SAISGSAGSTYYA	387	CARDTNDFWSGYSIFDPW
1-36	256	FTFSDYAMN	322	SIISGSGGSTYYA	388	CAREGYRDYLWYFDLW
1-15	257	ITFSSHAMS	323	SGISGSGGSTYYA	389	CAKHGSGTIFGVVIAKYFYDYW
1-51	258	FTFSSYAMS	324	SVISGSGGSTYYA	390	CAREGYRDYLWYFDLW
1-52	259	FTFSSNYAMS	325	SAISGSAGSTYYA	391	CARVRQGLRRTWYFYDYW
1-53	260	FTFSSYTMS	326	SVISGSGGSTYYA	392	CAREGYRDYLWYFDLW
1-54	261	FTFSSYAMY	327	SAISGSAGSTYYA	393	CARDTNDFWSGYSIFDPW
1-55	262	FTFSSYAMA	328	SAISGSGSSTYYA	394	CASSSSWQFDYW
1-56	263	FTFSSYAMY	329	SAISGSAGSTYYA	395	CARDTNDFWSGYSIFDPW
1-57	264	FTFSSYAMT	330	SAISGSGGSTFYA	396	CTRPPYGDYGDYW
1-58	265	FTFSSYAMG	331	SAISTSGGSTYYA	397	CAKGLWFGGGGFDPW
1-59	266	FTFSSYAMN	332	SVISGSGGSTYYA	398	CAREGYRDYLWYFDLW
1-6	267	FTFSSYAMY	333	SAISSSGGSTYYA	399	CTRPPYGDYGDYW
1-61	268	FTFSSYAVS	334	SDISGSGGSTYYA	400	CVKGTIPIFGVIRSAFDYW
1-62	269	FRFSSYAMS	335	SAISGSGGTTYA	401	CAKDFHGIAAAGIDYW
1-63	270	FMFSSYAMS	336	SAISGSGGSTYYT	402	CAKDASGWPWNHFDLW
1-64	271	FAFSSYAMS	337	SAISGSGGDTYYA	403	CAHSRDSSSWYVDYW
1-37	272	FTFTSYAMN	338	TAISGSGGSTFYA	404	CTRPPYGDYGDYW
1-38	273	ITFSSYAMS	339	SGISGSGGSTYYA	405	CAKHGSGTIFGVVIAKYFYDYW
1-39	274	FTFSSYAMY	340	TAISGSGGSTFYA	406	CTRPPYGYG DYW
1-41	275	FTFSSYDMS	341	SVISGSGGSTYYA	407	CAKGPLVGWYFDLW
1-43	276	FTFSSQAMS	342	TAISGSGGGTYA	408	CAKDTIFGEFYPW
1-44	277	FTFSSQAMS	343	TAISGSGGGTYA	409	CAKDRKFGFDPW
1-45	278	FTFSSYAMS	344	SIISGSAGSTYYA	410	CARDGYKYCLW
1-46	279	FTFTSYAMS	345	SAISGSGGSTYYA	411	CANDTSDFCFGYWIFDPW
1-50	280	FTFSSYDMS	346	SAISGSGGTTYA	412	CARGPLDFW
1-65	281	FTFSSHAMS	347	SAISGSAGSTYYA	413	CARDTKDFWSGYCIFDPW
1-66	282	FTFTSYAMS	348	SAISGSGGSTYYA	414	CANDTNDFWFGYWIFDPW

Table 22. SARS-CoV-2 S1 Variable Light Chain CDRs

Name	SEQ ID NO	CDRL1	SEQ ID NO	CDRL2	SEQ ID NO	CDRL3
1-21	415	TGTSSDVGSNNLVS	474	EGNKRPS	533	CCSYAGSVVF
1-22	416	TGTSTDVGSYNLVS	475	EGSQRPS	534	CCSYAGSSTVF

1-30	417	TGTSRDVGSYNLVS	476	EGTKRPS	535	CCSYAGSRTPVVF
1-35	418	TGTSSDVGSYSLVS	477	EGTKRPS	536	CCSYAGSYSYVVF
1-17	419	TGTSSDIGSYNLVS	478	EGTKRPS	537	CCSYAGSRTYVF
1-27	420	TGTSSYVGHYNLVS	479	EGSRRPS	538	CCSYAGSYTHYVF
1-12	421	TGTSSDVGHYNLVS	480	EGTKRPS	539	CCSYAGSSSFVVF
1-2	422	TGTSSGVGSYNLVS	481	EGIKRPS	540	CCSYAGSSSFVVF
1-3	423	TGISSDVGSYNLVS	482	EGTKRPS	541	CCSYAGSSSFVVF
1-23	424	TGTSDNVGSYNLVS	483	EGNKRPS	542	CCSYAGSYTVF
1-7	425	TGTSSDFGSYNLVS	484	EGNKRPS	543	CCSYAGSSTFVVF
1-31	426	TGTSSDVGKYNLVS	485	EGSQRPS	544	CCSYAGSRTPVVF
1-4	427	TGTSSDVGTYNLVS	486	EGSQRPS	545	CCSYAGSSSFVVF
1-8	428	TGTSDNVGSYNLVS	487	EGSKRPS	546	CCSYAGSRYVVF
1-9	429	TGTSSDVGDYNLVS	488	EGSNRPS	547	CCSYAGSSSFVVF
1-32	430	TGTSSDVGGYNLVS	489	EGNKRPS	548	CCSYAGSSTFPVVF
1-33	431	TGTSSDVGSYNLLS	490	EASKRPS	549	CCSYAGSRTPVVF
1-24	432	TGTSSDVGYYNLVS	491	EGNKRPS	550	CCSYAGGSVVF
1-5	433	TGTSSGVGSYNLVS	492	EASKRPS	551	CCSYAGSYTFVVF
1-28	434	TGTSSGVGSYNLVS	493	EGSQRPS	552	CCSYAGSSTHYVF
1-10	435	TGTSSNVGSYNLVS	494	EGTKRPS	553	CCSYAGSSSFVVF
1-26	436	TGTSSDVGGYNLVS	495	EGTKRPS	554	CCSFVRSSAHVVF
1-42	437	TGTSSDVGRYNLVS	496	EGGKRPS	555	CCSYASSSTLVF
1-20	438	TGTSSDVGHYNLVS	497	EGTKRPS	556	CCSYAGSSSFVVF
1-47	439	TGTSSGVGSYNLVS	498	EGIKRPS	557	CCSYAGSSSFVVF
1-29	440	TGTSSDIGSYNLVS	499	EGTKRPS	558	CCSYAGSRTYVF
1-1	441	TGTSSDIGSYNLVS	500	EGTNRPS	559	CCSYAGSRTYVF
1-19	442	TGTSSDVGHYNLVS	501	EGTKRPS	560	CCSYAGSSSFVVF
1-16	443	TGTSSDVGSYSLVS	502	EGDKRPS	561	CCSYAGSYSYVVF
1-34	444	TGTSSYVGHYNLVS	503	EGSRRPS	562	CCSYAGSYTHYVF
1-48	445	TGTSSDIGSYNLVS	504	EGTKRPS	563	CCSYAGSRTYVF
1-49	446	TGTSSDIGSYNLVS	505	EGTKRPS	564	CCSYAGSRTYVF
1-25	447	TGTSSDIGSYNLVS	506	EGTKRPS	565	CCSYAGSRTYVF
1-36	448	SGTSSDVGSYNLVS	507	EASKRPS	566	CCSYAGSSTVF
1-51	449	TGTSSDVGSYDLVS	508	EGNKRPS	567	CCSYAGSSVVF
1-52	450	TGTSSDVGSSNLVS	509	EGSKRPS	568	CCSYAGSLYVF
1-53	451	TGTSTDVGSYNLVS	510	EGTKRPS	569	CCSYAGSYTSVVF
1-54	452	TGTSSDIGSYNLVS	511	EGTKRPS	570	CCYHSRTRTHVF

