



(51) International Patent Classification:

C07K 16/10 (2006.01)

A61P 31/14 (2006.01)

A61K 39/42 (2006.01)

(21) International Application Number:

PCT/US2021/041328

(22) International Filing Date:

12 July 2021 (12.07.2021)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

63/052,140

15 July 2020 (15.07.2020)

US

63/075,701

08 September 2020 (08.09.2020)

US

(71) Applicant: **REGENTS OF THE UNIVERSITY OF MINNESOTA** [US/US]; 600 McNamara Alumni Center, 200 Oak Street SE, Minneapolis, Minnesota 55455-2020 (US).

(72) Inventors; and

(71) Applicants: **LEE, Frank Jonathan** [US/US]; 600 McNamara Alumni Center, 200 Oak Street SE, Minneapolis, Minnesota 55455-2020 (US). **YE, Gang** [CN/US]; 600 Mc-

Namara Alumni Center, 200 Oak Street SE, Minneapolis, Minnesota 55455-2020 (US). **SHANG, Jian** [CN/CN]; 600 McNamara Alumni Center, 200 Oak Street SE, Minneapolis, Minnesota 55455-2020 (US). **WAN, Yushun** [CN/CN]; 600 McNamara Alumni Center, 200 Oak Street SE, Minneapolis, Minnesota 55455-2020 (US). **LEBEAU, Aaron** [US/US]; 600 McNamara Alumni Center, 200 Oak Street SE, Minneapolis, Minnesota 55455-2020 (US). **GALLANT, Joseph** [US/US]; 600 McNamara Alumni Center, 200 Oak Street SE, Minneapolis, Minnesota 55455-2020 (US).

(74) Agent: **KRUEGER, Katherine A.D.** et al.; Viksnins Harris Pady's Malen LLP, 7851 Metro Parkway, Suite 325, Bloomington, Minnesota 55425 (US).

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

(54) Title: SARS-CoV-2 NANOBODIES AND METHODS OF USE THEREOF

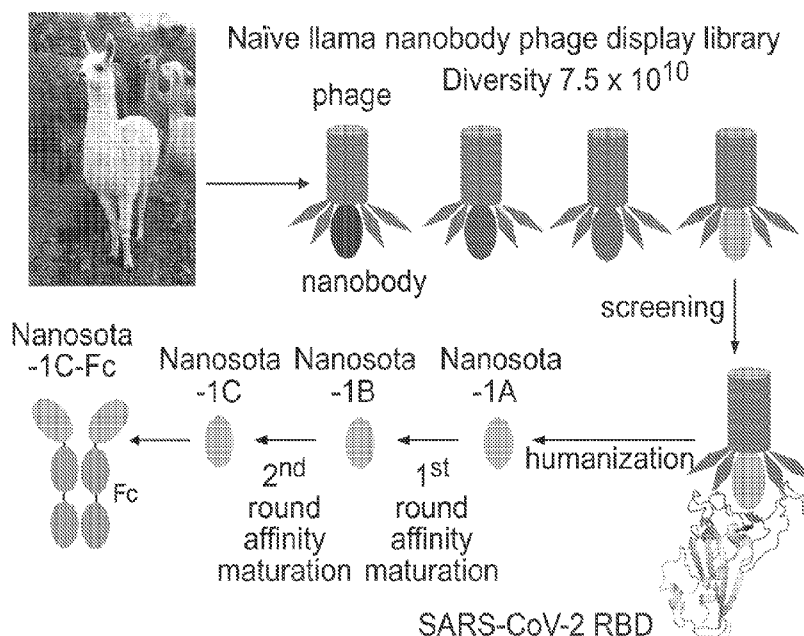


FIGURE 6

(57) Abstract: Certain embodiments of the invention provide isolated anti-SARS-CoV-2 nanobodies, as well as polypeptides and protein molecules comprising such nanobodies. Certain embodiments of the invention also provide methods of using these nanobodies, polypeptides and protein molecules for treating or preventing a SARS-CoV-2 infection.

SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN,
TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

- (84) Designated States** (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

- *with international search report (Art. 21(3))*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

SARS-CoV-2 NANOBODIES AND METHODS OF USE THEREOF

CROSS REFERENCE TO RELATED APPLICATIONS

This application claims priority to United States Provisional Application Number 63/052,140 filed on July 15, 2020 and United States Provisional Application Number 63/075,701 filed on September 8, 2020. The entire content of the applications referenced above is hereby incorporated by reference herein.

GOVERNMENT FUNDING

This invention was made with government support under AI157975, AI089728, GM118047, GM124165, and OD021527 awarded by the National Institutes of Health, DE-AC02-06CH11357 awarded by the Department of Energy, and 2029943 awarded by the National Science Foundation. The government has certain rights in the invention.

BACKGROUND OF THE INVENTION

Coronaviruses, which cause disease in mammals and birds, are a group of enveloped viruses that have a positive-sense single-stranded RNA genome and a nucleocapsid of helical symmetry. Their genome encodes four major structural proteins: spike (S), membrane (M), envelope (E) and nucleocapsid (N). In humans, coronaviruses cause respiratory tract infections that can range from mild to lethal. Mild illnesses include some cases of the common cold, while more lethal varieties can cause COVID-19, SARS and MERS. SARS-coronavirus 2 (SARS-CoV-2) causes COVID-19, a disease that has spread rapidly and created a global health emergency. Due to the recent emergence of SARS-CoV-2, there are very few drugs available to treat SARS-CoV-2 infections.

Generally, small molecule drugs have relatively low specificity due to their size, which results in off-target side effects. In contrast, due to their large size, traditional antibody drugs have relatively poor pharmacokinetics, are relatively unstable, and are expensive to produce in large quantities. Thus, there is a need for new approaches and new therapeutic agents for preventing or ameliorating SARS-CoV-2 infections.

SUMMARY OF THE INVENTION

Certain embodiments of the invention provide an isolated anti-SARS-CoV-2 nanobody comprising one or more complementarity determining regions (CDRs) selected from the group

consisting of:

(a) a CDR1 comprising an amino acid sequence having at least 80% sequence identity to an amino acid sequence of GFTFKNAD (SEQ ID NO:2);

5 (b) a CDR2 comprising an amino acid sequence having at least 80% sequence identity to an amino acid sequence of IYSDG(S/R)T (SEQ ID NO:20); and

(c) a CDR3 comprising an amino acid sequence having at least 80% sequence identity to an amino acid sequence of MAGSKSG(Q/H)ELDH (SEQ ID NO:21).

Certain embodiments of the invention provide a nanobody-Fc fusion polypeptide comprising an isolated anti-SARS-CoV-2 nanobody as described herein operably linked to an Fc
10 domain amino acid sequence.

Certain embodiments of the invention provide a protein molecule comprising two independently selected nanobody-Fc fusion polypeptides as described herein, wherein the two Fc polypeptides are linked to form a dimer.

Certain embodiments of the invention provide a composition comprising an isolated anti-
15 SARS-CoV-2 nanobody as described herein, a nanobody-Fc fusion polypeptide as described herein, or a protein molecule as described herein, and a carrier.

Certain embodiments of the invention provide an isolated polynucleotide comprising a nucleotide sequence encoding an isolated anti-SARS-CoV-2 nanobody as described herein or a nanobody-Fc fusion polypeptide as described herein.

20 Certain embodiments of the invention provide a vector comprising the polynucleotide as described herein.

Certain embodiments of the invention provide a cell comprising the polynucleotide as described herein or the vector as described herein.

Certain embodiments of the invention provide a method of detecting the presence of
25 SARS-CoV-2 in a cell, the method comprising contacting the cell with an isolated anti-SARS-CoV-2 nanobody as described herein, a nanobody-Fc fusion polypeptide as described herein, or a protein molecule as described herein, and detecting whether a complex is formed between 1) the nanobody, polypeptide or protein molecule; and 2) SARS-CoV-2.

Certain embodiments of the invention provide a method of inhibiting the activity of
30 SARS-CoV-2, comprising contacting SARS-CoV-2 with an isolated anti-SARS-CoV-2 nanobody as described herein, a nanobody-Fc fusion polypeptide as described herein, or the protein molecule as described herein.

Certain embodiments of the invention provide a method for treating or preventing a SARS-CoV-2 infection in a mammal, comprising administering an effective amount of an

isolated anti-SARS-CoV-2 nanobody as described herein, a nanobody-Fc fusion polypeptide as described herein, or a protein molecule as described herein, to the mammal.

Certain embodiments of the invention provide an isolated anti-SARS-CoV-2 nanobody as described herein, a nanobody-Fc fusion polypeptide as described herein, or a protein molecule as described herein, for the prophylactic or therapeutic treatment of a SARS-CoV-2 infection.

Certain embodiments of the invention provide an isolated anti-SARS-CoV-2 nanobody as described herein, a nanobody-Fc fusion polypeptide as described herein, or a protein molecule as described herein, for use in medical therapy.

Certain embodiments of the invention provide a kit comprising:

1) an isolated anti-SARS-CoV-2 nanobody as described herein, a nanobody-Fc fusion polypeptide as described herein, or a protein molecule as described herein;

2) packaging material; and

3) instructions for administering the nanobody, polypeptide or protein molecule, to a mammal to treat or prevent a SARS-CoV-2 infection.

The invention also provides processes and intermediates disclosed herein that are useful for preparing anti-SARS-CoV-2 binding molecules described herein (e.g., nanobodies, polypeptides, and protein molecules), as well as compositions described herein.

BRIEF DESCRIPTION OF THE FIGURES

Figures 1A-1E. Biacore measurement data. For the experiments, SARS-CoV-2 RBD was covalently immobilized (via amine group) to a sensor chip. Fig. 1A. Biacore measurement of the binding interactions between the Nanosota-1 drugs and SARS-CoV-2 RBD. Fig. 1B. Nanosota-1C-Fc binding to SARS-CoV-2 RBD by SPR using Biacore. Nanosota-1C-Fc flowed by at the following concentrations: 1.25 nM, 2.5 nM, 5 nM, 10 nM, and 20 nM. Fig. 1C. Nanosota-1C binding to SARS-CoV-2 RBD by SPR using Biacore. Nanosota-1C flowed by at the following concentrations: 2.5 nM, 5 nM, 10 nM, 20 nM, 40 nM, and 80 nM. Fig. 1D. Nanosota-1B binding to SARS-CoV-2 RBD by SPR using Biacore. Nanosota-1B flowed by at the following concentrations: 2.5 nM, 5 nM, 10 nM, 20 nM, 40 nM, and 80 nM. Fig. 1E. Nanosota-1A binding to SARS-CoV-2 RBD by SPR using Biacore. Nanosota-1A flowed by at the following concentrations: 10 nM, 20 nM, 40 nM, 80 nM, 160 nM, and 320 nM.

Figures 2A-2B. Inhibition of SARS-CoV-2 pseudovirus entry by Nanosota-1. Entry efficiency was characterized as luciferase signal accompanying entry. Error bars represent SEM (N=4).

Figures 3A-3B. Competitive Binding. Fig. 3A. Pull down of hACE2 and Nanosota-1C

from solution. Fig. 3B. Competitive binding of Nanosota-1C and human ACE2 (hACE2) to SARS-CoV-2 RBD as detected by gel filtration chromatography.

Figures 4A-4C. Bioavailability measurements. Fig. 4A. Bioavailability of Nanosota-1C-Fc in mice. Fig. 4B. Bioavailability of Nanosota-1C in mice. Fig. 4C. Bioavailability of PBS control in mice. The data are presented as mean A450 values; error bars represent the SEM of the mean (n = 3, meaning that there were three mice in each group).

Figure 5. Stability of Nanosota-1C-Fc at different temperatures. Nanosota-1C-Fc was stored at different temperature for 1 week, and then was assayed for its ability to bind to SARS-CoV-2 RBD (containing a C-terminal His tag) using ELISA.

Figure 6. Construction of a camelid nanobody phage display library and use of this library for screening of anti-SARS-CoV-2 nanobodies. A large-sized (diversity 7.5×10^{10}), naïve nanobody phage display library was constructed using B cells of over a dozen llamas and alpacas. Phages were screened for their high binding affinity for SARS-CoV-2 RBD. Nanobodies expressed from the selected phages were further screened for their potency in neutralizing SARS-CoV-2 pseudovirus entry. The best performing nanobody contained two humanization mutations and subjected to two rounds of affinity maturation.

Figures 7A-7B. Binding interactions among *Nanosota-1* drugs, human ACE2 and SARS-CoV-2 RBD. (Fig. 7A) Protein pull-down assay to assess binding among *Nanosota-1C*, ACE2, and RBD. *Nanosota-1C* and ACE2 (both His tagged) were mixed in different molar ratios, with the concentration of ACE2 remaining constant. RBD (Fc tagged) was used to pull down *Nanosota-1C* and ACE2 from solution. Western blot was performed to detect the amounts of *Nanosota-1C* and ACE2 pulled down by RBD. The assay was repeated three times (biological replication: new aliquots of proteins were used for each repeat). (Fig. 7B) Gel filtration chromatography to assess binding among *Nanosota-1C*, ACE2, and RBD. *Nanosota-1C*, ACE2 and RBD (all His tagged) were mixed together in solution, with both *Nanosota-1C* and ACE2 in molar excess of RBD, and then subjected to gel filtration chromatography. Protein components in each of the gel filtration chromatography peaks were analyzed with SDS-PAGE and stained by Coomassie blue. The assay was repeated three times (biological replication: new aliquots of proteins were used for each repeat).

Figures 8A-8B. Efficacy of *Nanosota-1* drugs in neutralizing SARS-CoV-2 infections. (Fig. 8A) Neutralization of SARS-CoV-2 pseudovirus entry into target cells by one of three inhibitors: *Nanosota-1C-Fc*, *Nanosota-1C*, and recombinant human ACE2. Retroviruses pseudotyped with SARS-CoV-2 spike protein (i.e., SARS-CoV-2 pseudoviruses) were used to enter HEK293T cells expressing human ACE2 in the presence of the inhibitor at different

concentrations. Entry efficiency was characterized as luciferase signal accompanying entry. Data are the mean $\pm$ SEM ($n = 4$). Nonlinear regression was performed using a log (inhibitor) versus normalized response curve and a variable slope model ($R^2 > 0.95$ for all curves). The efficacy of each inhibitor was expressed as the concentration capable of neutralizing 50% of the entry efficiency (i.e., 50% Neutralizing Dose or ND_{50}) or neutralizing 90% of the entry efficiency (i.e., ND_{90}). ND_{90} for ACE2 was not calculated due to insufficient data points. The assay was repeated three times (biological replication: new aliquots of pseudoviruses and cells were used for each repeat). (Fig. 8B) Neutralization of authentic SARS-CoV-2 infection of target cells by the same inhibitors as above. The potency of *Nanosota-1* drugs in neutralizing authentic SARS-CoV-2 infections was evaluated using a SARS-CoV-2 micro-neutralization assay. 100 infectious SARS-CoV-2 particles were used to infect Vero E6 cells in the presence of the inhibitor at different concentrations. Infection was characterized as the virus-induced formation of cytopathic effect (CPE). The efficacy of each inhibitor was expressed as the lowest concentration capable of completely preventing virus-induced CPE in 100% of the wells (i.e., 100% Neutralizing Dose or ND_{100}) or 50% of the wells (i.e., ND_{50}).

Figures 9A-9B. Efficacy of *Nanosota-1* drugs in protecting hamsters from SARS-CoV-2 infections. Four groups of hamsters (6 per group) were injected with a single dose of *Nanosota-1C-Fc* at the indicated time point and the indicated dosage. Along with a negative (control) group, they were challenged with SARS-CoV-2 (at a titer of 10^6 Median Tissue Culture Infectious Dose or $TCID_{50}$) on day 0. (Fig. 9A) Body weights of hamsters were monitored on each day and percent change in body weight relative to day 0 was calculated for each hamster. Data are the mean $\pm$ 1 SEM ($n = 6$). ANOVA on group as a between-group factor and day (1-10) as a within-group factor revealed significant differences between the control group and each of the following groups: 24 hour pre-challenge (20 mg/kg) group ($F(1, 10) = 17.80$, $p = .002$; effect size $\eta_p^2 = .64$), 4 hour post-challenge (20 mg/kg) group ($F(1, 10) = 5.02$, $p = .035$; $\eta_p^2 = .37$), and 4 hour post-challenge (10 mg/kg) group ($F(1, 10) = 7.04$, $p = .024$, $\eta_p^2 = .41$). The 24 hour post-challenge group showed a similar but insignificant trend ($F < 1$). All p -values are two-tailed. (Fig. 9B) Tissues of bronchial tubes from each of the hamsters were collected on day 10 and scored for the severity of bronchioloalveolar hyperplasia: 3 - moderate; 2 - mild; 1 - minimum; 0 - none. Data are the mean $\pm$ 1 SEM ($n = 6$). A comparison between the control group and each of the other groups was performed using Student's t-test. *** $p < .001$; ** $p < .01$; * $p < .05$; † $p < .10$ one-tailed.

Figures 10A-10D. Cost-effectiveness and pharmacokinetics of *Nanosota-1* drugs. (Fig. 10A) Purification of *Nanosota-1C-Fc* from bacteria. The protein was nearly 100% pure after gel

filtration chromatography, as demonstrated by its elution profile and SDS-PAGE (stained by Coomassie blue). The yield of the protein was 40 mg per liter bacterial culture, without any optimization of the expression. (Fig. 10B) *In vitro* stability of *Nanosota-1C-Fc*. The protein was stored at one of the indicated temperatures for a week, and was then subjected to ELISA for

5 detection of its remaining SARS-CoV-2 RBD-binding capability. For ELISA, RBD (His tagged) was coated, serially diluted *Nanosota-1C-Fc* was added, and binding was detected using anti-Fc antibody targeting the Fc tag of *Nanosota-1C-Fc*. Data are the mean $\pm$ 1 SEM (n = 4). (Fig. 10C) *In vivo* stability of *Nanosota-1C-Fc*. *Nanosota-1C-Fc* was injected into mice, mouse sera were collected at different time points, and *Nanosota-1C-Fc* remaining in the sera was detected for its

10 SARS-CoV-2 RBD-binding capability using ELISA as described above. Data are the mean $\pm$ 1 SEM (n = 3). (Fig. 10D) Biodistribution of [^{89}Zr]Zr-*Nanosota-1C-Fc*. *Nanosota-1C-Fc* was radioactively labeled with ^{89}Zr and injected into mice via tail vein. Different tissues or organs were collected at different time points (n=3 mice per time point). The amount of *Nanosota-1C-Fc* present in each tissue or organ was measured through examining the radioactivity of each

15 tissue or organ. Data are the mean $\pm$ 1 SEM (n = 3). For each tissue type, the following timepoints are shown from left to right: 6 hrs, 24 hrs., 2 days, and 3 days.

Figure 11. Schematic drawings of conventional antibodies and nanobodies. VH: variable domain of heavy chain. CH: constant domain of heavy chain. VL: variable domain of light chain. CL: constant domain of light chain. VHH: variable domain of heavy-chain only antibody.

20 scFv: single-chain variable fragment.

Figures 12A-12D. Measurement of the binding affinities between *Nanosota-1* drugs and SARS-CoV-2 RBD by surface plasmon resonance assay using Biacore. Purified recombinant SARS-CoV-2 RBD was covalently immobilized on a sensor chip through its amine groups. Purified recombinant nanobodies flowed over the RBD individually at one of five different

25 concentrations. The resulting data were fit to a 1:1 binding model and the value of K_d was calculated for each nanobody. The assay was repeated three times (biological replication: new aliquots of proteins and new sensor chips were used for each repeat).

Figure 13. Neutralization of SARS-CoV-2 pseudovirus, which contains the D614G mutation in the spike protein, by *Nanosota-1* drugs. The procedure was the same as described in

30 Fig. 8A, except that the mutant spike protein replaced the wild type spike protein. The assay was repeated three times (biological replication: new aliquots of pseudoviruses and cells were used for each repeat).

Figure 14. Additional data on the efficacy of *Nanosota-1* drugs in protecting hamsters from SARS-CoV-2 infections. Nasal swabs were collected from each hamster on days 1, 2, 3, 5,

7, and 10. Nasal swab samples from day 2 and day 3 were lost due to Hurricane Laura. qRT-PCR was performed to determine the virus loads in each of the samples. The qRT-PCR results are displayed on a log scale (qRT-PCR amplifies signals on a log scale). Data are the mean $\pm$ 1 SEM (n = 6 per group). Missing data from one animal in the 4-hour post-challenge (10mg/kg) group on Day 7 were replaced by the average of that animal's days 5 and 10 data. ANOVA analysis using group as a between-group factor and day (1, 5, 7, and 10) as a within-group factor revealed significant differences between the control group and each of the following groups: 24 hour pre-challenge (20 mg/kg) group ($F(1, 10) = 6.02, p = .017$, effect size $\eta_p^2 = .38$), 4 hour post-challenge (20 mg/kg) group ($F(1, 10) = 5.38, p = .037, \eta_p^2 = .31$), and 4 hour post-challenge (10 mg/kg) group ($F(1, 10) = 3.40, p = .048, \eta_p^2 = .25$). The 24 hour post-challenge group showed a numerical, but not significant trend of reduced nasal virus titer overall ($F < 1$); the reduction was significant on Day 5 ($t(10) = 2.76, p = .01$). All p -values are one-tailed for directional tests.

Figures 15A-15C. Pharmacokinetics of Nanosota-1C. Fig. 15 (A) *In vivo* stability and biodistribution of *Nanosota-1C* were measured in the same way as described in Fig. 10C and Fig. 10D, respectively, except that time points for *Nanosota-1C* differed from those for *Nanosota-1C-Fc* due to pharmacokinetic differences of the small molecular weight nanobody versus the larger Fc tagged nanobody. Fig. 15 (B) PBS buffer was used as a negative control for the *in vivo* stability experiment. Fig. 15 (C) *Nanosota-1C* was radiolabeled with zirconium-89 and injected systemically into mice. Tissues were collected at various time points and biodistribution of *Nanosota-1C* was quantified using a scintillation counter. For each tissue type, the following timepoints are shown from left to right: 1 hr, 2 hrs., 6 hrs. and 1 day.

Figures 16A-16B: Crystal structure of SARS-CoV-2 RBD complexed with *Nanosota-1C*. Fig.16 (A) Structure of SARS-CoV-2 RBD complexed with *Nanosota-1C*, viewed at two different angles. *Nanosota-1C*, the core structure of RBD, and the receptor-binding motif (RBM) of RBD are shown. Fig.16 (B) Overlay of the structures of the RBD/*Nanosota-1C* complex and RBD/ACE2 complex (PDB 6M0J). The structures of the two complexes were superimposed based on their common RBD structure. The *Nanosota-1C* loops that have clashes with ACE2 are circled.

Figure 17. Footprint of *Nanosota-1C* on SARS-CoV-2 RBD. RBD residues that bind to *Nanosota-1C* are labeled in gray, those that bind to ACE2 are labeled in light gray, and those that bind to both *Nanosota-1C* and ACE2 are labeled in black.

Figures 18A-18D. The binding of *Nanosota-1C* to SARS-CoV-2 spike protein in different conformations. Fig.18 (A) The binding of *Nanosota-1C* to the spike protein in the

closed conformation. The structures of the RBD/*Nanosota-1C* complex and SARS-CoV-2 spike protein in the closed conformation (PDB: 6ZWV) were superimposed based on their common RBD structure. Fig. 18 (B) The binding of ACE2 to the spike protein in the closed conformation. The structures of the RBD/ACE2 complex (PDB 6M0J) and SARS-CoV-2 spike protein in the closed conformation (PDB: 6ZWV) were superimposed based on their common RBD structure. Clashes between ACE2 and the rest of the spike protein were circled. Fig. 18 (C) The binding of *Nanosota-1C* to the spike protein in the open conformation (PDB: 6VSB). Fig. 18 (D) The binding of ACE2 to the spike protein in the open conformation (PDB: 6VSB).

Figure 19. Binding interactions between *Nanosota-1* and SARS-CoV-2 RBD. Binding interactions among *Nanosota-1C-Fc*, ACE2, and SARS-CoV-2 RBD as evaluated using a protein pull-down assay. Various concentrations of *Nanosota-1C-Fc* (Fc tagged) and a constant concentration of ACE2 (His tagged) were combined in different molar ratios. Biotinylated SARS-CoV-2 RBD (His tagged) was used to pull down *Nanosota-1C-Fc* and ACE2. A western blot was used to detect the presence of *Nanosota-1C* and ACE2 following pull down by SARS-CoV-2 RBD. The assay was repeated three times (biological replication: new aliquots of proteins were used for each repeat).

Figure 20. Efficacy of *Nanosota-1* in neutralizing SARS-CoV-2 infections *in vitro*. Neutralization of live SARS-CoV-2 infection of target cells by one of two inhibitors: *Nanosota-1C-Fc* and *Nanosota-1C*. The potency of *Nanosota-1* in neutralizing live SARS-CoV-2 infections was evaluated using a SARS-CoV-2 plaque reduction neutralization test (PRNT) assay. 80 PFU infectious SARS-CoV-2 particles were used to infect Vero E6 cells in the presence of individual inhibitor at various concentrations. Infection was characterized as the number of virus plaques formed in overlaid cells. Images of virus plaques for each inhibitor at the indicated concentrations are shown. Each image represents data from triplications. The efficacy of each inhibitor was calculated and expressed as the concentration capable of reducing the number of virus plaques by 50% (i.e., ND₅₀). The assay was repeated twice (biological replication: new aliquots of virus particles and cells were used for each repeat).

Figure 21. Detailed data on the neutralization of live SARS-CoV-2 infection of target cells by *Nanosota-1*. Data are the mean $\pm$ SEM (n = 3). Nonlinear regression was performed using a log (inhibitor) versus normalized response curve and a variable slope model ($R^2 > 0.95$ for all curves). The assay was repeated twice (biological replication: new aliquots of virus particles and cells were used for each repeat).

Figures 22A-22B. Efficacy of *Nanosota-1* in protecting mice from SARS-CoV-2 infections. Fig. 22 (A) 5 mice from each group were euthanized on day 2 post-challenge, and the

virus titers in their lungs were measured using a plaque assay. A comparison between the control group and each of the other groups was performed using one-tailed Student's t-test for directional tests. Data are the mean $\pm$ SEM (n = 5). $**p < 0.01$; $*p < 0.05$. Fig.22 (B) The remaining 2 mice from each group were euthanized on day 5 post-challenge. Lung tissues were collected and examined for pathological changes after staining with hematoxylin and eosin. PBS control group: severe inflammatory cell infiltration and bronchiole infiltration on the top panel; severe alveolar edema filled with liquid (labeled *) and proliferative alveolar epithelium (labeled $\rightarrow$) on the bottom panel. 24h pre-challenge 20 mg/kg group: close to normal. 24h pre-challenge 10 mg/kg group: minor proliferative alveolar epithelium (labeled $\rightarrow$) and inflammatory cell infiltration. 4h post-challenge 20 mg/kg group: close to normal. 4h post-10 mg/kg group: obvious cell inflammatory cell infiltration and vascular thrombosis (labeled #).

DETAILED DESCRIPTION

Nanobodies are much smaller than traditional antibodies (e.g., about 15kDa vs 150kDa), which provides certain advantages as therapeutics. For example, nanobodies are highly stable, expressed at high yields, easy to store and transport, highly cost effective, capable of accessing/recognizing less exposed epitopes, and are able to penetrate tissues and barriers in the human body efficiently. Moreover, like traditional antibodies, nanobodies can bind antigens with high affinity and specificity. While nanobodies can be cleared by the kidneys due to their small size, the pharmacokinetics and/or effector function of a nanobody can be modulated by constructing nanobody-Fc proteins.

As described herein, using a large-sized and humanized naïve llama nanobody library a series of anti-SARS-CoV-2 nanobody drugs and nanobody-Fc fusion polypeptides were developed. As shown in the Examples, these nanobodies and nanobody-Fc fusion polypeptides bind to the SARS-CoV-2 (COVID-19) spike protein with high affinity and inhibit SARS-CoV-2 pseudovirus entry and authentic SAR-CoV-2 infection in target cells. Additionally, animal testing showed that *Nanosota-1-Fc* protected hamsters from SARS-CoV-2 infections both prophylactically and therapeutically. Finally, these nanobodies and fusion polypeptides can be produced in bacteria in large quantities, are highly stable for storage and transportation, and demonstrate an extended in vivo half-life and high tissue permeability.

Anti-SARS-CoV-2 Binding Molecules

Accordingly, certain embodiments provide anti-SARS-CoV-2 binding molecules (e.g., targeting a SARS-CoV-2 spike protein). In certain embodiments, these binding molecules are single domain antibodies or nanobodies that specifically bind to SARS-CoV-2.

As used herein, the term “nanobody” refers to a single monomeric variable antibody domain, such as a VHH, a humanized VHH or a camelized VH (such as a camelized human VH) or generally a sequence optimized VHH (such as e.g. optimized for chemical stability and/or solubility, maximum overlap with known human framework regions and maximum expression), which is capable of binding to a specific antigen. Nanobodies are also described in WO2008/020079 and WO2009/138519, which publications are incorporated by reference herein for all purposes. The terms “nanobody” and “single domain antibody” are used interchangeably herein.

In some embodiments, an anti-SARS-CoV-2 nanobody, comprises: (1) one or more complementarity determining region (CDR) sequences as described herein; and/or (2) a heavy chain variable region sequence as described herein (e.g., as described in Table 1 below).

In certain embodiments, an isolated anti-SARS-CoV-2 nanobody comprises one or more CDRs selected from the group consisting of:

(a) a CDR1 comprising an amino acid sequence having at least 80% (e.g., 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100%) sequence identity to an amino acid sequence of GFTFKNAD (SEQ ID NO:2);

(b) a CDR2 comprising an amino acid sequence having at least 80% (e.g., 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100%) sequence identity to an amino acid sequence of IYSDG(S/R)T (SEQ ID NO:20); and

(c) a CDR3 comprising an amino acid sequence having at least 80% (e.g., 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100%) sequence identity to an amino acid sequence of MAGSKSG(Q/H)ELDH (SEQ ID NO:21).

In some embodiments, an anti-SARS-CoV-2 nanobody comprises two, or three CDRs as described above (e.g., each CDR is selected from one of (a)-(c)).

In certain embodiments, the anti-SARS-CoV-2 nanobody as described herein comprises one or more CDRs selected from the group consisting of:

(a) a CDR1 comprising the amino acid sequence of SEQ ID NO:2;

- (b) a CDR2 comprising the amino acid sequence of SEQ ID NO:20; and
- (c) a CDR3 comprising the amino acid sequence of SEQ ID NO:21.

In some embodiments, an anti-SARS-CoV-2 nanobody comprises two or three CDRs as described above (e.g., each CDR is selected from one of (a)-(c)).

5 For example, in certain embodiments, the anti-SARS-CoV-2 nanobody as described herein comprises:

- (a) a CDR1 comprising the amino acid sequence of SEQ ID NO:2;
- (b) a CDR2 comprising the amino acid sequence of SEQ ID NO:20; and
- (c) a CDR3 comprising the amino acid sequence of SEQ ID NO:21.

10 In certain embodiments, the isolated anti-SARS-CoV-2 nanobody as described herein comprises:

- (a) a CDR1 comprising the amino acid sequence of GFTFKNAD (SEQ ID NO:2);
- (b) a CDR2 comprising the amino acid sequence of IYSDGST (SEQ ID NO:3); and
- (c) a CDR3 comprising the amino acid sequence of MAGSKSGQELDH (SEQ ID

15 NO:4).

In certain embodiments, the isolated anti-SARS-CoV-2 nanobody as described herein comprises:

- (a) a CDR1 comprising the amino acid sequence of GFTFKNAD (SEQ ID NO:2);
- (b) a CDR2 comprising the amino acid sequence of IYSDGST (SEQ ID NO:3); and
- (c) a CDR3 comprising the amino acid sequence of MAGSKSGHELDH (SEQ ID

20 NO:8).

In certain embodiments, the isolated anti-SARS-CoV-2 nanobody as described herein comprises:

- (a) a CDR1 comprising the amino acid sequence of GFTFKNAD (SEQ ID NO:2);
- (b) a CDR2 comprising the amino acid sequence of IYSDGRT (SEQ ID NO:12); and
- (c) a CDR3 comprising the amino acid sequence of MAGSKSGHELDH (SEQ ID

25 NO:8).

In some embodiments, an anti-SARS-CoV-2 nanobody comprises a sequence (e.g., a variable domain of a heavy-chain only antibody (VHH), or a fragment thereof) derived from any
 30 of the following nanobodies described herein: Nanosota-1A, Nanosota-1B and Nanosota-1C. The amino acid sequences of these anti-SARS-CoV-2 nanobody clones are set forth in Table 1 below. In certain embodiments, an anti-SARS-CoV-2 nanobody comprises a sequence as described in any of the embodiments provided herein.

In certain embodiments, an anti-SARS-CoV-2 nanobody described herein comprises an amino acid sequence that has at least 80% (e.g., 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100%) sequence identity to any one of:

5 (a) MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYRQAPGQGL
EWVTSIYSDGSTVYADSVKGRFTVSRDNPSTVSLQMNSLKPEDTGVYHGMAGSKSGQ
ELDHWGQGTQVTVSS (SEQ ID NO:1);

(b) MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYRQAPGQGL
EWVTSIYSDGSTVYADSVKGRFTVSRDNPSTVSLQMNSLKPEDTGVYYCMAGSKSGH
10 ELDHWGQGTQVTVSS (SEQ ID NO:7); and

(c) MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYRQVPGQGL
EWVTSIYSDGRTVYADSVKGRFTVSRDNPSTVSLQMNSLKPEDTGVYYCMAGSKSG
HELDHWGQGTQVTVSS (SEQ ID NO:11).

In some embodiments, an anti-SARS-CoV-2 nanobody comprises an amino acid
15 sequence of any one of SEQ ID NOs:1, 7, and 11. In some embodiments, an anti-SARS-CoV-2
nanobody consists of the amino acid sequence of any one of SEQ ID NOs:1, 7, and 11.

Clone Nanosota-1A

In some embodiments, an anti-SARS-CoV-2 nanobody comprises a CDR1 comprising
the amino acid sequence of SEQ ID NO:2, a CDR2 comprising the amino acid sequence of SEQ
20 ID NO:3, and a CDR3 comprising the amino acid sequence of SEQ ID NO:4. In some
embodiments, an anti-SARS-CoV-2 nanobody comprises CDRs1-3 consisting of the amino acid
sequences of SEQ ID NOs:2, 3, and 4, respectively.

In some embodiments, an anti-SARS-CoV-2 nanobody comprises an amino acid
sequence that has at least about 80% (e.g., 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%,
25 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100%) sequence identity to SEQ ID
NO:1. In some embodiments, an anti-SARS-CoV-2 nanobody comprises an amino acid
sequence of SEQ ID NO:1. In some embodiments, an anti-SARS-CoV-2 nanobody consists of
the amino acid sequence of SEQ ID NO:1.

Clone Nanosota-1B

30 In some embodiments, an anti-SARS-CoV-2 nanobody comprises a CDR1 comprising
the amino acid sequence of SEQ ID NO:2, a CDR2 comprising the amino acid sequence of SEQ
ID NO:3, and a CDR3 comprising the amino acid sequence of SEQ ID NO:8. In some
embodiments, an anti-SARS-CoV-2 nanobody comprises CDRs1-3 consisting of the amino acid
sequences of SEQ ID NOs:2, 3, and 8, respectively.

In some embodiments, an anti-SARS-CoV-2 nanobody comprises an amino acid sequence that has at least about 80% (e.g., 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100%) sequence identity to SEQ ID NO:7. In some embodiments, an anti-SARS-CoV-2 nanobody comprises an amino acid sequence of SEQ ID NO:7. In some embodiments, an anti-SARS-CoV-2 nanobody consists of the amino acid sequence of SEQ ID NO:7.

Clone Nanosota-1C

In some embodiments, an anti-SARS-CoV-2 nanobody comprises a CDR1 comprising the amino acid sequence of SEQ ID NO:2, a CDR2 comprising the amino acid sequence of SEQ ID NO:12, and a CDR3 comprising the amino acid sequence of SEQ ID NO:8. In some embodiments, an anti-SARS-CoV-2 nanobody comprises CDRs1-3 consisting of the amino acid sequences of SEQ ID NOs:2, 12, and 8, respectively.

In some embodiments, an anti-SARS-CoV-2 nanobody comprises an amino acid sequence that has at least about 80% (e.g., 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100%) sequence identity to SEQ ID NO:11. In some embodiments, an anti-SARS-CoV-2 nanobody comprises an amino acid sequence of SEQ ID NO:11. In some embodiments, an anti-SARS-CoV-2 nanobody consists of the amino acid sequence of SEQ ID NO:11.

Detectable Agents

In certain embodiments, a binding molecule described herein is operably linked to at least one detectable agent. In certain embodiments, an isolated anti-SARS-CoV-2 nanobody as described herein is operably linked to at least one detectable agent. The location of the detectable agent is not critical, provided that it does not interfere with the function of the nanobody. In certain embodiments, the detectable agent is operably linked to the N-terminus of the nanobody. In certain embodiments, the detectable agent is operably linked to the C-terminus of the nanobody.

In certain embodiments, the at least one detectable agent is a tag, such as an affinity tag or an epitope tag. For example, such a tag may be useful for detecting, isolating and/or purifying the nanobody polypeptide. In certain embodiments, the tag is a polypeptide tag. Polypeptide tags are known in the art and include, but are not limited to, e.g., a His tag, Myc tag, HA tag or an Fc tag. In certain embodiments, the at least one detectable agent is an Fc tag (e.g., an IgG1, IgG2, IgG3, or IgG4 Fc). For example, as described herein, a nanobody may be operably linked to an Fc domain amino acid sequence, to produce a nanobody-Fc fusion

polypeptide. In certain embodiments, the at least one detectable agent is a His tag and/or an HA tag. As used herein, an HA tag is a short peptide tag comprising a sequence of YPYDVPDYA (SEQ ID NO:22). In certain embodiments, a sequence comprising an HA tag is about 9-15 amino acids (e.g., 9-12 aa) in length. In certain embodiments, multiple tags may be operably
 5 linked in tandem either directly or via a linker group (e.g., a His tag and an HA tag; *see, e.g.*, SEQ ID NO:18).

In certain embodiments, the nanobody is directly linked to the detectable agent, such as a polypeptide tag (e.g., through a peptide bond).

In certain other embodiments, the nanobody is linked to the detectable agent, such as a
 10 polypeptide tag, via one or more linker group(s). The nature of the linker group is not critical, provided that the linker group does not interfere with the function of the nanobody or the detectable agent. In certain embodiments, the linker group is an amino acid sequence (e.g., a sequence described herein). In certain embodiments, the linker group is an amino acid sequence that is at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more amino acids in length. In certain embodiments,
 15 the linker group is an amino acid sequence about 1 to about 25 amino acids in length, or about 1 to about 20 amino acids in length, or about 1 to about 15 amino acids in length, or about 1 to about 10 amino acids in length (e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 amino acids in length). In certain embodiments, the linker group sequence comprises an amino acid sequence having at least about 60% sequence identity to SEQ ID NO:15. In certain embodiments, the linker group
 20 sequence comprises an amino acid sequence having at least about 80% sequence identity to SEQ ID NO:15. In certain embodiments, the linker group sequence comprises an amino acid sequence having at least about 90% sequence identity to SEQ ID NO:15. In certain embodiments, the linker group sequence comprises SEQ ID NO:15. In certain embodiments, the linker group sequence consists of SEQ ID NO:15.

25 Accordingly, in certain embodiments, an anti-SARS-CoV-2 nanobody described herein operably linked to a detectable agent, comprises an amino acid sequence that has at least 80% (e.g., 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100%) sequence identity to any one of:

(a) MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYRQAPGQGL
 30 EWVTSIYSDGSTVYADSVKGRFTVSRDNPSTVSLQMNSLKPEDTGVYHGMAGSKSGQ
 ELDHWGQGTQVTVSSGPGGQHHHHHHGAYPYDVPDYAS (SEQ ID NO:5);

(b) MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYRQAPGQGL
 EWVTSIYSDGSTVYADSVKGRFTVSRDNPSTVSLQMNSLKPEDTGVYYCMAGSKSGH
 ELDHWGQGTQVTVSSGPGGQHHHHHHGAYPYDVPDYAS (SEQ ID NO:9); and

(c) MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYRQVPGQGL
EWVTSIYSDGRTVYADSVKGRFTVSRDNPVKSTVSLQMNSLKPEDTGVYYCMAGSKSG
HELDHWGQGTQVTVSSGPGGQHSHHHHHGAYPYDVPDYAS (SEQ ID NO:13).

In some embodiments, an anti-SARS-CoV-2 nanobody comprises an amino acid
5 sequence of any one of SEQ ID NOs:5, 9, and 13. In some embodiments, an anti-SARS-CoV-2
nanobody consists of the amino acid sequence of any one of SEQ ID NOs:5, 9, and 13.

In some embodiments, the nanobody polypeptide is encoded by a polynucleotide
comprising a nucleic acid sequence that has at least about 80% (e.g., at least about 80%, 85%,
90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100%) sequence identity to SEQ ID
10 NO:23, 25, or 27.

In certain embodiments, a nanobody as described herein is a recombinant nanobody. In
certain embodiments, a nanobody as described herein is a chimeric nanobody. In certain
embodiments, a nanobody or as described herein is humanized.

In certain embodiments, a nanobody of the invention is a monoclonal nanobody. In some
15 embodiments, the monoclonal nanobody recognizes an epitope within SARS-CoV-2.

In certain embodiments, an isolated anti-SARS-CoV-2 nanobody described herein is an
inhibitor of SARS-CoV-2.

The term “inhibitor of SARS-CoV-2” as used herein refers to nanobody that is capable
of inhibiting the function of SARS-CoV-2 (e.g., inhibits binding to ACE2). For example, in
20 certain embodiments, a nanobody as described herein detectably inhibits the biological activity
of SARS-CoV-2 as measured, e.g., using an assay described herein. In certain embodiments, the
nanobody inhibits the biological activity of SARS-CoV-2 by at least about 5%, at least about
10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least
about 60%, at least about 70%, at least about 80%, or at least about 90%.

In certain embodiments, the nanobody is a selective inhibitor of SARS-CoV-2. For
example, a nanobody of the invention may be at least 5, at least 10, at least 50, at least 100, at
least 500, or at least 1,000 fold selective for SARS-CoV-2 over another coronavirus in a selected
assay (e.g., an assay described in the Examples herein).

Certain embodiments of the invention provide a nanobody as described herein.

Polypeptides and Protein Molecules

In certain embodiments, an isolated anti-SARS-CoV-2 nanobody described herein is
further linked to one or more antibody domain sequences (e.g., heavy or light chain domain
sequences, such as variable or constant domain sequences). Accordingly, certain embodiments

provide a polypeptide comprising a nanobody of the invention operably linked to one or more antibody domain sequences. In certain embodiments, the one or more antibody domain sequences are derived from an antibody class or isotype as defined herein (e.g., IgG (IgG1, IgG2, IgG3, IgG4), IgM, IgA (IgA1 and IgA2), IgD, and IgE).

5 In certain embodiments, the nanobody is not linked to a light chain domain. In certain embodiments, the nanobody is not linked to a constant domain region. In certain embodiments, the nanobody is not linked to a CH1 region.

In certain embodiments, an isolated anti-SARS-CoV-2 nanobody described herein is linked (e.g., through a linker or a direct bond, such as a peptide bond) to at least one heavy chain
10 constant region (e.g., 1, 2, or 3). In certain embodiments, the nanobody is linked to two heavy chain constant regions (e.g., a CH2 and CH3 region). In certain embodiments, the nanobody is operably linked to an Fc domain amino acid sequence (e.g., an IgG Fc domain such as IgG1, IgG2, IgG3, or IgG4 Fc domain), to produce a nanobody-Fc fusion polypeptide.