1-55	453	TGTSSDLGSYNIVS	512	EGSRRPS	571	CCSYAGSYTHYVF
1-56	454	TGTSSDIGSYNLVS	513	EGTKRPS	572	CCYHSRTRTHVS
1-57	455	TGTSSDVVGKYNLVS	514	EGVKRPS	573	CCSYAGSRTPVVF
1-58	456	TGTRSDVGSYNLVS	515	EVSKRPS	574	CCSYAGDSFPYVF
1-59	457	TGTSSDVGSFNLVS	516	EVSKRPS	575	CCSYAGSSVVF
1-6	458	TGTSSDVGGYNLVS	517	EASKRPS	576	CCSYAGSYIPVVF
1-61	459	TGTSSDVGSYSLVS	518	EGGKRPS	577	CCSYAGSYSYVVF
1-62	460	TGTSSDVGSHNLVS	519	EGGKRPS	578	CCSYSGRYTYVF
1-63	461	TGTSSDVGSYYLVS	520	EGDKRPS	579	CCSHAGRYPYVF
1-64	462	TGTSSGVGSYNLVS	521	AGSKRPS	580	CCSYLGSGTFDVLV
1-37	463	TGTSSDIGSYNLVS	522	EGTKRPS	581	CCSYAGSRITYVF
1-38	464	TGTSSDIGSYNLVS	523	EGTKRPS	582	CCSYAGSRITYVF
1-39	465	TGTSSDIGSYNLVS	524	EGTKRPS	583	CCSYAGSRITYVF
1-41	466	TGTSSDIGSYNLVS	525	EGTKRPS	584	CCSYAGSRITYVF
1-43	467	TGTSSDIGSYNLVS	526	EGTKRPS	585	CCSYAGSRITYVF
1-44	468	TGTSSDIGSYNLVS	527	EGTKRPS	586	CCSYAGSRITYVF
1-45	469	TGTSSDIGSYNLVS	528	EGTKRPS	587	CCSYAGSRITYVF
1-46	470	TGTSSDIGSYNLVS	529	EGTKRPS	588	CCSYAGSRITYVF
1-50	471	TGTSSDIGSYNLVS	530	EGTKRPS	589	CCSYAGSRITYVF
1-65	472	TGTSSDIGSYNLVS	531	EGTKRPS	590	CCSYAGSRITYVF
1-66	473	TGTSSDIGSYNLVS	532	EGTKRPS	591	CCSYAGSRITYVF

Table 23. SARS-CoV-2 S1 Variant Sequences Variable Heavy Chain

Name	SEQ ID	Amino Acid Sequence
1-21	592	EVQLLES GGG LVQP GGS LRLSCAASGFTFS SYTMSWVRQAPGK GLEWVSII SGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAREGY RDYLYWYFDLWGQGTLVTVSS
1-22	593	EVQLLES GGG LVQP GGS LRLSCAASGFTFS SYAMNWVRQAPGK GLEWVSII SGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAREGY RDYLYWYFDLWGQGTLVTVSS
1-30	594	EVQLLES GGG LVQP GGS LRLSCAASGFTFS SYAMTWVRQAPGK GLEWVSA ISGSGGSTFYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCTRPPY GDYGDYWGQGTLVTVSS
1-35	595	EVQLLES GGG LVQP GGS LRLSCAASGFTFS NYAMSWVRQAPGK GLEWVSD ISGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCVKGTI PIFGVIRSAFDYWGQGTLVTVSS
1-17	596	EVQLLES GGG LVQP GGS LRLSCAASGFTFS SYAMYWVRQAPGK GLEWVSA ISGSAGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARDTN DFWSGYSIFDPWGQGTLVTVSS
1-27	597	EVQLLES GGG LVQP GGS LRLSCAASGFTFS SYAMTWVRQAPGK GLEWVSA ISGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCASSSS WQFDYWGQGTLVTVSS

1-12	598	EVQLLESGGGLVQPGGSLRLSCAASGITFSSYAMSWVRQAPGKGLEWVSGI SGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKHGS GTIFGVVIAKYYFDYWGGQTLVTVSS
1-2	599	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSQAMSWVRQAPGKGLEWVSA ISGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKDR RFGEFDPWGQGTLLTVSS
1-3	600	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSQAMSWVRQAPGKGLEWVSA ISGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKDR RFGEFDPWGQGTLLTVSS
1-23	601	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYAIWVRQAPGKGLEWVSIIS SGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAREGYR DYLWYFDLWGQGTLLTVSS
1-7	602	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYDMSWVRQAPGKGLEWVSV ISGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKGPL VGWYFDLWGQGTLLTVSS
1-31	603	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYAMNWVRQAPGKGLEWVSA ISGSGGSTFYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCTRPPY GDYGDYWGQGTLLTVSS
1-4	604	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSQAMSWVRQAPGKGLEWVSA ISGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKDR RFGEFDPWGQGTLLTVSS
1-8	605	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYDMSWVRQAPGKGLEWVSV ISGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKGPL VGWYFDLWGQGTLLTVSS
1-9	606	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYDMSWVRQAPGKGLEWVSV ISGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKGPL VGWYFDLWGQGTLLTVSS
1-32	607	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYAMYWVRQAPGKGLEWVSA ISGSGGSTFYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCTRPPY GDYGDYWGQGTLLTVSS
1-33	608	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYAMIWVRQAPGKGLEWVSAI SGSGGSTFYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCTRPPYG DYGDYWGQGTLLTVSS
1-24	609	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYAMNWVRQAPGKGLEWVSI SGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAREGY RDYLWYFDLWGQGTLLTVSS
1-40	610	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYAMYWVRQAPGKGLEWVSA ISGSAGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARDTN DFWSGYSIFDPWGQGTLLTVSS
1-5	611	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSQAMSWVRQAPGKGLEWVSA ISGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKDR RFGEFDPWGQGTLLTVSS
1-28	612	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYGMSWVRQAPGKGLEWVSA ISGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCASSSS WQFDYWGGQGTLLTVSS
1-10	613	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYDMSWVRQAPGKGLEWVSV ISGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKGPL VGWYFDLWGQGTLLTVSS
1-26	614	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYAIWVRQAPGKGLEWVSAI SGSGYSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARGAP DSSGYFQGEVYFDYWGGQGTLLTVSS
1-42	615	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYAMSWVRQAPGKGLEWVSA ITGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKHGS GTIFGVVIAKYYFDYWGGQGTLLTVSS
1-13	616	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSHAMSWVRQAPGKGLEWVSA

		ISGSAGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARDTK DFWSGYCIFDPWGQGLTVTVSS
1-14	617	EVQLES GGGLVQPGGSLRLS CAASGFTFTSYAMSWVRQAPGKGLEWVSA ISGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCANDTN DFWFGYWIFDPWGQGLTVTVSS
1-6	618	EVQLES GGGLVQPGGSLRLS CAASGFTFSSQAMSWVRQAPGKGLEWVTA ISGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKDTI FGEFYPWGQGLTVTVSS
1-20	619	EVQLES GGGLVQPGGSLRLS CAASGITFSSYAMSWVRQAPGKGLEWVSGI SGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCANHGS GTIFGVVIAKYYFDYWGQGLTVTVSS
1-47	620	EVQLES GGGLVQPGGSLRLS CAASGFTFSSQAMSWVRQAPGKGLEWVSA ISGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKDR RFGEFDPWGQGLTVTVSS
1-29	621	EVQLES GGGLVQPGGSLRLS CAASGFTFSSYAMYWVRQAPGKGLEWVSA ISGSAGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARDTN DFWSGYSIFDPWGQGLTVTVSS
1-1	622	EVQLES GGGLVQPGGSLRLS CAASGFTFSSHAMYWVRQAPGKGLEWVSA ISGSAGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARDTK DFWSGYSIFDPWGQGLTVTVSS
1-19	623	EVQLES GGGLVQPGGSLRLS CAASGITFSSYAMSWVRQAPGKGLEWVSGI SGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKHGS GTIFGVVIAKYYFDYCGQGLTVTVSS
1-16	624	EVQLES GGGLVQPGGSLRLS CAASGLTFSSYAMSWVRQAPGKGLEWVSD ISGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCVKGTI PIFGVIRSAFDYWGQGLTVTVSS
1-34	625	EVQLES GGGLVQPGGSLRLS CAASGFTFSSYAMTWVRQAPGKGLEWVSA ISGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCASSSS WQFDYWGQGLTVTVSS
1-48	626	EVQLES GGGLVQPGGSLRLS CAASGFTFSSHAMYWVRQAPGKGLEWVSA ISGSAGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAHDTK DFWSGYCIFDPWGQGLTVTVSS
1-49	627	EVQLES GGGLVQPGGSLRLS CAASGFTFSSYAMYWVRQAPGKGLEWVSA ISGSAGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARDTN DFWGQGLTVTVSS
1-18	628	EVQLES GGGLVQPGGSLRLS CAASGFTFSSHAMSWVRQAPGKGLEWVSA ISGSAGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAHDTK DFWSGYCIFDPWGQGLTVTVSS
1-11	629	EVQLES GGGLVQPGGSLRLS CAASGFTFSSYDMSWVRQAPGKGLEWVSA ISGSGGTTYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARGPL DFWGQGLTVTVSS
1-25	630	EVQLES GGGLVQPGGSLRLS CAASGFTFTSYAMYWVRQAPGKGLEWVSA ISGSAGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARDTN DFWSGYSIFDPWGQGLTVTVSS
1-36	631	EVQLES GGGLVQPGGSLRLS CAASGFTFSDYAMNWVRQAPGKGLEWVSII SGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAREGY RDYWLWYFDLWGQGLTVTVSS
1-15	632	EVQLES GGGLVQPGGSLRLS CAASGITFSSHAMSWVRQAPGKGLEWVSGI SGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKHGS GTIFGVVIAKYYFDYWGQGLTVTVSS
1-51	633	EVQLES GGGLVQPGGSLRLS CAASGFTFSSYAMSWVRQAPGKGLEWVSV ISGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAREGY RDYWLWYFDLWGQGLTVTVSS
1-52	634	EVQLES GGGLVQPGGSLRLS CAASGFTFSNYAMSWVRQAPGKGLEWVSA ISGSAGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARVRQ

		GLRRTWYFFDYWGQGLTVTVSS
1-53	635	EVQLLES GGGLVQPGGSLRLS CAASGFTFSSY TMSWVRQAPGKGLEWVSVI SGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAREGY RDYLYWYFDLWGQGLTVTVSS
1-54	636	EVQLLES GGGLVQPGGSLRLS CAASGFTFSSY AMYWVRQAPGKGLEWVSA ISGSAGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARDTN DFWSGYSIFDPWGQGLTVTVSS
1-55	637	EVQLLES GGGLVQPGGSLRLS CAASGFTFSSY AMAWVRQAPGKGLEWVSA ISGSGSSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCASSSS WQFDYWGQGLTVTVSS
1-56	638	EVQLLES GGGLVQPGGSLRLS CAASGFTFSSY AMYWVRQAPGKGLEWVSA ISGSAGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARDTN DFWSGYSIFDPWGQGLTVTVSS
1-57	639	EVQLLES GGGLVQPGGSLRLS CAASGFTFSSY AMTWVRQAPGKGLEWVSA ISGSGGSTFYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCTRPPY GDYGDYWGQGLTVTVSS
1-58	640	EVQLLES GGGLVQPGGSLRLS CAASGFTFSSY AMGWVRQAPGKGLEWVSA ISTSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKGL WFGGGGFDPWGQGLTVTVSS
1-59	641	EVQLLES GGGLVQPGGSLRLS CAASGFTFSSY AMNWVRQAPGKGLEWVSV ISGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAREGY RDYLYWYFDLWGQGLTVTVSS
1-6	642	EVQLLES GGGLVQPGGSLRLS CAASGFTFSSY AMYWVRQAPGKGLEWVSA ISSSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCTRPPY GDYGDYWGQGLTVTVSS
1-61	643	EVQLLES GGGLVQPGGSLRLS CAASGFTFSSY AVSWVRQAPGKGLEWVSDI SGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCVKGTIP IFGVIRSAFDYWGQGLTVTVSS
1-62	644	EVQLLES GGGLVQPGGSLRLS CAASGFRFSSY AMSWVRQAPGKGLEWVSA ISGSGGTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKDFH GIAAAGIDYWGQGLTVTVSS
1-63	645	EVQLLES GGGLVQPGGSLRLS CAASGFMFSSY AMSWVRQAPGKGLEWVSA ISGSGGSTYYTDSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKDGA SGWPNWHFDLWGQGLTVTVSS
1-64	646	EVQLLES GGGLVQPGGSLRLS CAASGFAFSSY AMSWVRQAPGKGLEWVSA ISGSGGDTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAHSRD SSSWYVDYWGQGLTVTVSS
1-37	647	EVQLLES GGGLVQPGGSLRLS CAASGFTFTSY AMNWVRQAPGKGLEWVTA ISGSGGSTFYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCTRPPY GDYGDYWGQGLTVTVSS
1-38	648	EVQLLES GGGLVQPGGSLRLS CAASGITFSSY AMSWVRQAPGKGLEWVSGI SGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKHGS GTIFGVVIAKYYFDYWGQGLTVTVSS
1-39	649	EVQLLES GGGLVQPGGSLRLS CAASGFTFSSY AMYWVRQAPGKGLEWVTA ISGSGGSTFYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCTRPPY GYYGDYWGQGLTVTVSS
1-41	650	EVQLLES GGGLVQPGGSLRLS CAASGFTFSSY DMSWVRQAPGKGLEWVSV ISGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKGPL VGWYFDLWGQGLTVTVSS
1-43	651	EVQLLES GGGLVQPGGSLRLS CAASGFTFSSQ AMSWVRQAPGKGLEWVTA ISGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKDTI FGEFYPWGQGLTVTVSS
1-44	652	EVQLLES GGGLVQPGGSLRLS CAASGFTFSSQ AMSWVRQAPGKGLEWVTA ISGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKDR KFGEFDPWGQGLTVTVSS

1-45	653	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYAMSWVRQAPGKGLEWVSII SGSAGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARDGY KYCLWGQGTLVTVSS
1-46	654	EVQLLESGGGLVQPGGSLRLSCAASGFTFTSYAMSWVRQAPGKGLEWVSA ISGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCANDTS DFCFGYWIFDPWGQGTLVTVSS
1-50	655	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYDMSWVRQAPGKGLEWVSA ISGSGGTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARGPL DFWGQGTLVTVSS
1-65	656	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSHAMSWVRQAPGKGLEWVSA ISGSAGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARDTK DFWSGYCIFDPWGQGTLVTVSS
1-66	657	EVQLLESGGGLVQPGGSLRLSCAASGFTFTSYAMSWVRQAPGKGLEWVSA ISGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCANDTN DFWFGYWIFDPWGQGTLVTVSS