Thus, certain embodiments of the invention provide a nanobody-Fc-fusion polypeptide
15 comprising a nanobody of the invention operably linked to a Fc domain amino acid sequence. In certain embodiments, the nanobody and Fc domain amino acid sequence are directly linked, e.g., through a peptide bond. In certain embodiments, the nanobody and Fc domain amino acid sequence are linked through an amino acid linking group. In certain embodiments, the Fc domain amino acid sequence is an IgG4 Fc domain amino acid sequence. In certain
20 embodiments, the Fc domain amino acid sequence is an IgG1 Fc domain amino acid sequence. In certain embodiments, the Fc domain amino acid sequence has at least about 80% (e.g., 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100%) sequence identity to SEQ ID NO:19. In certain embodiments, the Fc domain amino acid sequence comprises SEQ ID NO:19. In certain embodiments, the Fc
25 domain amino acid sequence consists of SEQ ID NO:19.

In certain embodiments, the nanobody-Fc fusion polypeptide comprises an amino acid sequence that has at least 80% (e.g., 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100%) sequence identity to any one of:

(a) MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYRQAPGQGL
30 EWVTSIYSDGSTVYADSVKGRFTVSRDNPSTVSLQMNSLKPEDTGVYHGMAGSKSGQ
ELDHWGQGTQVTVSSEPKSCDKTHTCPPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVT
CVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNG
KEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPS

DIAVEWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALH
NHYTQKSLSLSPGK (SEQ ID NO:6);

(b) MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYRQAPGQGL
EWVTSIYSDGSTVYADSVKGRFTVSRDNPSTVSLQMNSLKPEDTGVIYCMAGSKSGH
5 ELDHWGQGTQVTVSSEPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVT
CVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNG
KEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPS
DIAVEWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALH
NHYTQKSLSLSPGK (SEQ ID NO:10); and

(c) MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYRQVPGQGL
EWVTSIYSDGRTVYADSVKGRFTVSRDNPSTVSLQMNSLKPEDTGVIYCMAGSKSG
HELDHWGQGTQVTVSSEPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEV
TCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNG
GKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYP
15 SDIAVEWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEAL
HNHYTQKSLSLSPGK (SEQ ID NO:14).

In some embodiments, the nanobody-Fc fusion polypeptide comprises an amino acid sequence of any one of SEQ ID NOs:6, 10, and 14. In some embodiments, the nanobody-Fc fusion polypeptide consists of an amino acid sequence of any one of SEQ ID NOs:6, 10, and 14.

20 In some embodiments, the nanobody-Fc fusion polypeptide is encoded by a polynucleotide comprising a nucleic acid sequence that has at least about 80% (*e.g.*, at least about 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100%) sequence identity to SEQ ID NO:24, 26, or 28.

Certain embodiments of the invention also provide multivalent nanobodies (*e.g.*,
25 bivalent, trivalent, tetravalent, pentavalent or higher valence multivalent nanobodies). Thus, certain embodiments of the invention provide a protein molecule comprising two or more independently selected nanobodies as described herein, wherein the nanobodies are operably linked to each other (*e.g.*, to form a dimer, trimer, tetramer, pentamer or higher valence multimer nanobody). In certain embodiments, a multivalent nanobody or protein molecule as described
30 herein is a homo-multimer (*e.g.*, dimer, trimer, tetramer or pentamer). In certain embodiments, a multivalent nanobody or protein molecule as described herein is a hetero-multimer (*e.g.*, dimer, trimer, tetramer or pentamer).

In certain embodiments, the two or more nanobodies are operably linked via a linker group (*e.g.*, a peptide linker group), disulfide bond(s) and/or by non-covalent interactions. In

certain embodiments, the two or more nanobodies are operably linked via oligomerization of tag polypeptides (e.g., multimerization tags, such as a dimerization tags, trimerization tags, tetramerization tags, etc.).

In certain embodiments, the two or more nanobodies are operably linked via a linker group. The nature of the linker group is not critical, provided that the linker group does not interfere with the function of the nanobodies. In certain embodiments, the linker group is a peptide linker group. In certain embodiments, the peptide linker is a glycine-serine rich linker.

In certain embodiments, two independently selected nanobodies are linked via a linker group (e.g., a peptide linker group) to form dimeric nanobody. In certain embodiments, three independently selected nanobodies are linked via two linker groups (e.g., two peptide linker groups) to form a trimeric nanobody. In certain embodiments, four independently selected nanobodies are linked via three linker groups (e.g., three peptide linker groups) to form a tetrameric nanobody. In certain embodiments, five independently selected nanobodies are linked via four linker groups (e.g., four peptide linker groups) to form a pentameric nanobody.

In certain embodiments, the two or more nanobodies are operably linked via oligomerization of tag polypeptides. For example, a nanobody as described herein may be operably linked to a tag polypeptide to form a nanobody-tag fusion polypeptide, wherein the tag polypeptide is capable of oligomerizing. Accordingly, two or more nanobody-tag fusion polypeptides may be operably linked to form a dimer, trimer, tetramer, pentamer or a higher valence multimer via polypeptide tag-mediated oligomerization.

In certain embodiments, the nanobody and tag polypeptide are linked through a peptide linker to form the nanobody-tag fusion polypeptide. In certain embodiments, a nanobody and tag polypeptide are directly linked without an intervening peptide linker to form the nanobody-tag fusion polypeptide.

In certain embodiments, the tag polypeptide is a human Fc sequence, a human collagen XVIII trimerization domain or a coiled-coil peptide derived from human cartilage oligomeric matrix protein COMP48, which is capable of forming a multimer, such as a pentamer.

In certain embodiments, the tag polypeptide is a human Fc sequence. Accordingly, certain embodiments of the invention provide a protein molecule comprising: two independently selected nanobody-Fc fusion polypeptides as described herein, wherein the two Fc polypeptides are linked to form a dimer (e.g., linked by a covalent bond, such as a disulfide bond, or by non-covalent interactions such as electrostatic interactions, hydrogen bonding, etc.).

In certain embodiments, the two nanobody-Fc fusion polypeptides are the same. In certain embodiments, the two nanobody-Fc fusion polypeptides are different. In certain

embodiments, nanobody-Fc polypeptides as described herein can form homo-dimers. In certain embodiments, nanobody-Fc polypeptides as described herein can form hetero-dimers. In certain embodiments, nanobody-Fc polypeptides as described herein can form bispecific hetero-dimers having binding affinities for different SARS-Cov-2 protein(s) and/or epitopes.

5 In certain other embodiments, a single nanobody of the invention is operably linked to an Fc dimer.

In certain embodiments, a polypeptide or protein molecule as described herein is further operably linked to a detectable agent (e.g., a detectable agent described herein).

Certain embodiments of the invention also provide a polypeptide or protein molecule
10 comprising a nanobody as described herein.

Certain Nanobody, Polypeptide and Protein Molecule Embodiments

As described herein, a nanobody of the invention may be incorporated into a polypeptide or a protein molecule. For example, a nanobody of the invention may be linked to one or more antibody domains as described herein. Therefore, such polypeptides/protein molecules
15 comprising a nanobody of the invention and one or more additional antibody domains may be referenced herein as an antibody or antibody fragment. Additionally, such molecules may be further modified as described herein (e.g., humanized or to alter its affinity, etc.).

The terms “protein”, “protein molecule”, “peptide” and “polypeptide” are used interchangeably herein. In certain embodiments, the term “protein” or “protein molecule” may
20 refer to a single polypeptide or may refer to two or more polypeptides, wherein the two or more polypeptides may be linked by a covalent (e.g., disulfide bridge) or non-covalent interactions.

As used herein, the term “antibody” includes a single-chain variable fragment (scFv), a dimer of a nanobody, a nanobody-Fc fusion polypeptide or dimer thereof, humanized, fully human or chimeric antibodies, single-chain antibodies, diabodies, and antigen-binding fragments
25 of antibodies that do not contain the Fc region (e.g., Fab fragments). In certain embodiments, the antibody is a camelid antibody, human antibody or a humanized antibody. A “humanized” antibody contains only the three CDRs (complementarity determining regions) and sometimes a few carefully selected “framework” residues (the non-CDR portions of the variable regions) from each donor antibody variable region recombinantly linked onto the corresponding
30 frameworks and constant regions of a human antibody sequence. A “fully humanized antibody” is created in a hybridoma from mice genetically engineered to have only human-derived antibody genes or by selection from a phage-display library of human-derived antibody genes.

As used herein, the term “monoclonal nanobody” or “monoclonal antibody” refers to a nanobody/antibody obtained from a group of substantially homogeneous nanobodies/antibodies,

that is, a nanobody/antibody group wherein the nanobodies/antibodies constituting the group are homogeneous except for naturally occurring mutants that exist in a small amount. Monoclonal nanobodies/antibodies are highly specific and interact with a single antigenic site. Furthermore, each monoclonal nanobody/antibody targets a single antigenic determinant (epitope) on an antigen, as compared to common polyclonal nanobody/antibody preparations that typically contain various nanobodies/antibodies against diverse antigenic determinants. In addition to their specificity, monoclonal nanobodies/antibodies are advantageous in that they are typically produced from hybridoma cultures not contaminated with other immunoglobulins.

The adjective "monoclonal" indicates a characteristic of antibodies and nanobodies obtained from a substantially homogeneous group of antibodies/nanobodies, and does not specify antibodies/nanobodies produced by a particular method. For example, a monoclonal nanobody to be used in the present invention can be produced by, for example, hybridoma methods (Kohler and Milstein, *Nature* 256:495, 1975) or recombination methods (U.S. Pat. No. 4,816,567). The monoclonal nanobodies used in the present invention can be also isolated from a phage nanobody library (Clackson et al., *Nature* 352:624-628, 1991; Marks et al., *J. Mol. Biol.* 222:581-597, 1991). The monoclonal nanobodies of the present invention may be linked to other antibody domain sequences. Therefore, the resulting polypeptides/protein molecules may be "chimeric" immunoglobulins, wherein a part of the polypeptide is derived from a specific species or a specific antibody class or subclass, and the remaining portion is derived from another species, or another antibody class or subclass. Furthermore, mutant nanobodies, as well as mutant polypeptides/protein molecules comprising a nanobody of the invention, are also comprised in the present invention (U.S. Pat. No. 4,816,567; Morrison et al., *Proc. Natl. Acad. Sci. USA* 81:6851-6855, 1984).

As used herein, the term "mutant nanobody" or "mutant antibody" refers to a nanobody/antibody comprising a variant amino acid sequence in which one or more amino acid residues have been altered. For example, the variable region of a nanobody/antibody can be modified to improve its biological properties, such as antigen binding. Such modifications can be achieved by site-directed mutagenesis (see Kunkel, *Proc. Natl. Acad. Sci. USA* 82: 488 (1985)), PCR-based mutagenesis, cassette mutagenesis, and the like. Such mutants comprise an amino acid sequence which is at least 70% identical to the amino acid sequence of the heavy chain variable region of the nanobody, more specifically at least 75%, even more specifically at least 80%, still more specifically at least 85%, yet more specifically at least 90%, and most specifically at least 95% identical. Such mutants also comprise an amino acid sequence which is at least 70% identical to the amino acid sequence of a heavy or light chain variable region of the

antibody, more specifically at least 75%, even more specifically at least 80%, still more specifically at least 85%, yet more specifically at least 90%, and most specifically at least 95% identical. As used herein, the term “sequence identity” is defined as the percentage of residues identical to those in the nanobody’s/antibody’s original amino acid sequence, determined after the sequences are aligned and gaps are appropriately introduced to maximize the sequence identity as necessary.

Specifically, the identity of one nucleotide sequence or amino acid sequence to another can be determined using the algorithm BLAST, by Karlin and Altschul (Proc. Natl. Acad. Sci. USA, 90: 5873-5877, 1993). Programs such as BLASTN and BLASTX were developed based on this algorithm (Altschul et al., J. Mol. Biol. 215: 403-410, 1990). To analyze nucleotide sequences according to BLASTN based on BLAST, the parameters are set, for example, as score=100 and wordlength=12. On the other hand, parameters used for the analysis of amino acid sequences by BLASTX based on BLAST include, for example, score=50 and wordlength=3. Default parameters for each program are used when using the BLAST and Gapped BLAST programs. Specific techniques for such analyses are known in the art (see the website of the National Center for Biotechnology Information (NCBI), Basic Local Alignment Search Tool (BLAST); <http://www.ncbi.nlm.nih.gov>).

Polyclonal and monoclonal nanobodies/antibodies can be prepared by methods known to those skilled in the art.

In another embodiment, antibodies or antibody fragments (e.g., nanobodies) can be isolated from an antibody/nanobody phage library, produced by using the technique reported by McCafferty et al. (Nature 348:552-554 (1990)). Clackson et al. (Nature 352:624-628 (1991)), Marks et al. (J. Mol. Biol. 222:581-597 (1991)) and Muyldermans et al. (Annual Review of Biochemistry Volume 82, pp 775-797(2013)) reported on the respective isolation of mouse, camelid and human antibodies from phage libraries. There are also reports that describe the production of high affinity (nM range) human antibodies based on chain shuffling (Marks et al., Bio/Technology 10:779-783 (1992)), and combinatorial infection and in vivo recombination, which are methods for constructing large-scale phage libraries (Waterhouse et al., Nucleic Acids Res. 21:2265-2266 (1993)). These technologies can also be used to isolate monoclonal nanobodies/antibodies, instead of using conventional hybridoma technology for monoclonal nanobody/antibody production.

Nanobodies/antibodies to be used in the present invention can be purified by a method appropriately selected from known methods, such as the protein A-Sepharose method, hydroxyapatite chromatography, salting-out method with sulfate, ion exchange chromatography,

and affinity chromatography, or by the combined use of the same.

The present invention may use recombinant nanobodies/antibodies, produced by gene engineering. The genes encoding the nanobodies/antibodies obtained by a method described above are isolated from B cells or hybridomas. The genes are inserted into an appropriate vector, and then introduced into a host (see, e.g., Carl, A. K. Borrebaeck, James, W. Larrick, Therapeutic Monoclonal Antibodies, Published in the United Kingdom by Macmillan Publishers Ltd, 1990). The present invention provides the nucleic acids encoding the nanobodies/antibodies of the present invention, and vectors comprising these nucleic acids. Specifically, using a reverse transcriptase, cDNAs encoding the variable region(s) (V region) of the nanobodies/antibodies are synthesized from the mRNAs of B cells or hybridomas. After obtaining the DNAs encoding the variable region(s) of interest, they are optionally ligated with DNAs encoding desired constant regions (C regions), and the resulting DNA constructs are inserted into expression vectors. Alternatively, the DNAs encoding the variable region(s) may be inserted into expression vectors comprising the DNAs of the C regions. These are inserted into expression vectors so that the genes are expressed under the regulation of an expression regulatory region, for example, an enhancer and promoter. Then, host cells are transformed with the expression vectors to express the nanobodies/antibodies. The present invention provides cells expressing nanobodies/antibodies of the present invention. The cells expressing nanobodies/antibodies of the present invention include cells and hybridomas transformed with a gene of such a nanobody/antibody.

The nanobodies/antibodies of the present invention also include nanobodies/antibodies which comprise complementarity-determining regions (CDRs), or regions functionally equivalent to CDRs. The term "functionally equivalent" refers to comprising amino acid sequences similar to the amino acid sequences of CDRs of any of the monoclonal nanobodies isolated in the Examples. The term "CDR" refers to a region in a nanobody/antibody variable region (also called "V region"), and determines the specificity of antigen binding. The H chain and L chain (if present) each have three CDRs, designated from the N terminus as CDR1, CDR2, and CDR3. There are four regions flanking these CDRs: these regions are referred to as "framework," and their amino acid sequences are highly conserved. The CDRs can be transplanted into other nanobodies/antibodies, and thus a recombinant antibody can be prepared by combining CDRs with the framework of a desired nanobody/antibody. One or more amino acids of a CDR can be modified without losing the ability to bind to its antigen. For example, one or more amino acids in a CDR can be substituted, deleted, and/or added.

In certain embodiments, an amino acid residue is mutated into one that allows the

properties of the amino acid side-chain to be conserved. Examples of the properties of amino acid side chains comprise: hydrophobic amino acids (A, I, L, M, F, P, W, Y, V), hydrophilic amino acids (R, D, N, C, E, Q, G, H, K, S, T), and amino acids comprising the following side chains: aliphatic side-chains (G, A, V, L, I, P); hydroxyl group-containing side-chains (S, T, Y);
 5 sulfur atom-containing side-chains (C, M); carboxylic acid- and amide-containing side-chains (D, N, E, Q); base-containing side-chains (R, K, H); and aromatic-containing side-chains (H, F, Y, W). The letters within parenthesis indicate the one-letter amino acid codes. Amino acid substitutions within each group are called conservative substitutions. It is well known that a polypeptide comprising a modified amino acid sequence in which one or more amino acid
 10 residues is deleted, added, and/or substituted can retain the original biological activity (Mark D. F. et al., *Proc. Natl. Acad. Sci. U.S.A.* 81:5662-5666 (1984); Zoller M. J. and Smith M., *Nucleic Acids Res.* 10: 6487-6500 (1982); Wang A. et al., *Science* 224: 1431-1433; Dalbadie-McFarland G. et al., *Proc. Natl. Acad. Sci. U.S.A.* 79: 6409-6413 (1982)). The number of mutated amino acids is not limited, but in general, the number falls within 40% of amino acids of each CDR,
 15 and specifically within 35%, and still more specifically within 30% (e.g., within 25%). The identity of amino acid sequences can be determined as described herein.

In the present invention, recombinant nanobodies/antibodies artificially modified to reduce heterologous antigenicity against humans can be used. Examples include chimeric nanobodies/antibodies and humanized nanobodies/antibodies. These modified
 20 nanobodies/antibodies can be produced using known methods. A chimeric antibody includes an antibody comprising variable and constant regions of species that are different to each other, for example, an antibody comprising the antibody heavy chain and light chain variable regions of a nonhuman mammal such as a mouse, and the antibody heavy chain and light chain constant regions of a human. Additionally, a chimeric antibody or polypeptide may be produced by
 25 combining a nanobody of the invention with constant regions that are of different species to each other. Such an antibody (e.g., a camelid-human chimeric antibody) can be obtained by (1) ligating a DNA from the different regions; (2) incorporating this into an expression vector; and (3) introducing the vector into a host for production of the antibody.

A humanized nanobody/antibody, which is also called a reshaped human /nanobody
 30 antibody, may be obtained by replacing a CDR of a human antibody with an H or L chain CDR of a nanobody/antibody of a nonhuman mammal such as a mouse or camelid. Conventional genetic recombination techniques for the preparation of such antibodies are known (see, for example, Jones et al., *Nature* 321: 522-525 (1986); Reichmann et al., *Nature* 332: 323-329 (1988); Presta *Curr. Op. Struct. Biol.* 2: 593-596 (1992)). Specifically, a DNA sequence

designed to ligate a CDR of a mouse/camelid antibody with the framework regions (FRs) of a human antibody is synthesized by PCR, using several oligonucleotides constructed to comprise overlapping portions at their ends. A humanized antibody can be obtained by (1) ligating the resulting DNA to a DNA that encodes a human antibody constant region; (2) incorporating this
5 into an expression vector; and (3) transfecting the vector into a host to produce the antibody (see, European Patent Application No. EP 239,400, and International Patent Application No. WO 96/02576). Human antibody FRs that are ligated via the CDR are selected where the CDR forms a favorable antigen-binding site. The humanized antibody may comprise additional amino acid residue(s) that are not included in the CDRs introduced into the recipient antibody,
10 nor in the framework sequences. Such amino acid residues are usually introduced to more accurately optimize the antibody's ability to recognize and bind to an antigen. For example, as necessary, amino acids in the framework region of a variable region may be substituted such that the CDR of a reshaped human antibody forms an appropriate antigen-binding site (Sato, K. et al., *Cancer Res.* (1993) 53, 851-856).

15 Isotypes of nanobody-fusion polypeptides or antibodies comprising a nanobody of the present invention, or antibody fragments thereof, are not limited. The isotypes include, for example, IgG (IgG1, IgG2, IgG3, and IgG4), IgM, IgA (IgA1 and IgA2), IgD, and IgE.

As described herein, a nanobody of the present invention may be operably linked to one or more antibody domain sequences (e.g., a nanobody-Fc fusion polypeptide). Therefore, such
20 polypeptides/protein molecules comprising a nanobody of the invention and one or more additional antibody domains may be referenced herein as an antibody or antibody fragment. The term "antibody fragment" refers to a portion of a full-length antibody, and generally to a fragment comprising an antigen-binding domain or a variable region. Such antibody fragments include, for example, single domain antibody (sdAb), Fab, F(ab')₂, Fv, single-chain Fv (scFv)
25 which comprises a heavy chain Fv and a light chain Fv coupled together with an appropriate linker, diabody (diabodies), linear antibodies, and multispecific antibodies prepared from antibody fragments. Previously, antibody fragments were produced by digesting natural antibodies with a protease; currently, methods for expressing them as recombinant antibodies using genetic engineering techniques are also known (see Morimoto et al., *Journal of*
30 *Biochemical and Biophysical Methods* 24:107-117 (1992); Brennan et al., *Science* 229:81 (1985); Co, M. S. et al., *J. Immunol.*, 1994, 152, 2968-2976; Better, M. & Horwitz, A. H., *Methods in Enzymology*, 1989, 178, 476-496, Academic Press, Inc.; Plueckthun, A. & Skerra, A., *Methods in Enzymology*, 1989, 178, 476-496, Academic Press, Inc.; Lamoyi, E., *Methods in Enzymology*, 1989, 121, 663-669; Bird, R. E. et al., *TIBTECH*, 1991, 9, 132-137).

The nanobodies/antibodies can be purified to homogeneity. The nanobodies/antibodies can be isolated and purified by a method routinely used to isolate and purify proteins. The nanobodies/antibodies can be isolated and purified by the combined use of one or more methods appropriately selected from column chromatography, filtration, ultrafiltration, salting out, dialysis, preparative polyacrylamide gel electrophoresis, and isoelectro-focusing, for example (Strategies for Protein Purification and Characterization: A Laboratory Course Manual, Daniel R. Marshak et al. eds., Cold Spring Harbor Laboratory Press (1996); Antibodies: A Laboratory Manual. Ed Harlow and David Lane, Cold Spring Harbor Laboratory, 1988). Such methods are not limited to those listed above. Chromatographic methods include affinity chromatography, ion exchange chromatography, hydrophobic chromatography, gel filtration, reverse-phase chromatography, and adsorption chromatography. These chromatographic methods can be practiced using liquid phase chromatography, such as HPLC and FPLC. Columns to be used in affinity chromatography include protein A columns and protein G columns. For example, protein A columns include Hyper D, POROS, and Sepharose F. F. (Pharmacia).

Nanobodies/antibodies can also be purified by utilizing antigen binding, using carriers on which antigens have been immobilized.

The nanobodies/antibodies of the present invention can be formulated according to standard methods (see, for example, Remington's Pharmaceutical Science, latest edition, Mark Publishing Company, Easton, U.S.A), and may comprise pharmaceutically acceptable carriers and/or additives. The present invention relates to compositions (including reagents and pharmaceuticals) comprising the nanobodies/antibodies of the invention, and pharmaceutically acceptable carriers and/or additives. Exemplary carriers include surfactants (for example, PEG and Tween), excipients, antioxidants (for example, ascorbic acid), coloring agents, flavoring agents, preservatives, stabilizers, buffering agents (for example, phosphoric acid, citric acid, and other organic acids), chelating agents (for example, EDTA), suspending agents, isotonicizing agents, binders, disintegrators, lubricants, fluidity promoters, and corrigents. However, the carriers that may be employed in the present invention are not limited to this list. In fact, other commonly used carriers can be appropriately employed: light anhydrous silicic acid, lactose, crystalline cellulose, mannitol, starch, carmellose calcium, carmellose sodium, hydroxypropylcellulose, hydroxypropylmethyl cellulose, polyvinylacetaldihethylaminoacetate, polyvinylpyrrolidone, gelatin, medium chain fatty acid triglyceride, polyoxyethylene hydrogenated castor oil 60, sucrose, carboxymethylcellulose, corn starch, inorganic salt, and so on. The composition may also comprise other low-molecular-weight polypeptides, proteins such as serum albumin, gelatin, and immunoglobulin, and amino acids such as glycine,

glutamine, asparagine, arginine, and lysine. When the composition is prepared as an aqueous solution for injection, it can comprise an isotonic solution comprising, for example, physiological saline, dextrose, and other adjuvants, including, for example, D-sorbitol, D-mannose, D-mannitol, and sodium chloride, which can also contain an appropriate solubilizing agent, for example, alcohol (for example, ethanol), polyalcohol (for example, propylene glycol and PEG), and non-ionic detergent (polysorbate 80 and HCO-50).

If necessary, nanobodies/antibodies of the present invention may be encapsulated in microcapsules (microcapsules made of hydroxycellulose, gelatin, polymethylmethacrylate, and the like), and made into components (encapsulated or as a surface functionalization moiety) of colloidal drug delivery systems (liposomes, albumin microspheres, microemulsions, nanoparticles, and nano-capsules) (for example, see "Remington's Pharmaceutical Science 16th edition", Oslo Ed. (1980)). Moreover, methods for making sustained-release drugs are known, and these can be applied for the nanobodies/antibodies of the present invention (Langer et al., J. Biomed. Mater. Res. 15: 167-277 (1981); Langer, Chem. Tech. 12: 98-105 (1982); U.S. Pat. No. 3,773,919; EP Patent Application No. 58,481; Sidman et al., Biopolymers 22: 547-556 (1983); EP: 133,988).

Nucleic Acids, Expression Cassettes, Vectors and Cells

Certain embodiments of the invention provide an isolated nucleic acid encoding a nanobody as described herein or a polypeptide as described herein (e.g., an antibody or antibody fragment comprising a nanobody of the invention).

For example, certain embodiments of the invention provide an isolated nucleic acid comprising one or more CDR sequences (e.g., 1, 2 or 3 CDR sequences), wherein the CDR sequence has at least about 80% (e.g., at least about 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100%) sequence identity to a CDR sequence described herein (e.g., identified in SEQ ID NO:29, 30 or 31 of Table 1). In certain embodiments, the isolated nucleic acid comprises the three CDR sequences shown in SEQ ID NO:29 in Table 1. In certain embodiments, the isolated nucleic acid comprises the three CDR sequences shown in SEQ ID NO:30 in Table 1. In certain embodiments, the isolated nucleic acid comprises the three CDR sequences shown in SEQ ID NO:31 in Table 1.

Certain embodiments of the invention provide an isolated nucleic acid comprising a sequence that has at least about 80% (e.g., at least about 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100%) sequence identity to SEQ ID NO:29, 30, or 31. In certain embodiments, the isolated nucleic acid comprises or consists of SEQ ID NO:29. In certain

embodiments, the isolated nucleic acid comprises or consists of SEQ ID NO:30. In certain embodiments, the isolated nucleic acid comprises or consists of SEQ ID NO:31.

Certain embodiments of the invention also provide an isolated nucleic acid comprising a sequence that has at least about 80% (*e.g.*, at least about 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100%) sequence identity to SEQ ID NO:23, 25, or 27. In certain embodiments, the isolated nucleic acid comprises or consists of SEQ ID NO:23. In certain embodiments, the isolated nucleic acid comprises or consists of SEQ ID NO:25. In certain embodiments, the isolated nucleic acid comprises or consists of SEQ ID NO:27.

Certain embodiments of the invention also provide an isolated nucleic acid comprising a sequence that has at least about 80% (*e.g.*, at least about 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100%) sequence identity to SEQ ID NO:24, 26, or 28. In certain embodiments, the isolated nucleic acid comprises or consists of SEQ ID NO:24. In certain embodiments, the isolated nucleic acid comprises or consists of SEQ ID NO:26. In certain embodiments, the isolated nucleic acid comprises or consists of SEQ ID NO:28.

In certain embodiments, the nucleic acid further comprises a promoter.

Certain embodiments of the invention provide an expression cassette comprising a nucleic acid as described herein and a promoter.

Certain embodiments of the invention provide a vector (*e.g.*, a phagemid, Adeno-associated viruses (AAV)) comprising a nucleic acid or an expression cassette as described herein.

Certain embodiments of the invention provide a cell comprising a nucleic acid, expression cassette or vector as described herein. In certain embodiments, the cell is a bacterial cell.

In certain embodiments, the cell is a mammalian cell.

In certain embodiments, the cell is a human mammalian cell. In certain embodiments, the cell is a human embryonic kidney (HEK) 293 cell. In certain embodiments, the cell is a 293F cell. In certain embodiments, the cell is a 293T cell. In certain embodiments, the cell is a human embryonic retinal (PER.C6) cell. In certain embodiments, the cell is a HT-1080 cell. In certain embodiments, the cell is a Huh-7 cell.

In certain embodiments, the cell is a non-human mammalian cell. In certain embodiments, the cell is a Monkey kidney epithelial (Vero) cell. In certain embodiments, the cell is a Chinese Hamster Ovary (CHO) cell. In certain embodiments, the cell is a baby hamster kidney (BHK) cell.

In certain embodiments, the cell is a non-mammalian cell. In certain embodiments, the cell is an insect cell. In certain embodiments, the cell is a yeast cell.

Certain embodiments of the invention provide a phage particle comprising a vector as described herein.

5 The term "nucleic acid" refers to deoxyribonucleotides or ribonucleotides and polymers thereof in either single- or double-stranded form, composed of monomers (nucleotides) containing a sugar, phosphate and a base which is either a purine or pyrimidine. Unless specifically limited, the term encompasses nucleic acids containing known analogs of natural nucleotides that have similar binding properties as the reference nucleic acid and are
10 metabolized in a manner similar to naturally occurring nucleotides. Unless otherwise indicated, a particular nucleic acid sequence also implicitly encompasses conservatively modified variants thereof (*e.g.*, degenerate codon substitutions) and complementary sequences as well as the sequence explicitly indicated. Specifically, degenerate codon substitutions may be achieved by generating sequences in which the third position of one or more selected (or all) codons is
15 substituted with mixed-base and/or deoxyinosine residues (Batzner et al., Nucl. Acids Res., 19:508 (1991); Ohtsuka et al., JBC, 260:2605 (1985); Rossolini et al., Mol. Cell. Probes, 8:91 (1994). A "nucleic acid fragment" is a fraction of a given nucleic acid molecule. Deoxyribonucleic acid (DNA) in the majority of organisms is the genetic material while ribonucleic acid (RNA) is involved in the transfer of information contained within DNA into
20 proteins. The term "nucleotide sequence" refers to a polymer of DNA or RNA that can be single- or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases capable of incorporation into DNA or RNA polymers. The terms "nucleic acid," "nucleic acid molecule," "nucleic acid fragment," "nucleic acid sequence or segment," or "polynucleotide" may also be used interchangeably with gene, cDNA, DNA and RNA encoded
25 by a gene.

By "portion" or "fragment," as it relates to a nucleic acid molecule, sequence or segment of the invention, when it is linked to other sequences for expression, is meant a sequence having at least 80 nucleotides, more specifically at least 150 nucleotides, and still more specifically at least 400 nucleotides. If not employed for expressing, a "portion" or "fragment" means at least
30 9, specifically 12, more specifically 15, even more specifically at least 20, consecutive nucleotides, *e.g.*, probes and primers (oligonucleotides), corresponding to the nucleotide sequence of the nucleic acid molecules of the invention.

The invention encompasses isolated or substantially purified nucleic acid or protein compositions. In the context of the present invention, an "isolated" or "purified" DNA molecule

or an "isolated" or "purified" polypeptide is a DNA molecule or polypeptide that exists apart from its native environment and is therefore not a product of nature. An isolated DNA molecule or polypeptide may exist in a purified form or may exist in a non-native environment such as, for example, a transgenic host cell. For example, an "isolated" or "purified" nucleic acid molecule or protein, or biologically active portion thereof, is substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized. In one embodiment, an "isolated" nucleic acid is free of sequences that naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated nucleic acid molecule can contain less than about 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb, or 0.1 kb of nucleotide sequences that naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. A protein that is substantially free of cellular material includes preparations of protein or polypeptide having less than about 30%, 20%, 10%, 5%, (by dry weight) of contaminating protein. When the protein of the invention, or biologically active portion thereof, is recombinantly produced, culture medium may represent less than about 30%, 20%, 10%, or 5% (by dry weight) of chemical precursors or non-protein-of-interest chemicals. Fragments and variants of the disclosed nucleotide sequences and proteins or partial-length proteins encoded thereby are also encompassed by the present invention. By "fragment" or "portion" is meant a full length or less than full length of the nucleotide sequence encoding, or the amino acid sequence of, a polypeptide or protein.

"Naturally occurring" is used to describe an object that can be found in nature as distinct from being artificially produced. For example, a protein or nucleotide sequence present in an organism (including a virus), which can be isolated from a source in nature and which has not been intentionally modified by man in the laboratory, is naturally occurring.

A "variant" of a molecule is a sequence that is substantially similar to the sequence of the native molecule. For nucleotide sequences, variants include those sequences that, because of the degeneracy of the genetic code, encode the identical amino acid sequence of the native protein. Naturally occurring allelic variants such as these can be identified with the use of well-known molecular biology techniques, as, for example, with polymerase chain reaction (PCR) and hybridization techniques. Variant nucleotide sequences also include synthetically derived nucleotide sequences, such as those generated, for example, by using site-directed mutagenesis that encode the native protein, as well as those that encode a polypeptide having amino acid substitutions. Generally, nucleotide sequence variants of the invention will have at least 40, 50,

60, to 70%, *e.g.*, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, to 79%, generally at least 80%, *e.g.*, 81%-84%, at least 85%, *e.g.*, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, to 98%, sequence identity to the native (endogenous) nucleotide sequence.

“Conservatively modified variations” of a particular nucleic acid sequence refers to those
5 nucleic acid sequences that encode identical or essentially identical amino acid sequences, or where the nucleic acid sequence does not encode an amino acid sequence, to essentially identical sequences. Because of the degeneracy of the genetic code, a large number of functionally identical nucleic acids encode any given polypeptide. For instance the codons CGT, CGC, CGA, CGG, AGA, and AGG all encode the amino acid arginine. Thus, at every position where
10 an arginine is specified by a codon, the codon can be altered to any of the corresponding codons described without altering the encoded protein. Such nucleic acid variations are "silent variations" which are one species of "conservatively modified variations." Every nucleic acid sequence described herein which encodes a polypeptide also describes every possible silent variation, except where otherwise noted. One of skill will recognize that each codon in a nucleic
15 acid (except ATG, which is ordinarily the only codon for methionine) can be modified to yield a functionally identical molecule by standard techniques. Accordingly, each "silent variation" of a nucleic acid which encodes a polypeptide is implicit in each described sequence.

“Recombinant DNA molecule” is a combination of DNA sequences that are joined together using recombinant DNA technology and procedures used to join together DNA
20 sequences as described, for example, in Sambrook and Russell, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press (3rd edition, 2001).

The terms "heterologous DNA sequence," "exogenous DNA segment" or "heterologous nucleic acid," each refer to a sequence that originates from a source foreign to the particular host
25 cell or, if from the same source, is modified from its original form. Thus, a heterologous gene in a host cell includes a gene that is endogenous to the particular host cell but has been modified. The terms also include non-naturally occurring multiple copies of a naturally occurring DNA sequence. Thus, the terms refer to a DNA segment that is foreign or heterologous to the cell, or homologous to the cell but in a position within the host cell nucleic acid in which the element is
30 not ordinarily found. Exogenous DNA segments are expressed to yield exogenous polypeptides.

A "homologous" DNA sequence is a DNA sequence that is naturally associated with a host cell into which it is introduced.

"Wild-type" refers to the normal gene, or organism found in nature without any known mutation.

“Genome” refers to the complete genetic material of an organism.

A “vector” is defined to include, *inter alia*, any plasmid, cosmid, viral vector, phage or binary vector in double or single stranded linear or circular form which may or may not be self transmissible or mobilizable, and which can transform prokaryotic or eukaryotic host either by
5 integration into the cellular genome or exist extrachromosomally (*e.g.*, autonomous replicating plasmid with an origin of replication).

“Cloning vectors” typically contain one or a small number of restriction endonuclease recognition sites at which foreign DNA sequences can be inserted in a determinable fashion without loss of essential biological function of the vector, as well as a marker gene that is
10 suitable for use in the identification and selection of cells transformed with the cloning vector. Marker genes typically include genes that provide tetracycline resistance, hygromycin resistance or ampicillin resistance.

“Expression cassette” as used herein means a DNA sequence capable of directing expression of a particular nucleotide sequence in an appropriate host cell, comprising a promoter
15 operably linked to the nucleotide sequence of interest which is operably linked to termination signals. It also typically comprises sequences required for proper translation of the nucleotide sequence. The coding region usually codes for a protein of interest but may also code for a functional RNA of interest, for example antisense RNA or a nontranslated RNA, in the sense or antisense direction. The expression cassette comprising the nucleotide sequence of interest may
20 be chimeric, meaning that at least one of its components is heterologous with respect to at least one of its other components. The expression cassette may also be one that is naturally occurring but has been obtained in a recombinant form useful for heterologous expression. The expression of the nucleotide sequence in the expression cassette may be under the control of a constitutive promoter or of an inducible promoter that initiates transcription only when the host cell is
25 exposed to some particular external stimulus. In the case of a multicellular organism, the promoter can also be specific to a particular tissue or organ or stage of development.

Such expression cassettes will comprise the transcriptional initiation region of the invention linked to a nucleotide sequence of interest. Such an expression cassette is provided with a plurality of restriction sites for insertion of the gene of interest to be under the
30 transcriptional regulation of the regulatory regions. The expression cassette may additionally contain selectable marker genes.

The term “RNA transcript” refers to the product resulting from RNA polymerase catalyzed transcription of a DNA sequence. When the RNA transcript is a perfect complementary copy of the DNA sequence, it is referred to as the primary transcript or it may be

a RNA sequence derived from posttranscriptional processing of the primary transcript and is referred to as the mature RNA. "Messenger RNA" (mRNA) refers to the RNA that is without introns and that can be translated into protein by the cell. "cDNA" refers to a single- or a double-stranded DNA that is complementary to and derived from mRNA.

5 "Regulatory sequences" and "suitable regulatory sequences" each refer to nucleotide sequences located upstream (5' non-coding sequences), within, or downstream (3' non-coding sequences) of a coding sequence, and which influence the transcription, RNA processing or stability, or translation of the associated coding sequence. Regulatory sequences include enhancers, promoters, translation leader sequences, introns, and polyadenylation signal
10 sequences. They include natural and synthetic sequences as well as sequences that may be a combination of synthetic and natural sequences. As is noted above, the term "suitable regulatory sequences" is not limited to promoters. However, some suitable regulatory sequences useful in the present invention will include, but are not limited to constitutive promoters, tissue-specific promoters, development-specific promoters, inducible promoters and viral promoters.

15 "5' non-coding sequence" refers to a nucleotide sequence located 5' (upstream) to the coding sequence. It is present in the fully processed mRNA upstream of the initiation codon and may affect processing of the primary transcript to mRNA, mRNA stability or translation efficiency (Turner et al., *Mol. Biotech.*, 3:225 (1995)).

"3' non-coding sequence" refers to nucleotide sequences located 3' (downstream) to a
20 coding sequence and include polyadenylation signal sequences and other sequences encoding regulatory signals capable of affecting mRNA processing or gene expression. The polyadenylation signal is usually characterized by affecting the addition of polyadenylic acid tracts to the 3' end of the mRNA precursor.

The term "translation leader sequence" refers to that DNA sequence portion of a gene
25 between the promoter and coding sequence that is transcribed into RNA and is present in the fully processed mRNA upstream (5') of the translation start codon. The translation leader sequence may affect processing of the primary transcript to mRNA, mRNA stability or translation efficiency.

The term "mature" protein refers to a post-translationally processed polypeptide without
30 its signal peptide. "Precursor" protein refers to the primary product of translation of an mRNA. "Signal peptide" refers to the amino terminal extension of a polypeptide, which is translated in conjunction with the polypeptide forming a precursor peptide and which is required for its entrance into the secretory pathway. The term "signal sequence" refers to a nucleotide sequence that encodes the signal peptide.

"Promoter" refers to a nucleotide sequence, usually upstream (5') to its coding sequence, which controls the expression of the coding sequence by providing the recognition for RNA polymerase and other factors required for proper transcription. "Promoter" includes a minimal promoter that is a short DNA sequence comprised of a TATA- box and other sequences that serve to specify the site of transcription initiation, to which regulatory elements are added for control of expression. "Promoter" also refers to a nucleotide sequence that includes a minimal promoter plus regulatory elements that is capable of controlling the expression of a coding sequence or functional RNA. This type of promoter sequence consists of proximal and more distal upstream elements, the latter elements often referred to as enhancers. Accordingly, an "enhancer" is a DNA sequence that can stimulate promoter activity and may be an innate element of the promoter or a heterologous element inserted to enhance the level or tissue specificity of a promoter. Promoters may be derived in their entirety from a native gene, or be composed of different elements derived from different promoters found in nature, or even be comprised of synthetic DNA segments. A promoter may also contain DNA sequences that are involved in the binding of protein factors that control the effectiveness of transcription initiation in response to physiological or developmental conditions.

The "initiation site" is the position surrounding the first nucleotide that is part of the transcribed sequence, which is also defined as position +1. With respect to this site all other sequences of the gene and its controlling regions are numbered. Downstream sequences (*i.e.* further protein encoding sequences in the 3' direction) are denominated positive, while upstream sequences (mostly of the controlling regions in the 5' direction) are denominated negative.

Promoter elements, particularly a TATA element, that are inactive or that have greatly reduced promoter activity in the absence of upstream activation are referred to as "minimal or core promoters." In the presence of a suitable transcription factor, the minimal promoter functions to permit transcription. A "minimal or core promoter" thus consists only of all basal elements needed for transcription initiation, *e.g.*, a TATA box and/or an initiator.

"Constitutive expression" refers to expression using a constitutive or regulated promoter. "Conditional" and "regulated expression" refer to expression controlled by a regulated promoter.

As used herein, the term "operably linked" refers to a linkage of two elements in a functional relationship. For example, "operably linked" may refer to a linkage of polynucleotide elements or polypeptide elements in a functional relationship. A nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, a regulatory DNA sequence is said to be "operably linked to" or "associated with" a DNA sequence that codes for an RNA or a polypeptide if the two sequences are situated such

that the regulatory DNA sequence affects expression of the coding DNA sequence (*i.e.*, that the coding sequence or functional RNA is under the transcriptional control of the promoter).

Coding sequences can be operably-linked to regulatory sequences in sense or antisense orientation. "Operably-linked" also refers to the association two chemical moieties so that the function of one is affected by the other, *e.g.*, an arrangement of elements wherein the components so described are configured so as to perform their usual function.

"Expression" refers to the transcription and/or translation in a cell of an endogenous gene, transgene, as well as the transcription and stable accumulation of sense (mRNA) or functional RNA. In the case of antisense constructs, expression may refer to the transcription of the antisense DNA only. Expression may also refer to the production of protein.

"Transcription stop fragment" refers to nucleotide sequences that contain one or more regulatory signals, such as polyadenylation signal sequences, capable of terminating transcription. Examples of transcription stop fragments are known to the art.

"Translation stop fragment" refers to nucleotide sequences that contain one or more regulatory signals, such as one or more termination codons in all three frames, capable of terminating translation. Insertion of a translation stop fragment adjacent to or near the initiation codon at the 5' end of the coding sequence will result in no translation or improper translation. Excision of the translation stop fragment by site-specific recombination will leave a site-specific sequence in the coding sequence that does not interfere with proper translation using the initiation codon.