Table 24. SARS-CoV-2 S1 Variant Sequences Variable Light Chain

Name	SEQ ID NO	Amino Acid Sequence
1-21	658	QSALTQPASVSGSPGQSITISCTGTSSDVGSNNLVSWYQQHPGKAPKLMIEGNKR PSGVSNRFGSGSKSGNTASLTISGLQAEDEADYYCCSYAGSVVFVGGGTKLTVL
1-22	659	QSALTQPASVSGSPGQSITISCTGTSTDVGSYNLVSWYQQHPGKAPKLMIEGSQR PSGVSNRFGSGSKSGNTASLTISGLQAEDEADYYCCSYAGSSTVFVGGGTKLTVL
1-30	660	QSALTQPASVSGSPGQSITISCTGTSTRDVGSYNLVSWYQQHPGKAPKLMIEGTR PSGVSNRFGSGSKSGNTASLTISGLQAEDEADYYCCSYAGSRTPVVFVGGGTKLTVL
1-35	661	QSALTQPASVSGSPGQSITISCTGTSSDVGSYSLVSWYQQHPGKAPKLMIEGTR PSGVSNRFGSGSKSGNTASLTISGLQAEDEADYYCCSYAGSYSYVVFVGGGTKLTVL
1-17	662	QSALTQPASVSGSPGQSITISCTGTSSDIGSYNLVSWYQQHPGKAPKLMIEGTR SGVSNRFGSGSKSGNTASLTISGLQAEDEADYYCCSYAGSRTYVVFVGGGTKLTVL
1-27	663	QSALTQPASVSGSPGQSITISCTGTSSYVGHYNLVSWYQQHPGKAPKLMIEGSRR PSGVSNRFGSGSKSGNTASLTISGLQAEDEADYYCCSYAGSYTHYVVFVGGGTKLTVL
1-12	664	QSALTQPASVSGSPGQSITISCTGTSSDVGHYNLVSWYQQHPGKAPKLMIEGTR PSGVSNRFGSGSKSGNTASLTISGLQAEDEADYYCCSYAGSSSFVVFVGGGTKLTVL
1-2	665	QSALTQPASVSGSPGQSITISCTGTSSGVGSYNLVSWYQQHPGKAPKLMIEGIKRP SGVSNRFGSGSKSGNTASLTISGLQAEDEADYYCCSYAGSSSFVVFVGGGTKLTVL
1-3	666	QSALTQPASVSGSPGQSITISCTGISSDVGSYNLVSWYQQHPGKAPKLMIEGTR SGVSNRFGSGSKSGNTASLTISGLQAEDEADYYCCSYAGSSSFVVFVGGGTKLTVL
1-23	667	QSALTQPASVSGSPGQSITISCTGTSDNVGSYNLVSWYQQHPGKAPKLMIEGNKR PSGVSNRFGSGSKSGNTASLTISGLQAEDEADYYCCSYAGSYTVVFVGGGTKLTVL
1-7	668	QSALTQPASVSGSPGQSITISCTGTSSDFGSYNLVSWYQQHPGKAPKLMIEGNKR PSGVSNRFGSGSKSGNTASLTISGLQAEDEADYYCCSYAGSSTFVVFVGGGTKLTVL
1-31	669	QSALTQPASVSGSPGQSITISCTGTSSDVGKYNLVSWYQQHPGKAPKLMIEGSQR PSGVSNRFGSGSKSGNTASLTISGLQAEDEADYYCCSYAGSRTPVVFVGGGTKLTVL
1-4	670	QSALTQPASVSGSPGQSITISCTGTSSDVGTYNLVSWYQQHPGKAPKLMIEGSQR PSGVSNRFGSGSKSGNTASLTISGLQAEDEADYYCCSYAGSSSFVVFVGGGTKLTVL
1-8	671	QSALTQPASVSGSPGQSITISCTGTSDNVGSYNLVSWYQQHPGKAPKLMIEGSKR PSGVSNRFGSGSKSGNTASLTISGLQAEDEADYYCCSYAGSRYVVFVGGGTKLTVL
1-9	672	QSALTQPASVSGSPGQSITISCTGTSSDVGDYNLVSWYQQHPGKAPKLMIEGSNR PSGVSNRFGSGSKSGNTASLTISGLQAEDEADYYCCSYAGSSSFVVFVGGGTKLTVL
1-32	673	QSALTQPASVSGSPGQSITISCTGTSSDVGGYNLVSWYQQHPGKAPKLMIEGNKR PSGVSNRFGSGSKSGNTASLTISGLQAEDEADYYCCSYAGSSTFPVVFVGGGTKLTVL
1-33	674	QSALTQPASVSGSPGQSITISCTGTSSDVGSYNLLSWYQQHPGKAPKLMIEASKR PSGVSNRFGSGSKSGNTASLTISGLQAEDEADYYCCSYAGSRTPVVFVGGGTKLTVL

1-24	675	QSALTQPASVSGSPGQSITISCTGTSSDVGYNLSVSWYQQHPGKAPKLMIEGDKR PSGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGGSVVFGGGTCLTVL
1-5	676	QSALTQPASVSGSPGQSITISCTGTSSGVGSYNLSVSWYQQHPGKAPKLMIEASKR PSGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSYTFAVFGGGTCLTVL
1-28	677	QSALTQPASVSGSPGQSITISCTGTSSGVGSYNLSVSWYQQHPGKAPKLMIEGSQR PSGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSSTHYVFGGGTCLTVL
1-10	678	QSALTQPASVSGSPGQSITISCTGTSSNVGSYNLSVSWYQQHPGKAPKLMIEGDKR PSGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSSSFVVFGGGTCLTVL
1-26	679	QSALTQPASVSGSPGQSITISCTGTSSDVGGYNLSVSWYQQHPGKAPKLMIEGDKR PSGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSFVRSSAHVVFGGGTCLTVL
1-42	680	QSALTQPASVSGSPGQSITISCTGTSSDVGRYNLSVSWYQQHPGKAPKLMIEGGKR PSGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYASSSTLVFGGGTCLTVL
1-20	681	QSALTQPASVSGSPGQSITISCTGTSSDVGHYNLSVSWYQQHPGKAPKLMIEGDKR PSGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSSSFVVFGGGTCLTVL
1-47	682	QSALTQPASVSGSPGQSITISCTGTSSGVGSYNLSVSWYQQHPGKAPKLMIEGDKR SGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSSSFVVFGGGTCLTVL
1-29	683	QSALTQPASVSGSPGQSITISCTGTSSDIGSYNLSVSWYQQHPGKAPKLMIEGDKR SGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSRTYVFGGGTCLTVL
1-1	684	QSALTQPASVSGSPGQSITISCTGTSSDIGSYNLSVSWYQQHPGKAPKLMIEGTNR SGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSRTYVFGGGTCLTVL
1-19	685	QSALTQPASVSGSPGQSITISCTGTSSDVGHYNLSVSWYQQHPGKAPKLMIEGDKR PSGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSSSFVVFGGGTCLTVL
1-16	686	QSALTQPASVSGSPGQSITISCTGTSSDVGSYSLVSWYQQHPGKAPKLMIEGDKR PSGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSYSYVVFGGGTCLTVL
1-34	687	QSALTQPASVSGSPGQSITISCTGTSSYVGHYNLSVSWYQQHPGKAPKLMIEGSRR PSGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSYTHYVFGGGTCLTVL
1-48	688	QSALTQPASVSGSPGQSITISCTGTSSDIGSYNLSVSWYQQHPGKAPKLMIEGDKR SGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSRTYVFGGGTCLTVL
1-49	689	QSALTQPASVSGSPGQSITISCTGTSSDIGSYNLSVSWYQQHPGKAPKLMIEGDKR SGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSRTYVFGGGTCLTVL
1-25	690	QSALTQPASVSGSPGQSITISCTGTSSDIGSYNLSVSWYQQHPGKAPKLMIEGDKR SGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSRTYVFGGGTCLTVL
1-36	691	QSALTQPASVSGSPGQSITISCTGTSSDVGSYNLSVSWYQQHPGKAPKLMIEASKR PSGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSSTVFGGGTCLTVL
1-51	692	QSALTQPASVSGSPGQSITISCTGTSSDVGSYDLVSWYQQHPGKAPKLMIEGDKR PSGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSSVVFGGGTCLTVL
1-52	693	QSALTQPASVSGSPGQSITISCTGTSSDVGSYNLSVSWYQQHPGKAPKLMIEGSKRP SGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSLYVFGGGTCLTVL
1-53	694	QSALTQPASVSGSPGQSITISCTGTSTDVGSYNLSVSWYQQHPGKAPKLMIEGDKR PSGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSYTSVVFGGGTCLTVL
1-54	695	QSALTQPASVSGSPGQSITISCTGTSSDIGSYNLSVSWYQQHPGKAPKLMIEGDKR SGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCYHSRTRTHVFGGGTCLTVL
1-55	696	QSALTQPASVSGSPGQSITISCTGTSSDLGSYNIVSWYQQHPGKAPKLMIEGSRRP SGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSYTHYVFGGGTCLTVL
1-56	697	QSALTQPASVSGSPGQSITISCTGTSSDLGSYNIVSWYQQHPGKAPKLMIEGSRRP SGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSYTHYVFGGGTCLTVL
1-57	698	QSALTQPASVSGSPGQSITISCTGTSSDLGSYNIVSWYQQHPGKAPKLMIEGSRRP SGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSYTHYVFGGGTCLTVL
1-58	699	QSALTQPASVSGSPGQSITISCTGTSSDLGSYNIVSWYQQHPGKAPKLMIEGSRRP SGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSYTHYVFGGGTCLTVL
1-59	700	QSALTQPASVSGSPGQSITISCTGTSSDLGSYNIVSWYQQHPGKAPKLMIEGSRRP SGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSYTHYVFGGGTCLTVL
1-6	701	QSALTQPASVSGSPGQSITISCTGTSSDLGSYNIVSWYQQHPGKAPKLMIEGSRRP SGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSYTHYVFGGGTCLTVL

1-61	702	QSALTQPASVSGSPGQSITISCTGTSSDLGSYNIVSWYQQHPGKAPKLMIEGSRRP SGVSNRFSGSKSGNTASLTISGLQAEDEADYYCCSYAGSYTHYVFGGGTKLTVL
1-62	703	QSALTQPASVSGSPGQSITISCTGTSSDVGSNLSVSWYQQHPGKAPKLMIEGGKR PSGVSNRFSgSKSGNTasLTISGLQAEDEADYYCCSYSGRYTYVFGGGtKLTVL
1-63	704	QSALTQPASVSGSPGQSITISCTGTSSDVGSYLVSWYQQHPGKAPKLMIEGDKR PSGVSNRFSGSKSGNTASLTISGLQAEDEADYYCCSHAGRYPYVFGGGTKLTVL
1-64	705	QSALTQPASVSGSPGQSITISCTGTSSGVGSYNLSVSWYQQHPGKAPKLMIEAGSKR PSGVSNRFSGSKSGNTASLTISGLQAEDEADYYCCSYLGSGTFDVLFGGGTKLTVL
1-37	706	QSALTQPASVSGSPGQSITISCTGTSSDIGSYNLSVSWYQQHPGKAPKLMIEGTRKP SGVSNRFSGSKSGNTASLTISGLQAEDEADYYCCSYAGSRTYVFGGGTKLTVL
1-38	707	QSALTQPASVSGSPGQSITISCTGTSSDIGSYNLSVSWYQQHPGKAPKLMIEGTRKP SGVSNRFSGSKSGNTASLTISGLQAEDEADYYCCSYAGSRTYVFGGGTKLTVL
1-39	708	QSALTQPASVSGSPGQSITISCTGTSSDIGSYNLSVSWYQQHPGKAPKLMIEGTRKP SGVSNRFSGSKSGNTASLTISGLQAEDEADYYCCSYAGSRTYVFGGGTKLTVL
1-41	709	QSALTQPASVSGSPGQSITISCTGTSSDIGSYNLSVSWYQQHPGKAPKLMIEGTRKP SGVSNRFSGSKSGNTASLTISGLQAEDEADYYCCSYAGSRTYVFGGGTKLTVL
1-43	710	QSALTQPASVSGSPGQSITISCTGTSSDIGSYNLSVSWYQQHPGKAPKLMIEGTRKP SGVSNRFSGSKSGNTASLTISGLQAEDEADYYCCSYAGSRTYVFGGGTKLTVL
1-44	711	QSALTQPASVSGSPGQSITISCTGTSSDIGSYNLSVSWYQQHPGKAPKLMIEGTRKP SGVSNRFSGSKSGNTASLTISGLQAEDEADYYCCSYAGSRTYVFGGGTKLTVL
1-45	712	QSALTQPASVSGSPGQSITISCTGTSSDIGSYNLSVSWYQQHPGKAPKLMIEGTRKP SGVSNRFSGSKSGNTASLTISGLQAEDEADYYCCSYAGSRTYVFGGGTKLTVL
1-46	713	QSALTQPASVSGSPGQSITISCTGTSSDIGSYNLSVSWYQQHPGKAPKLMIEGTRKP SGVSNRFSGSKSGNTASLTISGLQAEDEADYYCCSYAGSRTYVFGGGTKLTVL
1-50	714	QSALTQPASVSGSPGQSITISCTGTSSDIGSYNLSVSWYQQHPGKAPKLMIEGTRKP SGVSNRFSGSKSGNTASLTISGLQAEDEADYYCCSYAGSRTYVFGGGTKLTVL
1-65	715	QSALTQPASVSGSPGQSITISCTGTSSDIGSYNLSVSWYQQHPGKAPKLMIEGTRKP SGVSNRFSGSKSGNTASLTISGLQAEDEADYYCCSYAGSRTYVFGGGTKLTVL
1-66	716	QSALTQPASVSGSPGQSITISCTGTSSDIGSYNLSVSWYQQHPGKAPKLMIEGTRKP SGVSNRFSGSKSGNTASLTISGLQAEDEADYYCCSYAGSRTYVFGGGTKLTVL

Table 25. Reformatted SARS-CoV-2 S1 Variant Sequences

Name	SEQ ID NO	Amino Acid Sequence
1-H1	717	EVQLLESggGLVQPGGSLRLScaasGFTfSSyTMSWVRQAPGKGLEWVSIISGSGG STYYADSVKGRFTISRDNskNTLYLQMNSLRAEDTAVYYCAREGYRDYLDWYFDL WGQGTLVTVSS
1-H2	718	EVQLLESggGLVQPGGSLRLScaasGFTfSSyAMNWVRQAPGKGLEWVSIISGSG GSTYYADSVKGRFTISRDNskNTLYLQMNSLRAEDTAVYYCAREGYRDYLDWYFD LWGQGTLVTVSS
1-H3	719	EVQLLESggGLVQPGGSLRLScaasGFTfSSyAMTWVRQAPGKGLEWVSAISGSG GSTFYADSVKGRFTISRDNskNTLYLQMNSLRAEDTAVYYCTRPPYGDYGDYWG QGTLVTVSS
1-H4	720	EVQLLESggGLVQPGGSLRLScaasGFTfSSyAMSWVRQAPGKGLEWVSDISGSG GSTYYADSVKGRFTISRDNskNTLYLQMNSLRAEDTAVYYCVKGTIPFGVIRSAF DYWGQGTLVTVSS
1-H5	721	EVQLLESggGLVQPGGSLRLScaasGFTfSSyAMYWVRQAPGKGLEWVSAISGSA GSTYYADSVKGRFTISRDNskNTLYLQMNSLRAEDTAVYYCARDTNDFWSGYSIF DPWGQGTLVTVSS
1-H6	722	EVQLLESggGLVQPGGSLRLScaasGFTfSSyAMTWVRQAPGKGLEWVSAISGSG GGTYADSVKGRFTISRDNskNTLYLQMNSLRAEDTAVYYCASSSSWQFDYWGQ GTLVTVSS

1-H8	723	EVQLLESGGGLVQPGGSLRLSCAASGITFSSYAMSWVRQAPGKGLEWVSGISGSG GSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKHGSGTIFGVVIA KYYFDYWGQGT LVT VSS
1-H9	724	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSQAMSWVRQAPGKGLEWVSAISGSG GGTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKDRRFGEFDPWG QGT LVT VSS
1-H10	725	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSQAMSWVRQAPGKGLEWVSAISGSG GGTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKDRRFGEFDPWG QGT LVT VSS
1-H11	726	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYAIWVRQAPGKGLEWVSIISGSGG STYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAREGYRDYLDWYFDL WGQGT LVT VSS
1-H12	727	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYDMSWVRQAPGKGLEWVSVISGSG GSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKGPLVGWYFDLW GQGT LVT VSS
1-H13	728	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYAMNWRQAPGKGLEWVSAISGSG GSTFYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCTRPPYGDYGDYWG QGT LVT VSS
1-H14	729	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSQAMSWVRQAPGKGLEWVSAISGSG GGTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKDRRFGEFDPWG QGT LVT VSS
1-H16	730	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYDMSWVRQAPGKGLEWVSVISGSG GSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKGPLVGWYFDLW GQGT LVT VSS
1-H17	731	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYDMSWVRQAPGKGLEWVSVISGSG GSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKGPLVGWYFDLW GQGT LVT VSS
1-H18	732	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYAMYWVRQAPGKGLEWVSAISGSG GSTFYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCTRPPYGDYGDYWG QGT LVT VSS
1-H19	733	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYAMIWVRQAPGKGLEWVSAISGSG GSTFYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCTRPPYGDYGDYWG QGT LVT VSS
1-H20	734	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYAMNWRQAPGKGLEWVSIISGSG GSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAREGYRDYLDWYFDL WGQGT LVT VSS
1-H23	735	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSQAMSWVRQAPGKGLEWVSAISGSG GGTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKDRRFGEFDPWG QGT LVT VSS
1-H25	736	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYGMSWVRQAPGKGLEWVSAISGSG GGTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCASSSSWQFDYWGQ GTLVT VSS
1-H26	737	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYDMSWVRQAPGKGLEWVSVISGSG GSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKGPLVGWYFDLW GQGT LVT VSS
1-H27	738	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYAIWVRQAPGKGLEWVSAISGSGY STYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARGAPDSSGYFQGE VYFDYWGQGT LVT VSS
1-H28	739	EVQLLESGGGLVQPGGSLRLSCAASGVT FSSYAMSWVRQAPGKGLEWVSAITGSG GSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKHGSGTIFGVVIA KYYFDYWGQGT LVT VSS
1-H31	740	EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYAMYWVRQAPGKGLEWVSAISGSA GSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARDTNDFWSGYSIF HPWGQGT LVT VSS
1-H36	741	EVQLLESGGGLVQPGGSLRLSCAASGITFSSYAMSWVRQAPGKGLEWVSGISGSG

		GSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCANHGS GTIFGVVIA KYYFDYWGQGLTVTVSS
1-H37	742	EVQLLESGGGLVQPGGSLRLS CAASGFTFSSQAMSWVRQAPGKGLEWVSAISGSG GGTYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKDRRFGEFDPWG QGLTVTVSS
1-H38	743	EVQLLESGGGLVQPGGSLRLS CAASGFTFSSYAMYWVRQAPGKGLEWVSAISGSA GSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARDTNDFWSGYSIF DPWGQGLTVTVSS
1-H39	744	EVQLLESGGGLVQPGGSLRLS CAASGFTFSSHAMYWVRQAPGKGLEWVSAISGSA GSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARDTKDFWSGYSIF DPWGQGLTVTVSS
1-H40	745	EVQLLESGGGLVQPGGSLRLS CAASGITFSSYAMSWVRQAPGKGLEWVSGISGSG GSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAKHGSGTIFGVVIA KYYFDYCGQGLTVTVSS
1-H41	746	EVQLLESGGGLVQPGGSLRLS CAASGLTFSSYAMSWVRQAPGKGLEWVSDISGSG GSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCVKGTIPIFGVIRSAF DYWGQGLTVTVSS
1-H42	747	EVQLLESGGGLVQPGGSLRLS CAASGFTFSSYAMTWVRQAPGKGLEWVSAISGSG GGTYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCASSSSWQFDYWGQ GTLTVTVSS
1-H43	748	EVQLLESGGGLVQPGGSLRLS CAASGFTFSSHAMYWVRQAPGKGLEWVSAISGSA GSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAHDTKDFWSGYCIF DPWGQGLTVTVSS
1-H44	749	EVQLLESGGGLVQPGGSLRLS CAASGFTFSSYAMYWVRQAPGKGLEWVSAISGSA GSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARDTNDFWGQGLT TVTVSS
1-H47	750	EVQLLESGGGLVQPGGSLRLS CAASGFTFSSYAMYWVRQAPGKGLEWVSAISSSG GSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCTRPYGDYGDYWG QGLTVTVSS
1-H48	751	EVQLLESGGGLVQPGGSLRLS CAASGFTFTSYAMYWVRQAPGKGLEWVSAISGSA GSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARDTNDFWSGYSIF DPWGQGLTVTVSS
1-H49	752	EVQLLESGGGLVQPGGSLRLS CAASGFTFSDYAMNWVRQAPGKGLEWVSIISGSG GSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAREGYRDYLDYFD LWGQGLTVTVSS
1-L1	753	QSALTQPASVSGSPGQSITISCTGTSSDVGSNNLVSWYQQHPGKAPKLMIEG NKR PSGVSNRFGSKSGNTASLTISGLQAED EADYYCCSYAGSVVFGGGTKLTVL
1-L2	754	QSALTQPASVSGSPGQSITISCTGTSTDVGSYNLVSWYQQHPGKAPKLMIEGSQR PSGVSNRFGSKSGNTASLTISGLQAED EADYYCCSYAGSSTVFGGGTKLTVL
1-L3	755	QSALTQPASVSGSPGQSITISCTGTSRDVGSYNLVSWYQQHPGKAPKLMIEGTRK PSGVSNRFGSKSGNTASLTISGLQAED EADYYCCSYAGSRTPVVFVFGGGTKLTVL
1-L4	756	QSALTQPASVSGSPGQSITISCTGTSSDVGSYSLVSWYQQHPGKAPKLMIEGTRK SGVSNRFGSKSGNTASLTISGLQAED EADYYCCSYAGSYVVFVFGGGTKLTVL
1-L5	757	QSALTQPASVSGSPGQSITISCTGTSSDIGSYNLVSWYQQHPGKAPKLMIEGTRK SGVSNRFGSKSGNTASLTISGLQAED EADYYCCSYAGSRTYVFVFGGGTKLTVL
1-L6	758	QSALTQPASVSGSPGQSITISCTGTSSYVGHYNLVSWYQQHPGKAPKLMIEGSRR PSGVSNRFGSKSGNTASLTISGLQAED EADYYCCSYAGSYTHYVFVFGGGTKLTVL
1-L8	759	QSALTQPASVSGSPGQSITISCTGTSSDVGHYNLVSWYQQHPGKAPKLMIEGTRK PSGVSNRFGSKSGNTASLTISGLQAED EADYYCCSYAGSSSFVVFVFGGGTKLTVL
1-L9	760	QSALTQPASVSGSPGQSITISCTGTSSGVGSYNLVSWYQQHPGKAPKLMIEGIKRP SGVSNRFGSKSGNTASLTISGLQAED EADYYCCSYAGSSSFVVFVFGGGTKLTVL
1-L10	761	QSALTQPASVSGSPGQSITISCTGISSDVGSYNLVSWYQQHPGKAPKLMIEGTRK SGVSNRFGSKSGNTASLTISGLQAED EADYYCCSYAGSSSFVVFVFGGGTKLTVL
1-L11	762	QSALTQPASVSGSPGQSITISCTGTSDVGSYNLVSWYQQHPGKAPKLMIEG NKR PSGVSNRFGSKSGNTASLTISGLQAED EADYYCCSYAGSYTVFVFGGGTKLTVL

1-L12	763	QSALTQPASVSGSPGQSITISCTGTSSDFGSYNLVSWSYQQHPGKAPKLMIEGNKR PSGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSSFTVVFVGGGTCLTVL
1-L13	764	QSALTQPASVSGSPGQSITISCTGTSSDVGKYNLVSWSYQQHPGKAPKLMIEGSQR PSGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSRTPVVFVGGGTCLTVL
1-L14	765	QSALTQPASVSGSPGQSITISCTGTSSDVGTYNLVSWSYQQHPGKAPKLMIEGSQR PSGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSSSFVVFVGGGTCLTVL
1-L16	766	QSALTQPASVSGSPGQSITISCTGTSSNDVGSYNLVSWSYQQHPGKAPKLMIEGSKR PSGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSRYVVFVGGGTCLTVL
1-L17	767	QSALTQPASVSGSPGQSITISCTGTSSDVGDYNLVSWSYQQHPGKAPKLMIEGSNR PSGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSSSFVVFVGGGTCLTVL
1-L18	768	QSALTQPASVSGSPGQSITISCTGTSSDVGGYNLVSWSYQQHPGKAPKLMIEGNKR PSGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSSFTVVFVGGGTCLTVL
1-L19	769	QSALTQPASVSGSPGQSITISCTGTSSDVGSYNLLSWYQQHPGKAPKLMIEASKRP SGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSRTPVVFVGGGTCLTVL
1-L20	770	QSALTQPASVSGSPGQSITISCTGTSSDVGYNLVSWSYQQHPGKAPKLMIEGNKR PSGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGGSSVFVGGGTCLTVL
1-L23	771	QSALTQPASVSGSPGQSITISCTGTSSVGSYNLVSWSYQQHPGKAPKLMIEASKR PSGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSYTFAVFVGGGTCLTVL
1-L25	772	QSALTQPASVSGSPGQSITISCTGTSSVGSYNLVSWSYQQHPGKAPKLMIEGSQR PSGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSSHTYVFVGGGTCLTVL
1-L26	773	QSALTQPASVSGSPGQSITISCTGTSSNVGSYNLVSWSYQQHPGKAPKLMIEGTRK PSGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSSSFVVFVGGGTCLTVL
1-L27	774	QSALTQPASVSGSPGQSITISCTGTSSDVGGYNLVSWSYQQHPGKAPKLMIEGTRK PSGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSFVRSSAHVVFVGGGTCLTVL
1-L28	775	QSALTQPASVSGSPGQSITISCTGTSSDVGRYNLVSWSYQQHPGKAPKLMIEGGKR PSGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYASSSTLVVFVGGGTCLTVL
1-L31	776	QSALTQPASVSGSPGQSITISCTGTSSDIGSYNLVSWSYQQHPGKAPKLMIEGTRK SGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSRTYVFVGGGTCLTVL
1-L36	777	QSALTQPASVSGSPGQSITISCTGTSSDVGHYNLVSWSYQQHPGKAPKLMIEGTRK PSGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSSSFVVFVGGGTCLTVL
1-L37	778	QSALTQPASVSGSPGQSITISCTGTSSVGSYNLVSWSYQQHPGKAPKLMIEGIKRP SGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSSSFVVFVGGGTCLTVL
1-L38	779	QSALTQPASVSGSPGQSITISCTGTSSDIGSYNLVSWSYQQHPGKAPKLMIEGTRK SGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSRTYVFVGGGTCLTVL
1-L39	780	QSALTQPASVSGSPGQSITISCTGTSSDIGSYNLVSWSYQQHPGKAPKLMIEGTNR SGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSRTYVFVGGGTCLTVL
1-L40	781	QSALTQPASVSGSPGQSITISCTGTSSDVGHYNLVSWSYQQHPGKAPKLMIEGTRK PSGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSSSFVVFVGGGTCLTVL
1-L41	782	QSALTQPASVSGSPGQSITISCTGTSSDVGSYSLVSWSYQQHPGKAPKLMIEGDKR PSGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSYVVFVGGGTCLTVL
1-L42	783	QSALTQPASVSGSPGQSITISCTGTSSYVGHYNLVSWSYQQHPGKAPKLMIEGSRR PSGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSYTHYVFVGGGTCLTVL
1-L43	784	QSALTQPASVSGSPGQSITISCTGTSSDIGSYNLVSWSYQQHPGKAPKLMIEGTRK SGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSRTYVFVGGGTCLTVL
1-L44	785	QSALTQPASVSGSPGQSITISCTGTSSDIGSYNLVSWSYQQHPGKAPKLMIEGTRK SGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSRTYVFVGGGTCLTVL
1-L47	786	QSALTQPASVSGSPGQSITISCTGTSSDVGGYNLVSWSYQQHPGKAPKLMIEASKR PSGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSYIPVVFVGGGTCLTVL
1-L48	787	QSALTQPASVSGSPGQSITISCTGTSSDIGSYNLVSWSYQQHPGKAPKLMIEGTRK SGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSRTYVFVGGGTCLTVL
1-L49	788	QSALTQPASVSGSPGQSITISCTGTSSDVGSYNLVSWSYQQHPGKAPKLMIEASKR PSGVSNRFGSGKSGNTASLTISGLQAEDEADYYCCSYAGSSTVFVGGGTCLTVL
1-H51	789	EVQLLESGGGLVQPGGSLRLSCAASGFTFSYAMSWVRQAPGKGLEWVSVISGSG GSTYYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAREGYRDYLVYFD LWGQGTCLTVSS