The terms "*cis*-acting sequence" and "*cis*-acting element" refer to DNA or RNA sequences whose functions require them to be on the same molecule.

The terms "*trans*-acting sequence" and "*trans*-acting element" refer to DNA or RNA sequences whose function does not require them to be on the same molecule.

The following terms are used to describe the sequence relationships between two or more sequences (*e.g.*, nucleic acids, polynucleotides or polypeptides): (a) "reference sequence," (b) "comparison window," (c) "sequence identity," (d) "percentage of sequence identity," and (e) "substantial identity."

(a) As used herein, "reference sequence" is a defined sequence used as a basis for sequence comparison. A reference sequence may be a subset or the entirety of a specified sequence; for example, as a segment of a full length cDNA, gene sequence or peptide sequence, or the complete cDNA, gene sequence or peptide sequence.

(b) As used herein, "comparison window" makes reference to a contiguous and specified segment of a sequence, wherein the sequence in the comparison window may comprise

additions or deletions (*i.e.*, gaps) compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. Generally, the comparison window is at least 20 contiguous nucleotides in length, and optionally can be 30, 40, 50, 100, or longer. Those of skill in the art understand that to avoid a high similarity to a reference
5 sequence due to inclusion of gaps in the sequence a gap penalty is typically introduced and is subtracted from the number of matches.

Methods of alignment of sequences for comparison are well known in the art. Thus, the determination of percent identity between any two sequences can be accomplished using a mathematical algorithm. Non-limiting examples of such mathematical algorithms are the
10 algorithm of Myers and Miller, CABIOS, 4:11 (1988); the local homology algorithm of Smith et al., Adv. Appl. Math., 2:482 (1981); the homology alignment algorithm of Needleman and Wunsch, JMB, 48:443 (1970); the search-for-similarity-method of Pearson and Lipman, Proc. Natl. Acad. Sci. USA, 85:2444 (1988); the algorithm of Karlin and Altschul, Proc. Natl. Acad. Sci. USA, 87:2264 (1990), modified as in Karlin and Altschul, Proc. Natl. Acad. Sci. USA,
15 90:5873 (1993).

Computer implementations of these mathematical algorithms can be utilized for comparison of sequences to determine sequence identity. Such implementations include, but are not limited to: CLUSTAL in the PC/Gene program (available from Intelligenetics, Mountain View, California); the ALIGN program (Version 2.0) and GAP, BESTFIT, BLAST, FASTA,
20 and TFASTA in the Wisconsin Genetics Software Package, Version 8 (available from Genetics Computer Group (GCG), 575 Science Drive, Madison, Wisconsin, USA). Alignments using these programs can be performed using the default parameters. The CLUSTAL program is well described by Higgins et al., Gene, 73:237 (1988); Higgins et al., CABIOS, 5:151 (1989); Corpet et al., Nucl. Acids Res., 16:10881 (1988); Huang et al., CABIOS, 8:155 (1992); and Pearson et al., Meth. Mol. Biol., 24:307 (1994). The ALIGN program is based on the algorithm of Myers and Miller, *supra*. The BLAST programs of Altschul et al., JMB, 215:403 (1990); Nucl. Acids Res., 25:3389 (1990), are based on the algorithm of Karlin and Altschul *supra*.
25

Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (available on the world wide web at ncbi.nlm.nih.gov).

This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold. These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are then
30

extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when the cumulative alignment score falls off by the quantity X from its maximum achieved value, the cumulative score goes to zero or below due to the accumulation of one or more negative-scoring residue alignments, or the end of either sequence is reached.

In addition to calculating percent sequence identity, the BLAST algorithm also performs a statistical analysis of the similarity between two sequences. One measure of similarity provided by the BLAST algorithm is the smallest sum probability ($P(N)$), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a test nucleic acid sequence is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid sequence to the reference nucleic acid sequence is less than about 0.1, more specifically less than about 0.01, and most specifically less than about 0.001.

To obtain gapped alignments for comparison purposes, Gapped BLAST (in BLAST 2.0) can be utilized as described in Altschul et al., *Nucleic Acids Res.* 25:3389 (1997).

Alternatively, PSI-BLAST (in BLAST 2.0) can be used to perform an iterated search that detects distant relationships between molecules. See Altschul et al., *supra*. When utilizing BLAST, Gapped BLAST, PSI-BLAST, the default parameters of the respective programs (*e.g.*, BLASTN for nucleotide sequences, BLASTX for proteins) can be used. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, a cutoff of 100, $M=5$, $N=-4$, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength (W) of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix. See the world wide web at ncbi.nlm.nih.gov. Alignment may also be performed manually by visual inspection.

For purposes of the present invention, comparison of sequences for determination of percent sequence identity to another sequence may be made using the BlastN program (version 1.4.7 or later) with its default parameters or any equivalent program. By "equivalent program" is intended any sequence comparison program that, for any two sequences in question, generates an alignment having identical nucleotide or amino acid residue matches and an identical percent sequence identity when compared to the corresponding alignment generated by the preferred program.

(c) As used herein, "sequence identity" or "identity" in the context of two nucleic acid or polypeptide sequences makes reference to a specified percentage of residues in the two sequences that are the same when aligned for maximum correspondence over a specified comparison window, as measured by sequence comparison algorithms or by visual inspection.

5 When percentage of sequence identity is used in reference to proteins it is recognized that residue positions which are not identical often differ by conservative amino acid substitutions, where amino acid residues are substituted for other amino acid residues with similar chemical properties (*e.g.*, charge or hydrophobicity) and therefore do not change the functional properties of the molecule. When sequences differ in conservative substitutions, the percent sequence
10 identity may be adjusted upwards to correct for the conservative nature of the substitution. Sequences that differ by such conservative substitutions are said to have "sequence similarity" or "similarity." Means for making this adjustment are well known to those of skill in the art. Typically this involves scoring a conservative substitution as a partial rather than a full mismatch, thereby increasing the percentage sequence identity. Thus, for example, where an
15 identical amino acid is given a score of 1 and a non-conservative substitution is given a score of zero, a conservative substitution is given a score between zero and 1. The scoring of conservative substitutions is calculated, *e.g.*, as implemented in the program PC/GENE (Intelligenetics, Mountain View, California).

(d) As used herein, "percentage of sequence identity" means the value determined by
20 comparing two optimally aligned sequences over a comparison window, wherein the portion of the sequence in the comparison window may comprise additions or deletions (*i.e.*, gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences
25 to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison, and multiplying the result by 100 to yield the percentage of sequence identity.

(e)(i) The term "substantial identity" of sequences means that a polynucleotide comprises a sequence that has at least 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, or 79%, at least
30 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, or 89%, at least 90%, 91%, 92%, 93%, or 94%, and at least 95%, 96%, 97%, 98%, or 99% sequence identity, compared to a reference sequence using one of the alignment programs described using standard parameters. One of skill in the art will recognize that these values can be appropriately adjusted to determine corresponding identity of proteins encoded by two nucleotide sequences by taking into account

codon degeneracy, amino acid similarity, reading frame positioning, and the like. Substantial identity of amino acid sequences for these purposes normally means sequence identity of at least 70%, at least 80%, 90%, at least 95%.

Another indication that nucleotide sequences are substantially identical is if two molecules hybridize to each other under stringent conditions (see below). Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. However, stringent conditions encompass temperatures in the range of about 1°C to about 20°C, depending upon the desired degree of stringency as otherwise qualified herein. Nucleic acids that do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides they encode are substantially identical. This may occur, *e.g.*, when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code. One indication that two nucleic acid sequences are substantially identical is when the polypeptide encoded by the first nucleic acid is immunologically cross reactive with the polypeptide encoded by the second nucleic acid.

(e)(ii) The term "substantial identity" in the context of a peptide indicates that a peptide comprises a sequence with at least 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, or 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, or 89%, at least 90%, 91%, 92%, 93%, or 94%, or 95%, 96%, 97%, 98% or 99%, sequence identity to the reference sequence over a specified comparison window. Optimal alignment is conducted using the homology alignment algorithm of Needleman and Wunsch, *J. Mol. Biol.* 48:443 (1970). An indication that two peptide sequences are substantially identical is that one peptide is immunologically reactive with antibodies raised against the second peptide. Thus, a peptide is substantially identical to a second peptide, for example, where the two peptides differ only by a conservative substitution.

For sequence comparison, typically one sequence acts as a reference sequence to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are input into a computer, subsequence coordinates are designated if necessary, and sequence algorithm program parameters are designated. The sequence comparison algorithm then calculates the percent sequence identity for the test sequence(s) relative to the reference sequence, based on the designated program parameters.

As noted above, another indication that two nucleic acid sequences are substantially identical is that the two molecules hybridize to each other under stringent conditions. The phrase "hybridizing specifically to" refers to the binding, duplexing, or hybridizing of a molecule only to a particular nucleotide sequence under stringent conditions when that sequence is present in a complex mixture (*e.g.*, total cellular) DNA or RNA. "Bind(s) substantially" refers

to complementary hybridization between a probe nucleic acid and a target nucleic acid and embraces minor mismatches that can be accommodated by reducing the stringency of the hybridization media to achieve the desired detection of the target nucleic acid sequence.

"Stringent hybridization conditions" and "stringent hybridization wash conditions" in the context of nucleic acid hybridization experiments such as Southern and Northern hybridizations are sequence dependent, and are different under different environmental parameters. Longer sequences hybridize specifically at higher temperatures. The thermal melting point (T_m) is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Specificity is typically the function of post-hybridization washes, the critical factors being the ionic strength and temperature of the final wash solution.

By "variant" polypeptide is intended a polypeptide derived from the native protein by deletion (so-called truncation) or addition of one or more amino acids to the N-terminal and/or C-terminal end of the native protein; deletion or addition of one or more amino acids at one or more sites in the native protein; or substitution of one or more amino acids at one or more sites in the native protein. Such variants may result from, for example, genetic polymorphism or from human manipulation. Methods for such manipulations are generally known in the art.

Thus, the polypeptides of the invention may be altered in various ways including amino acid substitutions, deletions, truncations, and insertions. Methods for such manipulations are generally known in the art. For example, amino acid sequence variants of the polypeptides can be prepared by mutations in the DNA. Methods for mutagenesis and nucleotide sequence alterations are well known in the art. See, for example, Kunkel, Proc. Natl. Acad. Sci. USA, 82:488 (1985); Kunkel et al., Meth. Enzymol., 154:367 (1987); U. S. Patent No. 4,873,192; Walker and Gaastra, Techniques in Mol. Biol. (MacMillan Publishing Co. (1983), and the references cited therein. Guidance as to appropriate amino acid substitutions that do not affect biological activity of the protein of interest may be found in the model of Dayhoff et al., Atlas of Protein Sequence and Structure (Natl. Biomed. Res. Found. 1978). Conservative substitutions, such as exchanging one amino acid with another having similar properties, are preferred.

Thus, the genes and nucleotide sequences of the invention include both the naturally occurring sequences as well as mutant forms. Likewise, the polypeptides of the invention encompass naturally occurring proteins as well as variations and modified forms thereof. Such variants will continue to possess the desired activity. The deletions, insertions, and substitutions of the polypeptide sequence encompassed herein are not expected to produce radical changes in the characteristics of the polypeptide. However, when it is difficult to predict the exact effect of

the substitution, deletion, or insertion in advance of doing so, one skilled in the art will appreciate that the effect will be evaluated by routine screening assays.

Individual substitutions deletions or additions that alter, add or delete a single amino acid or a small percentage of amino acids (typically less than 5%, more typically less than 1%) in an encoded sequence are "conservatively modified variations," where the alterations result in the substitution of an amino acid with a chemically similar amino acid. Conservative substitution tables providing functionally similar amino acids are well known in the art. The following five groups each contain amino acids that are conservative substitutions for one another: Aliphatic: Glycine (G), Alanine (A), Valine (V), Leucine (L), Isoleucine (I); Aromatic: Phenylalanine (F), Tyrosine (Y), Tryptophan (W); Sulfur-containing: Methionine (M), Cysteine (C); Basic: Arginine (R), Lysine (K), Histidine (H); Acidic: Aspartic acid (D), Glutamic acid (E), Asparagine (N), Glutamine (Q). In addition, individual substitutions, deletions or additions which alter, add or delete a single amino acid or a small percentage of amino acids in an encoded sequence are also "conservatively modified variations."

The term "transformation" refers to the transfer of a nucleic acid fragment into the genome of a host cell, resulting in genetically stable inheritance. Host cells containing the transformed nucleic acid fragments are referred to as "transgenic" cells, and organisms comprising transgenic cells are referred to as "transgenic organisms".

"Transformed," "transgenic," and "recombinant" refer to a host cell or organism into which a heterologous nucleic acid molecule has been introduced. The nucleic acid molecule can be stably integrated into the genome generally known in the art and are disclosed in Sambrook and Russell, *supra*. See also Innis et al., PCR Protocols, Academic Press (1995); and Gelfand, PCR Strategies, Academic Press (1995); and Innis and Gelfand, PCR Methods Manual, Academic Press (1999). Known methods of PCR include, but are not limited to, methods using paired primers, nested primers, single specific primers, degenerate primers, gene-specific primers, vector-specific primers, partially mismatched primers, and the like. For example, "transformed," "transformant," and "transgenic" cells have been through the transformation process and contain a foreign gene integrated into their chromosome. The term "untransformed" refers to normal cells that have not been through the transformation process.

Compositions and Kits

Certain embodiments provide a composition comprising an anti-SARS-CoV-2 nanobody as described herein, or a polypeptide/protein molecule comprising such a nanobody, or a vector as described herein and a carrier. In certain embodiments, the composition is a pharmaceutical

composition comprising a pharmaceutically acceptable carrier. In certain embodiments, the composition is a liquid composition. In certain embodiments, the composition is a solid composition (e.g., powder or lyophilized formulation). In certain embodiments, the composition is a lyophilized composition that further comprises one or more excipients selected from the group consisting of a cryo-lyoprotectant (e.g., trehalose, sucrose) and a bulking agent (e.g., mannitol, glycine). In certain embodiments, the solid composition may be reconstituted (e.g., with water, saline or Dextrose solution) prior to use.

Certain embodiments also provide a kit comprising an isolated anti-SARS-CoV-2 nanobody as described herein, or a polypeptide/protein molecule comprising such a nanobody, or a vector as described herein, packaging material, and instructions for administering the nanobody/polypeptide/protein molecule/vector, to a mammal to treat a SARS-CoV-2 infection. In certain embodiments, the kit further comprises at least one additional therapeutic agent. In certain embodiments, the at least one additional therapeutic agent is useful for preventing or treating a viral infection or inflammation. In certain embodiments, the at least one additional therapeutic agent is an antibody or a nanobody.

In certain embodiments, the kit comprises a syringe (e.g., a pre-filled syringe) or a vial comprising the composition as described herein.

In certain embodiments, the kit further comprises an atomizer nozzle for nasal or pulmonary delivery, wherein the atomizer nozzle is or could be fitted with the syringe or vial to produce a spray or mist.

In certain embodiments, the kit further comprises a needle that is or could be fitted with the syringe (e.g., to deliver subcutaneous, intradermal or intramuscular injection).

Methods of Use

Certain embodiments provide a method of detecting the presence of SARS-CoV-2 in a cell, the method comprising contacting the cell with an isolated anti-SARS-CoV-2 nanobody, polypeptide or protein molecule described herein and detecting whether a complex is formed between the anti-SARS-CoV-2 nanobody, polypeptide or protein molecule and SARS-CoV-2.

In certain embodiments, the cell is contacted *in vitro*. In certain embodiments, the cell is contacted *in vivo*.

Certain embodiments provide a method of inhibiting the activity of SARS-CoV-2, comprising contacting SARS-CoV-2 with an isolated anti-SARS-CoV-2 nanobody as described herein, or a polypeptide or protein molecule as described herein (e.g., under conditions suitable for binding between the nanobody/polypeptide/protein and SARS-CoV-2, such as between the

nanobody/polypeptide/protein and the SARS-CoV-2 RBD). In certain embodiments, binding between SARS-CoV-2 and ACE2 is inhibited. Thus, certain embodiments also provide a method for inhibiting the binding between SARS-CoV-2 and ACE2, comprising contacting SARS-CoV-2 with an isolated anti-SARS-CoV-2 nanobody as described herein, or a polypeptide or protein molecule as described herein (e.g., under conditions suitable for binding between the nanobody/polypeptide/protein and SARS-CoV-2, such as between the nanobody/polypeptide/protein and the SARS-CoV-2 RBD).

In certain embodiments, the SARS-CoV-2 protein is contacted *in vitro*. In certain embodiments, the SARS-CoV-2 protein is contacted *in vivo*. In certain embodiments, the SARS-CoV-2 protein is contacted extracellularly. In certain embodiments, the SARS-CoV-2 protein is contacted intracellularly (e.g., intracellularly delivered or expressed nanobody binds SARS-CoV-2 within an infected cell). Methods for measuring the activity of SARS-CoV-2 (e.g., ability to bind ACE2) are known in the art. For example, in certain embodiments, an assay described herein may be used. In certain embodiments, a nanobody, polypeptide or protein molecule of the invention inhibits the activity of SARS-CoV-2 (e.g., its ability to bind ACE2) by at least about 5%, at least about 10%, at least about 15%, at least about 20%, at least about 25%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 95%, at least about 99% or at least about 100% as compared to a control.

Certain embodiments also provide a method for treating or preventing a SARS-CoV-2 infection in a mammal, comprising administering an effective amount of an isolated anti-SARS-CoV-2 nanobody, polypeptide protein molecule, or a vector as described herein to the mammal.

In certain embodiments, the method further comprises administering at least one additional therapeutic agent to the mammal. In certain embodiments, the at least one additional therapeutic agent is useful for treating a viral infection or inflammation. In certain embodiments, the at least one additional therapeutic agent is an antibody or a nanobody.

Certain embodiments provide an isolated anti-SARS-CoV-2 nanobody, polypeptide, protein molecule, or vector as described herein for the prophylactic or therapeutic treatment of a SARS-CoV-2 infection.

Certain embodiments provide the use of an isolated anti-SARS-CoV-2 nanobody, polypeptide, protein molecule or vector as described herein to prepare a medicament for the treatment of a SARS-CoV-2 infection in a mammal.

Certain embodiments provide an isolated anti-SARS-CoV-2 nanobody, polypeptide, protein molecule or vector as described herein for use in medical therapy.

Administration

For *in vivo* use, a nanobody of the invention, a polypeptide or protein molecule comprising such a nanobody, or a vector as described herein is generally incorporated into a pharmaceutical composition prior to administration. Within such compositions, one or more nanobodies, polypeptides, protein molecules or vectors of the invention may be present as active ingredient(s) (i.e., are present at levels sufficient to provide a statistically significant effect on the symptoms of a relevant disease (e.g., a SARS-CoV-2 infection), as measured using a representative assay). A pharmaceutical composition comprises one or more such nanobodies, polypeptides, protein molecules or vectors in combination with any pharmaceutically acceptable carrier(s) known to those skilled in the art to be suitable for the particular mode of administration. In addition, other pharmaceutically active ingredients (including other therapeutic agents) may, but need not, be present within the composition.

The term “therapeutically effective amount,” in reference to treating a disease state/condition, refers to an amount of a nanobody, polypeptide, protein molecule or vector either alone or as contained in a pharmaceutical composition that is capable of having any detectable, positive effect on any symptom, aspect, or characteristics of a disease state/condition when administered as a single dose or in multiple doses. Such effect need not be absolute to be beneficial.

The terms “treat” and “treatment” refer to both therapeutic treatment and prophylactic or preventative measures, wherein the object is to prevent or decrease an undesired physiological change or disorder, such as a SARS-CoV-2 infection. For purposes of this invention, beneficial or desired clinical results include, but are not limited to, alleviation of symptoms, diminishment of extent of disease, stabilized (*i.e.*, not worsening) state of disease, delay or slowing of disease progression, amelioration or palliation of the disease state, and remission (whether partial or total), whether detectable or undetectable. “Treatment” can also mean prolonging survival as compared to expected survival if not receiving treatment. Those in need of treatment include those already with the condition or disorder as well as those prone to have the condition or disorder or those in which the condition or disorder is to be prevented.

In certain embodiments, the present nanobodies, polypeptides, or protein molecule may be systemically administered, *e.g.*, orally, in combination with a pharmaceutically acceptable vehicle such as an inert diluent or an assimilable edible carrier. They may be enclosed in hard or soft shell gelatin capsules, may be compressed into tablets, or may be incorporated directly with the food of the patient's diet. For oral therapeutic administration, the nanobody, polypeptide or protein molecule may be combined with one or more excipients and used in the form of

ingestible tablets, buccal tablets, troches, capsules, elixirs, suspensions, syrups, wafers, and the like. Such compositions and preparations should contain at least 0.1% of a nanobody, polypeptide or protein molecule of the present invention. The percentage of the compositions and preparations may, of course, be varied and may conveniently be between about 2 to about 5 60% of the weight of a given unit dosage form. The amount of nanobody, polypeptide or protein molecule in such therapeutically useful compositions is such that an effective dosage level will be obtained.

The tablets, troches, pills, capsules, and the like may also contain the following: binders such as gum tragacanth, acacia, corn starch or gelatin; excipients such as dicalcium phosphate; a 10 disintegrating agent such as corn starch, potato starch, alginic acid and the like; a lubricant such as magnesium stearate; and a sweetening agent such as sucrose, fructose, lactose or aspartame or a flavoring agent such as peppermint, oil of wintergreen, or cherry flavoring may be added. When the unit dosage form is a capsule, it may contain, in addition to materials of the above type, a liquid carrier, such as a vegetable oil or a polyethylene glycol. Various other materials 15 may be present as coatings or to otherwise modify the physical form of the solid unit dosage form. For instance, tablets, pills, or capsules may be coated with gelatin, wax, shellac or sugar and the like. A syrup or elixir may contain the nanobody, polypeptide or protein molecule, sucrose or fructose as a sweetening agent, methyl and propylparabens as preservatives, a dye and flavoring such as cherry or orange flavor. Of course, any material used in preparing any 20 unit dosage form should be pharmaceutically acceptable and substantially non-toxic in the amounts employed. In addition, the nanobody, polypeptide or protein molecule may be incorporated into sustained-release preparations and devices.