1-H52	790	EVQLLESGGGLVQPGGSLRLSCAASGFTFSNYAMSWVRQAPGKGLEWVSAISGSA GSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCARVRQGLRRTWYY FDYWGGQGLTVTVSS
1-H53	791	EVQLLESGGGLVQPGGSLRLSCAASGFTFSNYTMSWVRQAPGKGLEWVSVISGSG GSTYYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYCAREGYRDYLVYFD LWGQGLTVTVSS
1-L51	792	QSALTQPASVSGSPGQSITISCTGTSSDVGSYDLVSWYQQHPGKAPKLMIEGSKR PSGVSNRFGSKSGNTASLTISGLQAEDYCCSYAG-SSVVFGGGTKLTVL
1-L52	793	QSALTQPASVSGSPGQSITISCTGTSSDVGSNLSWYQQHPGKAPKLMIEGSKRP SGVSNRFGSKSGNTASLTISGLQAEDYCCSYA-GSLYVFGGGTKLTVL
1-L53	794	QSALTQPASVSGSPGQSITISCTGTSTDVGSYNLSWYQQHPGKAPKLMIEGSKR PSGVSNRFGSKSGNTASLTISGLQAZDEADYCCSYAGSYTSVVFGGGTKLTVL

[00292] 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.

CLAIMS

WHAT IS CLAIMED IS:

1. An antibody or antibody fragment comprising a variable domain, heavy chain region (VH) and a variable domain, light chain region (VL), wherein VH comprises complementarity determining regions CDRH1, CDRH2, and CDRH3, wherein VL comprises complementarity determining regions CDRL1, CDRL2, and CDRL3, and wherein (a) an amino acid sequence of CDRH1 is as set forth in any one of SEQ ID NOs: 1-36 or 217-282; (b) an amino acid sequence of CDRH2 is as set forth in any one of SEQ ID NOs: 37-72 or 283-348; (c) an amino acid sequence of CDRH3 is as set forth in any one of SEQ ID NOs: 73-108 or 349-414; (d) an amino acid sequence of CDRL1 is as set forth in any one of SEQ ID NOs: 109-144 or 415-473; (e) an amino acid sequence of CDRL2 is as set forth in any one of SEQ ID NOs: 145-180 or 415-473; and (f) an amino acid sequence of CDRL3 is as set forth in any one of SEQ ID NOs: 181-216 or 533-591.
2. The antibody or antibody fragment of claim 1, wherein (a) an amino acid sequence of CDRH1 is as set forth in SEQ ID NO: 30; (b) an amino acid sequence of CDRH2 is as set forth in SEQ ID NO: 66; (c) an amino acid sequence of CDRH3 is as set forth in SEQ ID NO: 102; (d) an amino acid sequence of CDRL1 is as set forth in SEQ ID NO: 138; (e) an amino acid sequence of CDRL2 is as set forth in SEQ ID NO: 174; and (f) an amino acid sequence of CDRL3 is as set forth in SEQ ID NO: 210.
3. The antibody or antibody fragment of claim 1, wherein (a) an amino acid sequence of CDRH1 is as set forth in SEQ ID NO: 35; (b) an amino acid sequence of CDRH2 is as set forth in SEQ ID NO: 71; (c) an amino acid sequence of CDRH3 is as set forth in SEQ ID NO: 107; (d) an amino acid sequence of CDRL1 is as set forth in SEQ ID NO: 143; (e) an amino acid sequence of CDRL2 is as set forth in SEQ ID NO: 179; and (f) an amino acid sequence of CDRL3 is as set forth in SEQ ID NO: 215.
4. The antibody or antibody fragment of claim 1, wherein (a) an amino acid sequence of CDRH1 is as set forth in SEQ ID NO: 12; (b) an amino acid sequence of CDRH2 is as set forth in SEQ ID NO: 48; (c) an amino acid sequence of CDRH3 is as set forth in SEQ ID NO: 84; (d) an amino acid sequence of CDRL1 is as set forth in SEQ ID NO: 120; (e) an amino acid sequence of CDRL2 is as set forth in SEQ ID NO: 156; and (f) an amino acid sequence of CDRL3 is as set forth in SEQ ID NO: 192.
5. The antibody or antibody fragment of claim 1, wherein (a) an amino acid sequence of CDRH1 is as set forth in SEQ ID NO: 31; (b) an amino acid sequence of CDRH2 is as set forth in SEQ ID NO: 67; (c) an amino acid sequence of CDRH3 is as set forth in SEQ ID NO: 103; (d) an amino acid sequence of CDRL1 is as set forth in SEQ ID NO: 139; (e) an amino acid sequence of CDRL2 is as set forth in SEQ ID NO: 175; and (f) an amino acid sequence of CDRL3 is as set forth in SEQ ID NO: 211.
6. The antibody or antibody fragment of claim 1, wherein (a) an amino acid sequence of CDRH1 is as set forth in SEQ ID NO: 240; (b) an amino acid sequence of CDRH2 is as set forth in SEQ ID NO: 306; (c) an amino acid sequence of CDRH3 is as set forth in SEQ ID NO: 372; (d) an amino acid sequence of CDRL1 is as set forth in SEQ ID NO: 437; (e) an amino acid sequence of CDRL2 is as set forth in SEQ ID NO: 496; and (f) an amino acid sequence of CDRL3 is as set forth in SEQ ID NO: 555.