The nanobody, polypeptide or protein molecule, or a vector as described herein may also be administered subcutaneously, intradermally, intranasally, intramuscularly, intravenously or 25 intraperitoneally by infusion or injection. The nanobody, polypeptide or protein molecule, or a vector as described herein may also be administered via intranasal and/or pulmonary delivery (e.g., delivered as a spray or mist). Additionally, the nanobody, polypeptide, protein molecule or vector may be administered by local injection, such as by intrathecal injection, epidural injection or peri-neural injection using a scope. Solutions of the nanobody, polypeptide, protein 30 molecule or vector may be prepared in water, optionally mixed with a nontoxic surfactant. Dispersions can also be prepared in glycerol, liquid polyethylene glycols, triacetin, and mixtures thereof and in oils. Under ordinary conditions of storage and use, these preparations contain a preservative to prevent the growth of microorganisms.

The pharmaceutical dosage forms suitable for injection or infusion can include sterile

aqueous solutions or dispersions or sterile powders comprising the nanobody, polypeptide, protein molecule or vector that are adapted for the extemporaneous preparation of sterile injectable or infusible solutions or dispersions, optionally encapsulated in liposomes. In all cases, the ultimate dosage form should be sterile, fluid and stable under the conditions of manufacture and storage. The liquid carrier or vehicle can be a solvent or liquid dispersion medium comprising, for example, water, ethanol, a polyol (for example, glycerol, propylene glycol, liquid polyethylene glycols, and the like), vegetable oils, nontoxic glyceryl esters, and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the formation of liposomes, by the maintenance of the required particle size in the case of dispersions or by the use of surfactants. The prevention of the action of microorganisms can be brought about by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thimerosal, and the like. In many cases, it will be useful to include isotonic agents, for example, sugars, buffers or sodium chloride. Prolonged absorption of the injectable compositions can be brought about by the use in the compositions of agents delaying absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions are prepared by incorporating the nanobody, polypeptide, protein molecule or vector in the required amount in the appropriate solvent with various of the other ingredients enumerated above, as required, followed by filter sterilization. In the case of sterile powders for the preparation of sterile injectable solutions, the methods of preparation are vacuum drying and the freeze drying techniques, which yield a powder of the nanobody, polypeptide, protein molecule or vector plus any additional desired ingredient present in the previously sterile-filtered solutions.

For topical administration, the present nanobodies, polypeptides or protein molecules may be applied in pure form, *i.e.*, when they are liquids. However, it will generally be desirable to administer them to the skin as compositions or formulations, in combination with a dermatologically acceptable carrier, which may be a solid or a liquid.

Useful solid carriers include finely divided solids such as talc, clay, microcrystalline cellulose, silica, alumina and the like. Useful liquid carriers include water, alcohols or glycols or water-alcohol/glycol blends, in which the present nanobodies, polypeptides or protein molecules can be dissolved or dispersed at effective levels, optionally with the aid of non-toxic surfactants. Adjuvants such as fragrances and additional antimicrobial agents can be added to optimize the properties for a given use. The resultant liquid compositions can be applied from absorbent pads, used to impregnate bandages and other dressings, or sprayed onto the affected area using pump-type or aerosol sprayers.

Thickeners such as synthetic polymers, fatty acids, fatty acid salts and esters, fatty alcohols, modified celluloses or modified mineral materials can also be employed with liquid carriers to form spreadable pastes, gels, ointments, soaps, and the like, for application directly to the skin of the user.

5 Examples of useful dermatological compositions that can be used to deliver the nanobodies, polypeptides or protein molecules of the present invention to the skin are known to the art; for example, see Jacquet et al. (U.S. Pat. No. 4,608,392), Geria (U.S. Pat. No. 4,992,478), Smith et al. (U.S. Pat. No. 4,559,157) and Wortzman (U.S. Pat. No. 4,820,508).

10 Useful dosages of the nanobodies, polypeptides, protein molecules or vectors of the present invention can be determined by comparing their *in vitro* activity, and *in vivo* activity in animal models. Methods for the extrapolation of effective dosages in mice, and other animals, to humans are known to the art; for example, see U.S. Pat. No. 4,938,949.

15 The amount of a nanobody, polypeptide, protein molecule or vector of the present invention required for use in treatment will vary with the route of administration, the nature of the condition being treated and the age and condition of the patient and will be ultimately at the discretion of the attendant physician or clinician.

20 The desired dose may conveniently be presented in a single dose or as divided doses administered at appropriate intervals, for example, as two, three, four or more sub-doses per day. The sub-dose itself may be further divided, *e.g.*, into a number of discrete loosely spaced administrations.

25 Nanobodies, polypeptides, protein molecules or vectors of the invention can also be administered in combination with other therapeutic agents and/or treatments, such as other agents or treatments that are useful for the treatment of a SARS-CoV-2 infection. In certain embodiments such an agent is an antibody or a nanobody. Additionally, one or more nanobodies, polypeptides, protein molecules or vectors of the invention, may be administered (*e.g.*, a combination of nanobodies, polypeptides, protein molecules and/or vectors may be administered). Accordingly, one embodiment the invention also provides a composition comprising a nanobody, polypeptide, protein molecule or vector of the invention, at least one other therapeutic agent, and a pharmaceutically acceptable diluent or carrier. The invention also provides a kit comprising a nanobody, polypeptide, protein molecule or vector of the invention, at least one other therapeutic agent, packaging material, and instructions for administering a nanobody, polypeptide, protein molecule or vector of the invention, and the other therapeutic agent or agents to an animal to treat a SARS-CoV-2 infection.

As used herein, the term “therapeutic agent” refers to any agent or material that has a

beneficial effect on the mammalian recipient.

The invention will now be illustrated by the following non-limiting Examples.

5 EXAMPLE 1.

The impact of COVID-19 pandemic on global health and economy has been unprecedented in modern medicine era. To battle the COVID-19 global pandemic, potent and economical anti-SARS-CoV-2 drugs are urgently needed for the world's mass populations. For fast evolving viral pandemics such as this one, small-molecule drugs often take too long to develop and their specificity is generally low due to their small size, leading to side effects; repurposed drugs particularly lack specificity since they were developed against other targets; and traditional antibodies are easy to develop and their specificity is usually high, but they are expensive and also typically have poor stability and pharmacokinetics due to their large size.

Described herein is the development of a series of nanobodies and nanobody-Fc fusion polypeptides, which bind to the SARS-CoV-2 spike protein with high affinity and potentially block SARS-CoV-2 pseudovirus infection in human cells. In addition to their potency, the Nanosota-1 drugs can be prepared in bacteria in large quantities, demonstrates superior stability, and have an extended in vivo half-life. Overall, the Nanosota-1 series are potent and economical anti-SARS-CoV-2 drugs, which may contribute to the global battle against COVID-19.

20 MATERIALS AND METHODS

Camelids nanobody library

A high diversity naïve nanobody phage display library was developed from camelids. This library was constructed using B cells isolated from the spleen, bone marrow and blood of over a dozen non-immunized llamas and alpacas. As a result of the large number of animals used in the construction of the library and the high quality of the source tissue, this naïve nanobody library has an unparalleled diversity of 7.5×10^{10} . This library routinely yields subnanomolar binders when screen against recombinant proteins immobilized on immunotubes or magnetic beads.

30 Camelids nanobody library screen

The naïve camelids nanobody phage library was used in the bio-panning. The nanobody genes were constructed in the PADL22c vector with a His₆ tag and an HA tag at the end of the gene. Four rounds of panning were performed to obtain the RBD specific nanobodies with high binding affinity. The amounts of the RBD-Fc antigen used to coat in the immune tubes in each

round were 75 µg, 50 µg, 25 µg and 10 µg, respectively. For each round, the coated tube was blocked with 2% milk for 2 hours, and then incubated with the pre-blocked phage for 1 hour. After incubation, the coated tube was washed 5 times with PBS (with 0.1% Tween-20) and 5 times with PBS in round 1, 10 times with PBS (with 0.1% Tween-20) and 10 times with PBS in round 2, 20 times with PBS (with 0.1% Tween-20) and 20 times with PBS in round 3, 25 times with PBS (with 0.1% Tween-20) and 25 times with PBS in round 4, respectively. The retained phages were eluted using 1 ml 100 mM triethylamine and neutralized with 500 µl 1 M Tris-HCl pH 7.5. The eluted phages were amplified in TG1 *E. coli* cells and rescued with M13K07 helper phage. The eluted phages from round 4 were used to infect ss320 *E. coli*, 192 single clones were picked into 2YT A medium and nanobody expression were induced with 1 mM IPTG. The supernatants were used to perform ELISA to select strong binders. Briefly, RBD-His were coated onto 96-well plate (Costar[®] Assay Plate, CORNING) in PBS at 4 °C overnight. 2 µl ss320 supernatant was added into each well and incubated for 1 hour after being blocked with 2% BSA for 1 hour. Then each well was incubated with HRP-Conjugated anti-HA antibody (3F10, sigma-aldrich) for another 1 hour. After washing, the wells were incubated with 50 µl 3,3',5,5'-Tetramethylbenzidine (TMB) substrate and then stopped with 10 µl 1 N H₂SO₄. Finally, ELISA signals were measured at 450 nM. 24 binders with the highest binding were sent for sequencing. The strongest binder after initial screening was named Nanosota-1A (*see*, Table 1).

Affinity maturation of nanobody Nanosota-1A

Mutations were introduced into the whole gene of the nanobody Nanosota-1A by using error-prone PCR. Two rounds of error-prone PCR were performed using the GeneMorph II Random Mutagenesis Kit (Agilent Technologies). The PADL22c vector was linearized using PCR, and then the error-prone PCR products were cloned into the linearized PADL22c vector using the NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs). The ligated products were purified using DNA Clean and Concentrator[™]-5 (ZYMO RESEARCH) and then electroporated into the TG1 *E. coli*. The library size was about 6×10^8 . Three rounds of bio-panning were performed using 25 ng, 10 ng and 2 ng RBD-Fc, respectively. The wash steps were as follows, 5 times with PBS (with 0.1% Tween-20) and 5 times with PBS in round 1, 20 times with PBS (with 0.1% Tween-20) and 20 times with PBS in round 2, and 30 times with PBS (with 0.1% Tween-20) and 30 times with PBS in round 3. The eluted phages from round 3 were used to infect ss320 cells, 96 single clones were picked for ELISA to select strong binders as described above. The strongest binder after affinity maturation was named Nanosota-1B (*see*,

Table 1).

Production of Nanosota-1B

Nanosota-1B was purified from the periplasm of the ss320 *E. coli* from the last round of bio-panning. The ss320 *E. coli* grown in TB medium with 100 mg/L ampicillin, induced by 1 mM IPTG when OD₆₀₀ reached about 1.0, then grown overnight at 30 °C. The bacterial cells were collected and re-suspended in 15 ml TES buffer (0.2 M Tris pH 8, 0.5 mM EDTA, 0.5 M sucrose), incubated on ice for 1 hour and then incubated with 40 ml of TES/4 buffer followed by another one-hour incubation on ice. The supernatant was collected and incubated with imidazole at a final concentration of 20 mM. Then the supernatant was filtered through 0.45 µM filters and loaded onto a Ni-NTA column. Proteins were eluted using gradient imidazole and then further purified using a Superdex 200 PG gel filtration column (GE Healthcare).

For the Fc-tagged Nanosota-1B, the gene was fused to the Fc (human IgG1) and inserted into the pET42b vector. The recombinant plasmid was transformed into Shuffle[®] T7 *E. coli* (New England Biolabs). The induction of protein expression was the same as above. After induction, the bacterial cells were collected, resuspended in PBS and disrupted using the BRANSON Digital Sonifier[®]. The supernatant was filtered and loaded onto a protein A column. Proteins were eluted using glycine pH 2.7 and then purified further with a Superdex 200 PG gel filtration column.

Production of Nanosota-1C

A second round of affinity maturation was performed with Nanosota-1B using methods similar to those described above to generate Nanosota-1C (*see*, Table 1). Nanosota-1C was produced using methods similar to those described above.

For the Fc-tagged Nanosota-1C, the gene was fused to the Fc (human IgG1) and inserted into the pET42b vector. The recombinant plasmid was transformed into Shuffle[®] T7 *E. coli* (New England Biolabs). The induction of protein expression was the same as above. After induction, the bacterial cells were collected, resuspended in PBS and disrupted using the BRANSON Digital Sonifier[®]. The supernatant was filtered and loaded onto a protein A column. Proteins were eluted using glycine pH 2.7 and then purified further with a Superdex 200 PG gel filtration column.

Biacore Measurement

Biacore measurement of the binding interactions between Nanosota-1 drugs (Nanosota-

1A, Nanosota-1B, Nanosota-1C and Nanosota-1C-Fc) and SARS-CoV-2 RBD was performed using methods similar to those described in Shang et al., Nature 581: 221-224 (2020).

Briefly, the SARS-CoV-2 RBD (His tagged on the C terminus) was covalently immobilized (via amine group) to a sensor chip, and Nanosota-1 drugs at different concentrations flowed by. Specifically, Nanosota-1C-Fc flowed by at the following concentrations: 1.25 nM, 2.5 nM, 5 nM, 10 nM, and 20 nM; Nanosota-1C flowed by at the following concentrations: 2.5 nM, 5 nM, 10 nM, 20 nM, 40 nM, and 80 nM; Nanosota-1B flowed by at the following concentrations: 2.5 nM, 5 nM, 10 nM, 20 nM, 40 nM, and 80 nM; and Nanosota-1A flowed by at the following concentrations: 10 nM, 20 nM, 40 nM, 80 nM, 160 nM, and 320 nM.

Pull Down Assay

Competitive binding of Nanosota-1C and human ACE2 (hACE2) to SARS-CoV-2 RBD was evaluated by a protein pull down assay. hACE2 (containing a C-terminal His tag) and Nanosota-1C (containing a C-terminal His tag) were mixed together in different molar ratios, with the amount of hACE2 being kept constant at 2 µg). 0.5 µg SARS-CoV-2 RBD (containing a C-terminal Fc tag) was used to pull down hACE2 and Nanosota-1C, which were detected by Western blot using an anti-His tag antibody.

Gel Filtration Chromatography

Competitive binding of Nanosota-1C and human ACE2 (hACE2) to SARS-CoV-2 RBD was evaluated by gel filtration chromatography. Briefly, hACE2 (containing a C-terminal His tag), Nanosota-1C (containing a C-terminal His tag), and SARS-CoV-2 RBD (containing a C-terminal His tag) were mixed together in molar ratio 1.5:1.5:1. The mixture was subjected to gel filtration chromatography.

Inhibition of SARS-CoV-2 pseudovirus entry

Retroviruses pseudotyped with SARS-CoV-2 spike protein (i.e., SARS-CoV-2 pseudoviruses) were used to enter 293T cells expressing human ACE2 in the presence of different concentrations of Nanosota-1A (i.e., Nanosota-1 before affinity maturation), Nanosota-1B (i.e., Nanosota-1 after affinity maturation), Nanosota-1B-Fc (i.e., Nanosota-1B containing a C-terminal Fc tag), recombinant ACE2, and recombinant ACE2-Fc (containing a C-terminal Fc tag).

A second experiment was performed using similar methods. Briefly, retroviruses

pseudotyped with SARS-CoV-2 spike protein (i.e., SARS-CoV-2 pseudoviruses) were used to enter human ACE2-expressing HEK293T cell in the presence of different concentrations of Nanosota-1C or Nanosota-1C-Fc. Entry efficiency was characterized as luciferase signal accompanying entry. Pseudovirus entry in the absence of any drug was taken as 100%.

5

***In vivo* Bioavailability Assay**

Bioavailability of Nanosota-1 was evaluated in mice. Briefly, mice were intravenously injected with Nanosota-1C-Fc or Nanosota-1C (100 µg in 100 µl for each). PBS buffer was used as a negative control. Sera were collected at varying time points. Sera were then tested by ELISA for SARS-CoV-2 binding using ELISA. The data are presented as mean A450 values; error bars represent the SEM of the mean (n = 3, meaning that there were three mice in each group).

10

Stability Assay

The stability of Nanosota-1C-Fc was evaluated at different temperatures. Nanosota-1C-Fc was stored at different temperatures for 1 week, and then was assayed for its ability to bind to SARS-CoV-2 RBD (containing a C-terminal His tag) using ELISA.

15

RESULTS

20

Binding interactions between Nanosota-1 drugs and SARS-CoV-2 RBD were evaluated using Biacore measurements (Figs. 1A-1E). These measurements indicated that Nanosota-1C-Fc binds to SARS-CoV-2 RBD nearly 3,000 times more strongly than the ACE2 receptor binds to the RBD (*see*, Fig. 1A).

25

Competitive binding of Nanosota-1C and human ACE2 (hACE2) to SARS-CoV-2 RBD was detected using a protein pull down assay (Fig. 3A). The results showed that as the amount of Nanosota-1C in the solution increased, the amount of Nanosota-1C pulled down also increased, while the amount of hACE2 pulled down decreased. At 1:1 molar ratio, more Nanosota-1C was pulled down than hACE2. Thus, Nanosota-1C and hACE2 competitively bind to SARS-CoV-2 RBD, and Nanosota-1C has a higher RBD-binding affinity than hACE2.

30

Competitive binding of Nanosota-1C and human ACE2 (hACE2) to SARS-CoV-2 RBD was also evaluated by gel filtration chromatography. As shown in Fig. 3B, four peaks were obtained. SDS-PAGE (stained using coomassie blue) showed that peak 1 contained the complex of hACE2 and RBD, peak 2 contained hACE2, peak 3 contained the complex of Nanosota-1C and RBD, and peak 4 contained Nanosota-1C. No ternary complex of hACE2, Nanosota-1C and

RBD was detected. Therefore, these results demonstrate that hACE2 and Nanosota-1C bind competitively to RBD.

The results of a pseudovirus entry inhibition assay (Figures 2A-2B) showed that in a dose dependent manner Nanosota-1A, Nanosota-1B, Nanosota-1B-Fc, Nanosota-1C and Nanosota-1C-Fc potently inhibit SARS-CoV-2 pseudovirus entry into human cell expressing ACE2.

The bioavailability of Nanosota-1C and Nanosota-1C-Fc was also evaluated in mice. The results showed that the half-life of Nanosota-1C-Fc in mice is > 10 days, while that of Nanosota-1C is several hours (Figs. 4A-4C).

Additionally, the stability of Nanosota-1C-Fc was evaluated at different temperatures. As compared to -80 °C, Figure 5 shows that Nanosota-1C-Fc's RBD-binding function was intact after being stored at 25 °C or 4°C for a week and only dropped slightly after being stored at 37 °C for a week.

EXAMPLE 2. *Novel nanobodies as potent and cost-effective anti-SARS-CoV-2 drugs*

The COVID-19 pandemic is causing catastrophic devastation to global health and economy. Anti-SARS-CoV-2 drugs are urgently needed to save lives and control the damage caused by the disease. Yet drug development against such a fast-evolving viral pandemic faces daunting challenges. Small-molecule drugs not only take too long to develop, but also generally have low specificity and cause side effects due to their small size. Conventional antibodies take less time to develop and have high specificity, but their high production cost makes them an unrealistic solution to the global pandemic. Overcoming these challenges, a novel single-domain antibody - a nanobody named *Nanosota-1* - from llama nanobody phage display library has been developed. This nanobody bound to the receptor-binding domain of SARS-CoV-2 spike protein with high affinity and potently inhibited SARS-CoV-2 infection of target cells. Importantly, it is believed that for the first time it has been demonstrated that nanobodies like *Nanosota-1* effectively protect animal models from SARS-CoV-2 challenge. The cost effectiveness and pharmacokinetics that make *Nanosota-1* a realistic solution to the global pandemic have also been characterized - low-cost production in bacteria in large quantities, superior stability for storage and transportation, and excellent *in vivo* half-life and tissue permeability. As a potent and cost-effective anti-SARS-CoV-2 drug, *Nanosota-1* can potentially contribute to the global battle against COVID-19.

Abstract

Battling the COVID-19 global pandemic requires potent and cost-effective anti-SARS-

CoV-2 drugs for the world's vast population. As described herein, a novel single-domain antibody (i.e., nanobody), *Nanosota-1*, from a naïve llama nanobody phage display library has been developed. After two rounds of affinity maturation, *Nanosota-1* containing an Fc tag (*Nanosota-1-Fc*) bound to the receptor-binding domain of SARS-CoV-2 spike protein ~3000 times more tightly than the viral receptor ACE2 did, blocking ACE2 binding to SARS-CoV-2. Moreover, *Nanosota-1-Fc* inhibited SARS-CoV-2 pseudovirus infection ~160 times more efficiently than recombinant ACE2 did. Importantly, animal testing in two models showed that, administered at a single dose, *Nanosota-1-Fc* protected hamsters and mice from SARS-CoV-2 infections both prophylactically and therapeutically. Unlike conventional antibody drugs, *Nanosota-1-Fc* can be produced in bacteria with high yield at low cost and is highly stable in the environment. *Nanosota-1-Fc* also demonstrated an excellent *in vivo* half-life, stability, and a high tissue permeability. If further validated in human clinical trials, *Nanosota-1-Fc* can be an effective, inexpensive, and realistic solution to the COVID-19 pandemic.