7. The antibody or antibody fragment of claim 1, wherein (a) an amino acid sequence of CDRH1 is as set forth in SEQ ID NO: 244; (b) an amino acid sequence of CDRH2 is as set forth in SEQ ID NO: 310; (c) an amino acid sequence of CDRH3 is as set forth in SEQ ID NO: 376; (d) an amino acid sequence of CDRL1 is as set forth in SEQ ID NO: 437; (e) an amino acid sequence of CDRL2 is as set forth in SEQ ID NO: 496; and (f) an amino acid sequence of CDRL3 is as set forth in SEQ ID NO: 555.
8. The antibody or antibody fragment of claim 1, wherein (a) an amino acid sequence of CDRH1 is as set forth in SEQ ID NO: 270; (b) an amino acid sequence of CDRH2 is as set forth in SEQ ID NO: 336; (c) an amino acid sequence of CDRH3 is as set forth in SEQ ID NO: 402; (d) an amino acid sequence of CDRL1 is as set forth in SEQ ID NO: 461; (e) an amino acid sequence of CDRL2 is as set forth in SEQ ID NO: 520; and (f) an amino acid sequence of CDRL3 is as set forth in SEQ ID NO: 579.
9. The antibody or antibody fragment of claim 1, wherein the antibody or antibody fragment binds to a spike glycoprotein.
10. The antibody or antibody fragment of claim 1, wherein the antibody or antibody fragment binds to a receptor binding domain of the spike glycoprotein.
11. The antibody or antibody fragment of claim 1, wherein the antibody or antibody fragment comprises a K_D of less than 50 nM.
12. The antibody or antibody fragment of claim 1, wherein the antibody or antibody fragment comprises a K_D of less than 25 nM.
13. The antibody or antibody fragment of claim 1, wherein the antibody or antibody fragment comprises a K_D of less than 10 nM.
14. The antibody or antibody fragment of claim 1, wherein the antibody or antibody fragment comprises a K_D of less than 5 nM.
15. An antibody or antibody fragment comprising a variable domain, heavy chain region (VH) and a variable domain, light chain region (VL), wherein the VH comprises an amino acid sequence at least about 90% identical to a sequence as set forth in any one of SEQ ID NOs: 592-657, and wherein the VL comprises an amino acid sequence at least about 90% identical to a sequence as set forth in any one of SEQ ID NOs: 658-716.
16. The antibody or antibody fragment of claim 15, wherein the VH comprises an amino acid sequence at least about 90% identical to a sequence as set forth in SEQ ID NO: 594, and wherein the VL comprises an amino acid sequence at least about 90% identical to a sequence as set forth in SEQ ID NO: 660.
17. The antibody or antibody fragment of claim 15, wherein the VH comprises an amino acid sequence at least about 90% identical to a sequence as set forth in SEQ ID NO: 595, and wherein the VL comprises an amino acid sequence at least about 90% identical to a sequence as set forth in SEQ ID NO: 661.
18. The antibody or antibody fragment of claim 15, wherein the VH comprises an amino acid sequence at least about 90% identical to a sequence as set forth in SEQ ID NO: 598, and wherein the VL comprises an amino acid sequence at least about 90% identical to a sequence as set forth in SEQ ID NO: 664.

19. The antibody or antibody fragment of claim 15, wherein the VH comprises an amino acid sequence at least about 90% identical to a sequence as set forth in SEQ ID NO: 603, and wherein the VL comprises an amino acid sequence at least about 90% identical to a sequence as set forth in SEQ ID NO: 669.
20. The antibody or antibody fragment of claim 15, wherein the VH comprises an amino acid sequence at least about 90% identical to a sequence as set forth in SEQ ID NO: 615, and wherein the VL comprises an amino acid sequence at least about 90% identical to a sequence as set forth in SEQ ID NO: 680.
21. The antibody or antibody fragment of claim 15, wherein the VH comprises an amino acid sequence at least about 90% identical to a sequence as set forth in SEQ ID NO: 631, and wherein the VL comprises an amino acid sequence at least about 90% identical to a sequence as set forth in SEQ ID NO: 691.
22. The antibody or antibody fragment of claim 15, wherein the VH comprises an amino acid sequence at least about 90% identical to a sequence as set forth in SEQ ID NO: 645, and wherein the VL comprises an amino acid sequence at least about 90% identical to a sequence as set forth in SEQ ID NO: 704.
23. An antibody, wherein the antibody comprises a sequence comprising at least 90% sequence identity to any one of SEQ ID NOs: 1-716; and wherein the antibody is a monoclonal antibody, a polyclonal antibody, a bi-specific antibody, a multispecific antibody, a grafted antibody, a human antibody, a humanized antibody, a synthetic antibody, a chimeric antibody, a camelized antibody, a single-chain Fvs (scFv), a single chain antibody, a Fab fragment, a F(ab')₂ fragment, a Fd fragment, a Fv fragment, a single-domain antibody, an isolated complementarity determining region (CDR), a diabody, a fragment comprised of only a single monomeric variable domain, disulfide-linked Fvs (sdFv), an intrabody, an anti-idiotypic (anti-Id) antibody, or ab antigen-binding fragments thereof.
24. A nucleic acid composition comprising: a) a first nucleic acid encoding a variable domain, heavy chain region (VH) comprising an amino acid sequence at least about 90% identical to a sequence as set forth in any one of SEQ ID NOs: 592-657; b) a second nucleic acid encoding a variable domain, light chain region (VL) comprising at least about 90% identical to a sequence as set forth in any one of SEQ ID NOs: 658-716; and an excipient.
25. A method of treating a SARS-CoV-2 infection, comprising administering the antibody or antibody fragment of any one of claims 1-23.
26. The method of claim 25, wherein the antibody is administered prior to exposure to SARS-CoV-2.
27. The method of claim 26, wherein the antibody is administered at least about 1 week prior to exposure to SARS-CoV-2.
28. The method of claim 26, wherein the antibody is administered at least about 1 month prior to exposure to SARS-CoV-2.
29. The method of claim 26, wherein the antibody is administered at least about 5 months prior to exposure to SARS-CoV-2.
30. The method of claim 25, wherein the antibody is administered after exposure to SARS-CoV-2.
31. The method of claim 30, wherein the antibody is administered at most about 24 hours after exposure to SARS-CoV-2.

32. The method of claim 30, wherein the antibody is administered at most about 1 week after exposure to SARS-CoV-2.
33. The method of claim 30, wherein the antibody is administered at most about 1 month after exposure to SARS-CoV-2.
34. A method of treating an individual with a SARS-CoV-2 infection with the antibody or antibody fragment of any one of claims 1-23 comprising:
- (a) obtaining or having obtained a sample from the individual;
 - (b) performing or having performed an expression level assay on the sample to determine expression levels of SARS-CoV-2 antibodies; and
 - (c) if the sample has an expression level of the SARS-CoV-2 antibodies then administering to the individual the antibody or antibody fragment of any one of claims 1-23, thereby treating the SARS-CoV-2 infection.
35. A method for optimizing an antibody comprising:
- (a) providing a plurality of polynucleotide sequences encoding for an antibody or antibody fragment, wherein the antibody or antibody fragment is derived from a subject having SARS-CoV-2;
 - (b) generating a nucleic acid library comprising the plurality of sequences that when translated encode for antibodies or antibody fragments that bind SARS-CoV-2 or ACE2 protein, wherein each of the sequences comprises a predetermined number of variants within a CDR relative to an input sequence that encodes an antibody; wherein the library comprises at least 50,000 variant sequences; and
 - (c) synthesizing the at least 50,000 variant sequences.
36. The method of claim 35, wherein the antibody library comprises at least 100,000 sequences.
37. The method of claim 35, wherein the method further comprises enriching a subset of the variant sequences.
38. The method of claim 35, wherein the method further comprises expressing the antibody or antibody fragments corresponding to the variant sequences.
39. The method of claim 35, wherein the polynucleotide sequence is a murine, human, or chimeric antibody sequence.
40. The method of claim 35, wherein each sequence of the plurality of variant sequences comprises at least one variant in each CDR of a heavy chain or light chain, relative to the input sequence.
41. The method of claim 35, wherein each sequence of the plurality of variant sequences comprises at least two variants in each CDR of a heavy chain or light chain relative to the input sequence.
42. The method of claim 35, wherein at least one sequence when translated encodes for an antibody or antibody fragment having at least 5X higher binding affinity than a binding affinity of the input sequence.
43. The method of claim 35, wherein at least one sequence when translated encodes for an antibody or antibody fragment having at least 25X higher binding affinity than a binding affinity of the input sequence.

44. The method of claim 35, wherein at least one sequence when translated encodes for an antibody or antibody fragment having at least 50X higher binding affinity than a binding affinity of the input sequence.
45. The method of claim 35, wherein each sequence comprises at least one variant in each CDR of a heavy chain or light chain relative to a germline sequence of the input sequence.
46. The method of claim 35, wherein the nucleic acid library has a theoretical diversity of at least 10^{10} sequences.
47. The method of claim 35, wherein the nucleic acid library has a theoretical diversity of at least 10^{12} sequences.