Introduction

The novel coronavirus SARS-CoV-2 has led to the COVID-19 global pandemic (Li Q, *et al.* (2020) Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *N Engl J Med.*; Huang C, *et al.* (2020) Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet.*), causing devastating health and economic damages. Beside vaccines, anti-SARS-CoV-2 drugs are urgently needed to treat patients, save lives, and revive economies. Yet daunting challenges confront the development of such drugs. Repurposed drugs, developed against other viruses, generally have low specificity against SARS-CoV-2. Small-molecule drugs can specifically target SARS-CoV-2, but they often take years to develop and tend to cause side effects due to their small size. Conventional antibodies are quick to develop and have high specificity; however, they can only be produced in mammalian cells, leading to low expression yields and high production costs (Salazar *et al.*, (2017) Antibody therapies for the prevention and treatment of viral infections. *NPJ vaccines* 2:19; Breedveld FC (2000) *Lancet* 355(9205):735-740). Realistic solutions to COVID-19 require potent and cost-effective drugs. Herein it is reported that a novel anti-SARS-CoV-2 nanobody drug, *Nanosota-1*, has been developed and has been validated in animal testing, providing a possible solution to the COVID-19 pandemic.

Nanobodies - single-chain antibodies - are unique antibodies derived from heavy chain-only antibodies found in the camelidae family such as llamas, alpacas, and camels (Fig. 11) (Könning D, *et al.* (2017) *Curr Opin Struct Biol* 45:10-16; De Meyer *et al.* (2014) *Trends*

Biotechnol 32(5):263-270). Because of their nanometer size, low molecular weight, (2.5 nm by 4 nm; ~12-15 kDa) and unique binding domains (e.g., single domain structure), nanobodies offer many advantages and can be superior therapeutics to conventional antibodies: they can be produced in bacteria with high yields, are stable for storage and transportation, can bind to hidden cryptic epitopes on their targets, and possess high tissue permeability (Muyldermans S (2013) *Annu Rev Biochem* 82:775-797; Steeland et al. (2016) *Drug discovery today* 21(7):1076-1113). Although small, nanobodies are capable of binding to their targets with high affinity and specificity due to an extended antigen-binding region (Muyldermans S (2013) *Annu Rev Biochem* 82:775-797; Steeland et al. (2016) *Drug discovery today* 21(7):1076-1113).

Furthermore, they have low toxicity and immunogenicity in humans (Muyldermans S (2013) *Annu Rev Biochem* 82:775-797; Steeland et al. (2016) *Drug discovery today* 21(7):1076-1113). One potential drawback of nanobodies - rapid clearance through kidneys due to their small size - can be overcome by adding tags, for example, by increasing the molecular weight to a desired level that is above the kidney clearance threshold but still much lower than conventional antibodies' molecular weight. Confirming the potency and safety of nanobodies as human therapeutics, a nanobody drug was recently approved to clinically treat a blood clotting disorder (Scully M, et al. (2019) *N Engl J Med* 380(4):335-346). Additionally, due to their superior stability, nanobodies can be inhaled to treat lung diseases (Van Heeke G, et al. (2017) *Pharmacology & therapeutics* 169:47-56) or ingested to treat intestine diseases (Vega CG, et al. (2013) *PLoS Pathog* 9(5):e1003334). Nanobodies are currently being developed to target SARS-CoV-2 (Huo J, et al. (2020) Neutralizing nanobodies bind SARS-CoV-2 spike RBD and block interaction with ACE2. *Nat Struct Mol Biol.*; Schoof M, et al. (2020) *bioRxiv*:2020.2008.2008.238469). However, to date, no nanobodies have been tested for therapeutic efficacy in animal models, and no studies have characterized their cost-effectiveness and pharmacokinetics.

A prime target for nanobodies is the receptor-binding domain (RBD) of SARS-CoV-2 spike protein (Li F (2015) *J Virol* 89(4):1954-1964). A virus-envelope-anchored spike protein guides coronavirus entry into host cells by first binding to a receptor on the host cell surface and then fusing the viral and host membranes (Li F (2016) *Annual review of virology* 3(1):237-261; Perlman S & Netland J (2009) *Nature Reviews Microbiology* 7(6):439-450). Like the related SARS-CoV-1, the RBD of SARS-CoV-2 recognizes human angiotensin-converting enzyme 2 (ACE2) as its receptor (Wan Yet al. (2020) *J Virol* 94(7)). It has been previously shown that SARS-CoV-2 RBD has significantly higher ACE2-binding affinity than SARS-CoV-1 RBD due to several structural changes (Shang J, et al. (2020) *Nature* 581(7807):221-224; Li F, et al.

(2005) *Science* 309(5742):1864-1868). It has been further shown that SARS-CoV-2 RBD is more hidden than SARS-CoV-1 RBD in the entire spike protein as a possible viral strategy for immune evasion (Shang J, *et al.* (2020) *Proc Natl Acad Sci U S A* 117(21):11727-11734). Hence, to efficiently block SARS-CoV-2 binding to ACE2, nanobodies must bind to SARS-CoV-2 RBD more tightly than ACE2 does.

Nanobodies that potently inhibit SARS-CoV-2 have been developed herein. Identified by screening a camelid nanobody phage display library against the SARS-CoV-2 RBD, the *Nanosota-1* series bound potently to the SARS-CoV-2 RBD and were effective at inhibiting SARS-CoV-2 infection. Their RBD-binding features, anti-SARS-CoV-2 potency, cost-effectiveness and pharmacokinetics have been characterized. Importantly, using hamsters and mice as animal models, the nanobodies' prophylactic and therapeutic efficacy has been demonstrated. Produced at high yields, *Nanosota-1C-Fc* is easily scalable for mass production. It also demonstrated excellent *in vitro* thermostability, *in vivo* stability and bioavailability, providing compelling evidence that these nanobodies are potentially realistic solutions to the COVID-19 pandemic.

Results

Nanosota-1 was identified by phage display

To rapidly develop high-quality virus-targeting nanobodies, a highly diverse nanobody phage display library from naïve camelids was constructed. This library was constructed using B cells isolated from the spleen, bone marrow, and blood of over a dozen non-immunized llamas and alpacas. This naïve nanobody library has an unparalleled diversity of 7.5×10^{10} . This library was used to screen nanobodies targeting SARS-CoV-2 RBD (Fig. 6).

To develop SARS-CoV-2 RBD-targeting nanobodies, SARS-CoV-2 RBD was expressed and purified in mammalian cells and then screened against the library for phages that bound to the RBD. Nanobodies from the selected phages were further tested for their ability to neutralize SARS-CoV-2 pseudovirus entry into target cells (see below about the assay). The nanobody that demonstrated the highest preliminary neutralization potency was named *Nanosota-1A* and was subjected to two rounds of affinity maturation. For each round, random mutations were introduced to the whole gene of *Nanosota-1A* through error-prone PCR, and mutant phages were selected for enhanced binding to SARS-CoV-2 RBD. Nanobodies contain four framework regions (FRs) as structural scaffolds and three complementarity-determining regions (CDRs) for antigen binding. The nanobody after the first round of affinity maturation, named *Nanosota-1B*, possessed one mutation in CDR3 and two mutations in FR3 (near CDR3). Affinity maturation of

Nanosota-1B resulted in *Nanosota-1C*, which possessed one mutation in CDR2 and another mutation in FR2. To increase the molecular weight, an Fc-tagged version of *Nanosota-1C*, named *Nanosota-1C-Fc* was constructed to create a bivalent construct with increased molecular weight. These nanobodies contained two humanized mutations Q44G and R45L (Vincke C, *et al.* (2009) *J Biol Chem* 284(5):3273-3284), reducing the already low immunogenicity of nanobodies (Steeland S, Vandenbroucke RE, & Libert C (2016) Nanobodies as therapeutics: big opportunities for small antibodies. *Drug discovery today* 21(7):1076-1113). This series of nanobodies, particularly *Nanosota-1C* and *Nanosota-1C-Fc*, were evaluated as potential anti-SARS-CoV-2 drugs.

***Nanosota-1* tightly bound to the SARS-CoV-2 RBD and completely blocked the binding of ACE2**

To understand the structural basis for how the nanobodies interact with SARS-CoV-2 RBD, the crystal structure of SARS-CoV-2 RBD complexed with *Nanosota-1C* was determined. The structure showed that *Nanosota-1C* binds close to the center of the SARS-CoV-2 RBM (Fig. 16A). Among the fourteen RBM residues that directly interact with *Nanosota-1C*, six also directly interact with human ACE2 (Fig. 17). When the structures of the RBD/*Nanosota-1C* complex and the RBD/ACE2 complex were superimposed together, significant clashes occurred between ACE2 and *Nanosota-1C* (Fig. 16B), suggesting that *Nanosota-1C* binding to the RBD blocks ACE2 binding to the RBD. Moreover, trimeric SARS-CoV-2 spike protein is present in two different conformations: the RBD stands up in the open conformation but lies down in the closed conformation (J. Shang et al., PNAS 117, 11727-11734 (2020); D. Wrapp et al., Science, 367(6483):1260-1263, (2020); and Z. Ke et al., Nature, 588(7838):498-502 (2020)). When the structures of the RBD/*Nanosota-1C* complex and the closed spike were superimposed, no clash was found between RBD-bound *Nanosota-1C* and the rest of the spike protein (Fig. 18A). In contrast, severe clashes were identified between RBD-bound ACE2 and the rest of the spike protein in the closed conformation (Fig. 18B). Additionally, neither RBD-bound *Nanosota-1C* nor RBD-bound ACE2 had clashes with the rest of the spike protein in the open conformation (Fig. 18C, 18D). Thus, *Nanosota-1C* can access the spike protein in both its open and closed conformations, whereas ACE2 can only access the spike protein in its open conformation. Overall, our structural data reveal that *Nanosota-1C* is an ideal RBD-targeting drug candidate that not only blocks virus binding to its receptor, but also accesses its target in the spike protein in different conformations.

To corroborate the structural data on the RBD/*Nanosota-1* interactions, binding experiments were performed between the nanobodies and RBD using recombinant human ACE2

as a comparison. First, the binding affinity between the nanobodies and RBD was measured using surface plasmon resonance (Table B; Figs. 12A-12D). *Nanosota-1A*, *-1B*, and *-1C* bound to the RBD with increasing affinity (K_d - from 228 nM to 14 nM), confirming the success of affinity maturation. *Nanosota-1C-Fc* had the highest RBD-binding affinity ($K_d = 15.7$

5 picomolar), ~3,000 times stronger than the RBD-binding affinity of ACE2. *Nanosota-1C* bound to RBD ~3 times more tightly than ACE2 did. Moreover, compared with ACE2, *Nanosota-1C-Fc* bound to the RBD with a substantially higher k_{on} and a lower k_{off} , demonstrating significantly faster binding and slower dissociation. Second, the competitive binding among *Nanosota-1C*, ACE2, and RBD was investigated using protein pull-down assay (Fig. 7A). ACE2 and

10 *Nanosota-1C* were mixed in different ratios in solution, with the concentration of ACE2 remaining constant; RBD-Fc was added to pull down ACE2 and *Nanosota-1C* from solution. As the concentration of *Nanosota-1C* increased, less ACE2 was pulled down by the RBD. Thus, ACE2 and *Nanosota-1C* bound competitively to the RBD. The above protein pull-down assay was repeated with *Nanosota-1C-Fc* replacing *Nanosota-1C* (Fig. 19). The result confirmed that

15 *Nanosota-1C-Fc* and ACE2 bound competitively to the RBD; it further showed that *Nanosota-1C-Fc* bound to the RBD much more strongly than ACE2 did, consistent with the binding affinity measurement. Third, competitive binding among *Nanosota-1C*, ACE2, and RBD was analyzed using gel filtration chromatography (Fig. 7B). ACE2, *Nanosota-1C*, and RBD were mixed together, with both ACE2 and *Nanosota-1C* in molar excess over the RBD. The mixture

20 was subjected to gel filtration chromatography. No ternary complex of ACE2, *Nanosota-1C*, and RBD had formed; instead, only binary complexes of ACE2/RBD and *Nanosota-1C*/RBD were detected. Hence, the bindings of ACE2 and *Nanosota-1C* to the RBD were mutually exclusive. Overall, *Nanosota-1C-Fc* and *Nanosota-1C* are potent binders of SARS-CoV-2 RBD and their binding to the RBD blocks ACE2 binding to the RBD.

25 ***Nanosota-1C-Fc* potentially neutralized SARS-CoV-2 infection *in vitro* and *in vivo***

Next, the potency of *Nanosota-1* drugs in neutralizing SARS-CoV-2 infection *in vitro* was investigated. Both a SARS-CoV-2 pseudovirus entry assay and an authentic live SARS-CoV-2 infection assay was performed in the presence of *Nanosota-1* drugs (Figs. 8A-8B). For the pseudovirus entry assay, retroviruses pseudotyped with SARS-CoV-2 spike protein (i.e.,

30 SARS-CoV-2 pseudoviruses) were used to enter human ACE2-expressing HEK293T cells in the presence of each drug at different concentrations. The efficacy of each inhibitor was expressed as the concentration capable of neutralizing either 50% or 90% of the entry efficiency (i.e., 50% Neutralizing Dose or ND₅₀ and 90% Neutralizing Dose or ND₉₀). *Nanosota-1C-Fc* neutralized SARS-CoV-2 pseudovirus entry at an ND₅₀ of 0.27 µg/ml and an ND₉₀ of 3.12 µg/ml, both of

which were ~10 times more potent than monovalent *Nanosota-1C* (ND₅₀ of 2.52 µg/ml; ND₉₀ of 23.3 µg/ml) and the first of which was over 100 times (~160 times) more potent than ACE2 (Fig. 8A). Moreover, *Nanosota-1C-Fc* potentially neutralized SARS-CoV-2 pseudovirus bearing a D614G mutation in the spike protein (Fig. 13); this mutation has become prevalent in SARS-CoV-2 strains isolated in many regions of the world (Korber B, *et al.* (2020) *Cell* 182(4):812-827.e819). For the authentic SARS-CoV-2 infection assay, live SARS-CoV-2 was used to infect Vero cells in the presence of each drug at different concentrations. The efficacy of each drug was described as the lowest concentration capable of completely preventing virus-induced cytopathic effect (CPE) in 100% of the wells (i.e., 100% Neutralizing Dose or ND₁₀₀) or 50% of the wells (i.e., ND₅₀). *Nanosota-1C-Fc* had an ND₁₀₀ of 12 µg/ml, whereas *Nanosota-1C* and ACE2 had an ND₅₀ of 37 µg/ml and 2 mg/ml, respectively (Fig. 8B). Hence, *Nanosota-1C-Fc* inhibited authentic live SARS-CoV-2 infection at least ~200 times more effectively than ACE2 did. *Nanosota-1C-Fc* had an ND₅₀ of 0.16 µg/ml, which was significantly more potent than monovalent *Nanosota-1C* and ACE2 (Fig. 20; Fig. 21). Overall, both *Nanosota-1C-Fc* and *Nanosota-1C* potentially inhibit SARS-CoV-2 pseudovirus entry and authentic live SARS-CoV-2 infection in target cells.

The effectiveness of *Nanosota-1C-Fc* at protecting a non-lethal animal model – hamsters – from authentic SARS-CoV-2 infections was then evaluated (Sia S.F., *et al.* (2020) *Nature* 583(7818):834-838). Hamsters were challenged with SARS-CoV-2 (at a titer of 1×10^6 TCID₅₀) via intranasal inoculation. Other than a no-treatment control, four experimental groups of hamsters were injected via intraperitoneal injection with a single dose of *Nanosota-1C-Fc*: (i) 24 hours pre-challenge at 20 mg/kg body weight, (ii) 4 hours post-challenge at 20 mg/kg, (iii) 4 hours post-challenge at 10 mg/kg, or (iv) 24 hours post-challenge at 20 mg/kg. As previously validated in this model (Sia S.F., *et al.* (2020) *Nature* 583(7818):834-838), the body weight and tissue pathology were used as metrics of therapeutic efficacy in this Example to assess treatment effectiveness, and virus titers in nasal swabs of the hamsters were also monitored to assess treatment effectiveness. In the no-treatment control group, SARS-CoV-2 was fast acting: hamsters started losing weight precipitously on day 1 post challenge, reached the lowest weight on day 6, and then started regaining weight (Fig. 9A). Nasal virus titers were high on day 1, remained high on day 5, and then declined (Fig. 14). Pathology analysis on tissues collected on day 10 revealed moderate hyperplasia in the bronchial tubes (i.e., bronchioloalveolar hyperplasia) (Fig. 9B), with little hyperplasia in the lungs. These data are consistent with previous reports showing that SARS-CoV-2 mainly infects the nasal mucosa and bronchial epithelial cells of this hamster model (Sia SF, *et al.* (2020) *Nature* 583(7818):834-838). In

contrast, hamsters that received *Nanosota-1C-Fc* 24-hours pre-challenge were protected from SARS-CoV-2 infections, as evidenced by no weight loss, no bronchioloalveolar hyperplasia, and significantly reduced nasal virus titers (Figs. 9A-9B, Fig. 14). If administered 4 hours post-challenge, *Nanosota-1C-Fc* also effectively protected hamsters from SARS-CoV-2 infections at either dosage (20 or 10 mg/kg), as evidenced by the favorable therapeutic metrics of significantly reduced weight loss and bronchioloalveolar hyperplasia, and also reduced nasal virus titers (Figs. 9A-9B, Fig. 14). Hamsters receiving *Nanosota-1C-Fc* 24 hours post-challenge showed a trend toward reduced weight loss, bronchioloalveolar hyperplasia, and nasal virus titers (Figs. 9A-9B, Fig. 14). Overall, *Nanosota-1C-Fc* effectively protects hamsters from SARS-CoV-2 infections both prophylactically and therapeutically.

To further examine the *in vivo* efficacy of *Nanosota-1C-Fc*, its therapeutic efficacy was evaluated in human ACE2-transgenic mice challenged with SARS-CoV-2 (at a titer of 5×10^3 PFU) via intranasal inoculation. Instead of monitoring the body weights of the mice through the viral infection and recovery process, the virus titers in the lungs were monitored at the peak of the viral infection. To this end, four experimental groups of mice (seven per group) received a single dose of *Nanosota-1C-Fc* via intraperitoneal injection: (i) 24 hours pre-challenge at 20 mg/kg body weight, (ii) 24 hours pre-challenge at 10 mg/kg body weight, (iii) 4 hours post-challenge at 20 mg/kg, and (iv) 4 hours post-challenge at 10 mg/kg. Five out of the seven mice from each group were euthanized on day 2 post-challenge, and the virus titers in their lungs were measured using a virus titer plaque assay. Compared to the untreated control group, the mice that received *Nanosota-1C-Fc* had much lower virus titers in the lungs (~1000 times lower in the pre-challenge groups and ~100 times lower in the post-challenge groups), Fig.22A. In addition to the above virus titer measurements, the remaining two mice in each group were euthanized on day 5 post-challenge for pathologic analysis of lung tissues. In the untreated control group, histological examination revealed extensive inflammatory cell infiltration, especially in the peribronchial region, alveolar edema, and proliferative alveolar epithelium. These results are consistent with previous reports on the SARS-CoV-2-induced lung pathology of these mice (J. Zheng et al., Nature, 2021 Jan;589(7843):603-607). The mice that received *Nanosota-1C-Fc* showed a near absence of lung pathology for both the pre- and post-challenge groups at 20 mg/kg body weight, and minor lung pathological changes for the pre-challenge group at 10 mg/kg body weight, Fig.22B. Pathological changes were still present in mice treated post-challenge with 10 mg/kg *Nanosota-1C-Fc*. Overall, *Nanosota-1C-Fc* effectively protected the mouse model from SARS-CoV-2 infection of their lungs.

***Nanosota-1C-Fc* is stable *in vitro* and *in vivo* with excellent bioavailability**

Finally the cost-effectiveness and pharmacokinetics of *Nanosota-1C-Fc* was analyzed. First, the production of *Nanosota-1C-Fc* was characterized by expressing it in bacteria (Fig. 10A). After being purified on protein A column and gel filtration column, the purity of the protein reached nearly 100%. Without any optimization, the typical yield of the protein reached 40 mg per liter of bacterial culture. Second, the *in vitro* stability of *Nanosota-1C-Fc* was investigated (Fig. 10B). *Nanosota-1C-Fc* was incubated at one of four temperatures (-80°C, 4°C, 25°C or 37°C) for a week and then its SARS-CoV-2 RBD-binding capability was measured using ELISA. Using -80°C as a baseline, *Nanosota-1C-Fc* retained most of its RBD-binding capability after being incubated at other temperatures for a week. Third, the *in vivo* stability of *Nanosota-1C-Fc* was measured (Fig. 10C). *Nanosota-1C-Fc* was injected via tail vein into mice. Mouse sera were obtained at different time points and measured for their SARS-CoV-2 RBD-binding affinity using ELISA. *Nanosota-1C-Fc* retained most of its RBD-binding capability after 10 days *in vivo*. In contrast, *Nanosota-1C* was stable for only several hours in mice (Fig. 15A, Fig. 15B). Last, the biodistribution of *Nanosota-1C-Fc* in mice was examined (Fig. 10D). *Nanosota-1C-Fc* was radiolabeled with zirconium-89 and injected systemically into mice. Tissues or organs were collected at different time points and the biodistribution of *Nanosota-1C-Fc* was quantified through the presence of radioactivity using a scintillation counter. After three days, *Nanosota-1C-Fc* remained at high levels in the blood, lung, heart, kidney, liver and spleen, all of which are targets for SARS-CoV-2 (Puelles VG, *et al.* (2020) *N Engl J Med* 383(6):590-592); it remained at low levels in the intestine, muscle and bones. In contrast, *Nanosota-1C* had reduced biodistribution (Fig. 15C). Overall, *Nanosota-1C-Fc* can be produced at high yield, is highly stable, and has an extended *in vivo* half-life and excellent tissue distribution.

Discussion

Nanobody drugs derived from camelid antibodies are promising therapeutics with advantages relative to conventional antibodies. Herein, a series of nanobody drugs, named *Nanosota-1*, are reported that specifically target SARS-CoV-2 RBD. These nanobody drugs were derived from a naïve camelid nanobody phage display library with an unparalleled large size. The drug contained two humanized mutations. It underwent two successful rounds of affinity maturation, yielding a product that bound to the RBD with high affinity and specificity. Importantly, it is believed that the *Nanosota-1* drugs are the first anti-SARS-CoV-2 nanobodies that have demonstrated therapeutic efficacy in an animal model; it is also believed that they are the first anti-SARS-CoV-2 nanobodies that have been characterized for their superior cost-effectiveness, *in vitro* and *in vivo* stabilities, and tissue biodistribution. These features of

Nanosota-1 drugs make them promising anti-COVID-19 therapeutics.

By using recombinant ACE2 as a comparison, a rigorous method to evaluate the anti-SARS-CoV-2 potency of *Nanosota-1* drugs was chosen. Many common criteria for evaluating drug potency depend on specific experimental settings. For example, target-binding affinity of drugs is affected by coating methods (Shang J, *et al.* (2020) *Nature* 581(7807):221-224) and virus-neutralizing capability of drugs is affected by virus titers, receptor and protease expression in target cells, and conformations of the spike proteins (Shang J, *et al.* (2020) *Proc Natl Acad Sci U S A* 117(21):11727-11734). By using recombinant ACE2 as a comparison, this method does not depend on specific experimental settings. It was shown that *Nanosota-1* drugs directly competed with cell-surface ACE2 for the same binding site on SARS-CoV-2 RBD. Compared with ACE2, *Nanosota-1C-Fc* and *Nanosota-1C* bound to the RBD ~3000 fold and ~3 fold more strongly, respectively, blocking ACE2 binding to the RBD. Furthermore, *Nanosota-1C-Fc* and *Nanosota-1C* inhibited SARS-CoV-2 pseudovirus entry ~100 fold and ~10 fold more effectively than ACE2, and they inhibited authentic SARS-CoV-2 infections ~200 fold and ~50 fold more effectively than ACE2. As a potent anti-SARS-2 inhibitory itself (Monteil V, *et al.* (2020) *Cell* 181(4):905-913.e907), recombinant ACE2 is currently undergoing clinical trials in Europe as an anti-COVID-19 drug. This study showed that both *Nanosota-1C-Fc* and *Nanosota-1C* are much more potent than recombinant ACE2 in inhibiting SARS-CoV-2. Compared with ACE2, the much higher anti-SARS-CoV-2 potency of *Nanosota-1C-Fc* was partly due to the small size of its antigen-binding domain and its ideal binding site on the RBD, allowing *Nanosota-1C-Fc* to access the RBD in both the open spike during viral infection and the closed spike during viral immune evasion. Thus, the *Nanosota-1* series are ideal RBD-targeting drug candidates that can inhibit SARS-CoV-2 viral particles regardless of whether the viral spike molecules are in an open or a closed conformation. Importantly, it was shown that the effectiveness of *Nanosota-1C-Fc* was not limited to *in vitro* experiments, but translated directly to *in vivo* experiments by demonstrating efficacy in animal models. Remarkably, a single dose of *Nanosota-1C-Fc* prevented SARS-CoV-2 infections in hamsters; it also treated SARS-CoV-2 infections in hamsters, especially when administered relatively early after infection. Because SARS-CoV-2 rapidly replicates in hamsters, the time points and dosages for drug administration in hamsters cannot be directly translated to humans. These parameters will need to be determined in future clinical trials. It is worth noting that the molecular weight of *Nanosota-1C-Fc* (78 kDa) is above the kidney clearance threshold (60 kDa) (S. Steeland, et al., *Drug discovery today* 21, 1076-1113, 2016), but still only half of conventional antibodies' molecular weight (150 kDa). Its ideal size contributed to ease of production, excellent *in vitro* thermostability, good *in vivo* stability,

and high tissue bioavailability. All of these features are important for the implementation of *Nanosota-1C-Fc* as a potential COVID-19 therapeutic.

Equally important to their anti-SARS-CoV-2 potency are the superior cost-effectiveness and excellent pharmacokinetics of *Nanosota-1* drugs. Whereas conventional antibody drugs are produced in mammalian cells (Gräslund S, *et al.* (2008) *Nature methods* 5(2):135-146), *Nanosota-1* drugs are produced in bacteria in this Example. Even without any production optimization, the yield of the *Nanosota-1C-Fc* reached 40 mg per liter bacterial culture. In industrial settings, the production scale can be dramatically increased with optimized media and bioreactor conditions, minimizing their production cost (Sivashanmugam A, *et al.* (2009)

Protein science : a publication of the Protein Society 18(5):936-948). The purification procedure is also streamlined: after protein A column and gel filtration column, the purity of *Nanosota-1C-Fc* reached nearly 100%. *Nanosota-1* drugs are highly stable *in vitro*, facilitating their storage and transportation. By adding an Fc tag, the molecular weight of *Nanosota-1C-Fc* was increased to about 75 kDa, allowing it to have an extended half-life in the body (i.e., >10 days).

Importantly, *Nanosota-1C-Fc* is distributed at high levels in the blood, lung, heart, liver and kidney, all major targets for SARS-CoV-2. Thus, *Nanosota-1C-Fc* can battle SARS-CoV-2 for extended periods in infected organs while having low off-target activities. Nanobodies have low toxicity and low immunogenicity (Muyldermans S (2013) *Annu Rev Biochem* 82:775-797; Steeland et al. (2016) *Drug discovery today* 21(7):1076-1113), as evidenced by the recent approval of a nanobody drug for treating a blood clotting disorder (Scully M, *et al.* (2019) *N Engl J Med* 380(4):335-346). Nevertheless, *Nanosota-1* drugs contain two humanization mutations that to further reduce their immunogenicity. All of these salient features make *Nanosota-1* drugs ideal for combating the COVID-19 pandemic.

How can *Nanosota-1* drugs be used to curtail the COVID-19 pandemic? First, as evidenced by both animal studies, *Nanosota-1C-Fc* can prevent SARS-CoV-2 infections. Because of its long half-life, a single injected dose of *Nanosota-1C-Fc* can potentially protect a person from SARS-CoV-2 infections for days or weeks, reducing the spread of COVID-19. Second, data from both animal studies herein showed that *Nanosota-1C-Fc* can be injected to treat SARS-CoV-2 infections, saving lives and alleviating symptoms in infected patients. Third, despite its short half-life in the blood, *Nanosota-1C* can be inhaled / used as an inhaler to treat infections in the respiratory tracts (Van Heeke G, *et al.* (2017) *Pharmacology & therapeutics* 169:47-56) or ingested as an oral drug to treat infections in the intestines (Vega CG, *et al.* (2013) *PLoS Pathog* 9(5):e1003334). The high yield, low cost, and long *in vitro* and *in vivo* half-life of *Nanosota-1* drugs make them a feasible treatment option for the world's vast population.

Clinical trials of the *Nanosota-1* drugs are urgently needed to evaluate their efficacy in ending the COVID-19 global pandemic.

Materials and Methods

5 *Cell lines, plasmids and virus*

HEK293T cells (American Type Culture Collection) were cultured in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% fetal bovine serum, 2 mM L-glutamine, 100 units/mL penicillin, and 100 µg/mL streptomycin (Life Technologies). ss320 *E. coli* (Lucigen), TG1 *E. coli* (Lucigen), SHuffle T7 *E. coli* (New England Biolabs) were grown in 10 TB medium or 2YT medium with 100 mg/L ampicillin. Vero E6 cells (American Type Culture Collection) were grown in Eagle's minimal essential medium (EMEM) supplemented with penicillin (100 units/ml), streptomycin (100 µg/ml), and 10% fetal bovine serum (FBS). HEK293T cells were authenticated by ATCC using STR profiling. HEK293T cells and Vero E6 cells were tested for mycoplasma contamination. No commonly misidentified cell lines were 15 used.

SARS-CoV-2 spike (GenBank accession number QHD43416.1) and ACE2 (GenBank accession number NM_021804) were described previously (Shang J, *et al.* (2020) *Nature* 581(7807):221-224). SARS-CoV-2 RBD (residues 319-535) was subcloned into Lenti-CMV vector (Vigene Biosciences) with an N-terminal tissue plasminogen activator (tPA) signal 20 peptide and a C-terminal human IgG4 Fc tag or His tag. The ACE2 ectodomain (residues 1–615) was constructed in the same way except that its own signal peptide was used. Nanosota-1A, 1B and 1C were each cloned into PADL22c vector (Lucigen) with a N-terminal PelB leader sequence and C-terminal His tag and HA tag. Nanosota-1C was cloned into pET42b vector (Novagen) with a C-terminal human IgG₁ Fc tag to produce Nanosota-1C-Fc.

25 SARS-CoV-2 (US_WA-1 isolate) from CDC (Atlanta) was used throughout the study. The titer of the viral stock was 1×10^7 TCID₅₀ (50% tissue culture infectious doses)/ml. All experiments involving infectious SARS-CoV-2 were conducted in an approved biosafety level 3 laboratory.

Construction of camelid nanobody phage display library

30 The camelid nanobody phage display library was constructed as previously described (Abbadly et al., (2011) *Veterinary immunology and immunopathology* 142(1-2):49-56; Olichon A & de Marco A (2012) Preparation of a naïve library of camelid single domain antibodies. *Methods in molecular biology* (Clifton, N.J.) 911:65-78). Briefly, total mRNA was isolated from B cells from the spleen, bone marrow and blood of over a dozen non-immunized llamas and

alpacas. cDNA was prepared from the mRNA. The cDNA was then used in nested PCR reactions to construct the DNA for the library. The first PCR reaction was to amplify the gene fragments encoding the variable domain of the nanobody. The second PCR reaction (PCR2) was used to add restriction sites (SFI-I), a PelB leader sequence, a His₆ tag, and a HA tag. The PCR2 product was digested with SFI-I (New England Biolabs) and then was ligated with SFI-I-digested PADL22c vector. The ligated product was transformed via electroporation into TG1 *E. coli* (Lucigen). Aliquots of cells were spread onto 2YT agar plates supplemented with ampicillin and glucose, incubated at 30°C overnight, and then scraped into 2YT media. After centrifugation, the cell pellet was suspended into 50% glycerol and stored at -80°C. The library size was 7.5×10^{10} . To display nanobodies on phages, aliquots of the TG1 *E. coli* bank were inoculated into 2YT media, grown to early logarithmic phase, and infected with M13K07 helper phage. *See also*, Ye, et al., (2020). The Development of a Novel Nanobody Therapeutic for SARS-CoV-2. *bioRxiv : the preprint server for biology*, 2020.11.17.386532, doi.org/10.1101/2020.11.17.386532, which is incorporated by reference herein for all purposes.

Camelids nanobody library screening

The above camelids nanobody phage library was used in the bio-panning as previously described (Hintz et al., (2019) *Bioconjugate chemistry* 30(5):1466-1476). Briefly, four rounds of panning were performed to obtain the SARS-CoV-2 RBD-targeting nanobodies with high RBD-binding affinity. The amounts of the RBD antigen used in coating the immune tubes in each round were 75 µg, 50 µg, 25 µg, and 10 µg, respectively. The retained phages were eluted using 1 ml 100 mM triethylamine and neutralized with 500 µl 1 M Tris-HCl pH 7.5. The eluted phages were amplified in TG1 *E. coli* and rescued with M13K07 helper phage. The eluted phages from round 4 were used to infect ss320 *E. coli*. Single colonies were picked into 2YT media and nanobody expressions were induced with 1 mM IPTG. The supernatants were subjected to ELISA for selection of strong binders. The strong binders were then expressed and purified and subjected to SARS-CoV-2 pseudovirus entry assay for selection of anti-SARS-CoV-2 efficacy. The strongest binder after initial screening was named *Nanosota-1A*.

Affinity maturation of *Nanosota-1A*

Affinity maturation of *Nanosota-1A* was performed as previously described (Hust M & Lim TS (2018) *Phage display : methods and protocols*.). Briefly, mutations were introduced into the whole gene of *Nanosota-1A* using error-prone PCR. Two rounds of error-prone PCR were performed using the GeneMorph II Random Mutagenesis Kit (Agilent Technologies). The PCR product was cloned into the PADL22c vector and transformed via electroporation into the TG1 *E. coli*. The library size was 6×10^8 . Three rounds of bio-panning were performed using 25 ng,

10 ng and 2 ng RBD-Fc, respectively. The strongest binder after affinity maturation was named *Nanosota-1B*. A second round of affinity maturation was performed in the same way as the first round, except that three rounds of bio-panning were performed using 10 ng, 2 ng and 0.5 ng RBD-Fc, respectively. The strongest binder after the second round of affinity maturation was
5 named *Nanosota-1C*.

Production of Nanosota-1 drugs

Nanosota-1A, *1B* and *1C* were each purified from the periplasm of ss320 *E. coli* after the cells were induced by 1 mM IPTG. The cells were collected and re-suspended in 15 ml TES buffer (0.2 M Tris pH 8, 0.5 mM EDTA, 0.5 M sucrose), shaken on ice for 1 hour and then
10 incubated with 40 ml TES buffer followed by shaking on ice for another hour. The protein in the supernatant was sequentially purified using a Ni-NTA column and a Superdex200 gel filtration column (GE Healthcare) as previously described (Shang J, *et al.* (2020) *Nature* 581(7807):221-224).

Nanosota-1C-Fc was purified from the cytoplasm of Shuffle T7 *E. coli*. The induction of
15 protein expression was the same as above. After induction, the cells were collected, re-suspended in PBS and disrupted using Branson Digital Sonifier (ThermoFisher). The protein in the supernatant was sequentially purified on protein A column and Superdex200 gel filtration column as previously described (Shang J, *et al.* (2020) *Nature* 581(7807):221-224).

Production of SARS-CoV-2 RBD and ACE2

20 HEK293T cells stably expressing SARS-CoV-2 RBD (containing a C-terminal His tag or Fc tag) or human ACE2 ectodomain (containing a C-terminal His tag) were made according to the E and F sections of the pLKO.1 Protocol from Addgene (addgene.org/protocols/plko/). The proteins were secreted to cell culture media, harvested, and purified on either Ni-NTA column (for His-tagged protein) or protein A column (for Fc-tagged protein) and then on Superdex200
25 gel filtration column as previously described (Shang J, *et al.* (2020) *Nature* 581(7807):221-224).

ELISA

ELISA was performed to detect the binding between SARS-CoV-2 RBD and *Nanosota-1* drugs (either purified recombinant drugs or drugs in the mouse serum) (Zhao G, *et al.* (2018) A Novel Nanobody Targeting Middle East Respiratory Syndrome Coronavirus (MERS-CoV) Receptor-Binding Domain Has Potent Cross-Neutralizing Activity and Protective Efficacy
30 against MERS-CoV. *J Virol* 92(18)). Briefly, ELISA plates were coated with recombinant SARS-CoV-2 RBD-His or RBD-Fc, and were then incubated sequentially with nanobody drugs, HRP-conjugated anti-llama antibody (1:5,000) (Sigma) or HRP-conjugated anti-human-Fc antibody (1:5,000) (Jackson ImmunoResearch). ELISA substrate (Invitrogen) was added to the

plates, and the reactions were stopped with 1N H₂SO₄. The absorbance at 450 nm (A₄₅₀) was measured using a Synergy LX Multi-Mode Reader (BioTek).

Determination of the structure of SARS-CoV-2 RBD complexed with Nanosota-1C

To prepare the RBD/*Nanosota-1C* complex for crystallization, the two proteins were mixed together in solution and purified using a Superdex200 gel filtration column (GE Healthcare). The complex was concentrated to 10 mg/ml in buffer 20 mM Tris pH 7.2 and 200 mM NaCl. Crystals were screened at High-Throughput Crystallization Screening Center (Hauptman-Woodward Medical Research Institute) as previously described (J. R. Luft et al., Journal of structural biology 142, 170-179 (2003)), and were grown in sitting drops at room temperature over wells containing 50 mM MnCl₂, 50 mM MES pH 6.0, and 20% (W/V) PEG 4000. Crystals were soaked briefly in 50 mM MnCl₂, 50 mM MES pH 6.0, 25% (W/V) PEG 4000 and 30% ethylene glycol before being flash-frozen in liquid nitrogen. X-ray diffraction data were collected at the Advanced Photon Source beamline 24-ID-E. The structure was determined by molecular replacement using the structures of SARS-CoV-2 RBD (PDB 6M0J) and another nanobody (PDB 6QX4) as the search templates. Structure data and refinement statistics are shown in Table A.

Table A. X-ray data collection and structure refinement statistics (SARS-CoV-2 RBD/*Nanosota-1C* complex)

Data collection	
Wavelength	0.979
Resolution range	45.48 - 3.19 (3.30 - 3.19)
Space group	P 43 21 2
Unit cell	60.849 60.849 410.701 90 90 90
Total reflections	64167 (5703)
Unique reflections	13607 (1308)
Multiplicity	4.7 (4.4)
Completeness (%)	96.82 (97.60)
Mean I/sigma(I)	8.41 (1.80)
Wilson B-factor	83.24
R-merge	0.145 (0.928)
R-meas	0.1638 (1.053)
R-pim	0.07385 (0.4858)
CC1/2	0.995 (0.861)
CC*	0.999 (0.962)
Refinement	

Reflections used in refinement	13567 (1301)
Reflections used for R-free	674 (62)
R-work	0.2483 (0.3521)
R-free	0.2959 (0.4153)
CC(work)	0.963 (0.819)
CC(free)	0.909 (0.615)
Number of non-hydrogen atoms	4890
macromolecules	4833
ligands	57
Protein residues	621
RMS(bonds)	0.002
RMS(angles)	0.45
Ramachandran favored (%)	93.11
Ramachandran allowed (%)	6.89
Ramachandran outliers (%)	0.00
Rotamer outliers (%)	3.23
Clashscore	5.25
Average B-factor	90.29
macromolecules	89.84
ligands	127.91

Statistics for the highest-resolution shell are shown in parentheses.

Surface plasmon resonance assay

Surface plasmon resonance assay using a Biacore S200 system (GE Healthcare) were carried out as previously described (Shang J, *et al.* (2020) *Nature* 581(7807):221-224). Briefly, SARS2-CoV-2 RBD-His was immobilized to a CM5 sensor chip (GE Healthcare). Serial dilutions of purified recombinant *Nanosota-1* drugs were injected at different concentrations: 10 nM, 20 nM, 40 nM, 80 nM, 160 nM, 320 nM for *Nanosota-1A*; 2.5 nM, 5 nM, 10 nM, 20 nM, 40 nM, 80 nM for *Nanosota-1B* and *Nanosota-1C*; 1.25 nM, 2.5 nM, 5 nM, 10 nM, 20 nM for *Nanosota-1C-Fc*. The resulting data were fit to a 1:1 binding model using Biacore Evaluation Software (GE Healthcare).

Protein pull-down assay

Protein pull-down assay was performed using Immunoprecipitation kit (Invitrogen) as previously described (Shang J, *et al.* (2020) *Nature* 581(7807):221-224). Briefly, 10 µl protein A beads were incubated with 1 µg SARS-CoV-2 RBD-Fc at room temperature for 1 hour. Then different amounts (7.04, 3.52, 1.76, 0.88, 0.44, 0.22, or 0 µg) of *Nanosota-1C* (with a C-terminal His tag) and 4 µg human ACE2 (with a C-terminal His tag) were added to the RBD-bound beads. After one-hour incubation at room temperature, the bound proteins were eluted using elution buffer (0.1 M glycine pH 2.7). The samples were then subjected to SDS-PAGE and analyzed through Western blot using an anti-His antibody.

To pull down ACE2-His and *Nanosota-1C-Fc* (containing a C-terminal Fc tag), 10 µl

streptavidin beads were incubated with 1 µg SARS-CoV-2 RBD-His (biotinylated using EZ-Link™ Sulfo-NHS-LC-Biotinylation Kit, Thermo Scientific) at room temperature for 1 hour. Then different amounts (17.80, 8.90, 4.45, 2.23, 1.11, 0.56, or 0 µg) of *Nanosota-1C-Fc* and 4 µg ACE2-His were added to the RBD-bound beads. After incubation at room temperature for 1 hour, the bound proteins were eluted using elution buffer (0.1 M glycine pH 2.7). The samples were then subjected to SDS-PAGE and analyzed through Western blot using an anti-His antibody (for detecting ACE2-His) or an anti-Fc antibody (for detecting *Nanosota-1C-Fc*).

Gel filtration chromatography assay

Gel filtration chromatography assay was performed on a Superdex200 column. 500 µg human ACE2, 109 µg *Nanosota-1C* and 121 µg SARS-CoV-2 RBD were incubated together at room temperature for 30 min. Then the mixture was subjected to gel filtration chromatography. Samples from each peak off the column were then subjected to SDS-PAGE and analyzed through Coomassie blue staining.

SARS-CoV-2 pseudovirus entry assay

The potency of *Nanosota-1* drugs in neutralizing SARS-CoV-2 pseudovirus entry was evaluated as previously described (Shang J, *et al.* (2020) *Nature* 581(7807):221-224; Shang J, *et al.* (2020) *Proc Natl Acad Sci U S A* 117(21):11727-11734). Briefly, HEK293T cells were co-transfected with a plasmid carrying an Env-defective, luciferase-expressing HIV-1 genome (pNL4-3.luc.R-E-) and pcDNA3.1(+) plasmid encoding SARS-CoV-2 spike protein.

Pseudoviruses were collected 72 hours after transfection, incubated with individual drugs at different concentrations 37°C for one hour, and then were used to enter HEK293T cells expressing human ACE2 (Shang J, *et al.* (2020) *Nature* 581(7807):221-224). After pseudoviruses and target cells were incubated together for 6 hours at 37°C, the medium was changed to fresh medium, followed by incubation of another 60 hours. Cells were then washed with PBS buffer and lysed. Aliquots of cell lysates were transferred to plates, followed by the addition of luciferase substrate. Relative light units (RLUs) were measured using an EnSpire plate reader (PerkinElmer). The efficacy of each inhibitor was expressed as the concentration capable of neutralizing 50% or 90% of the entry efficiency (i.e., ND₅₀ or ND₉₀, respectively).

SARS-CoV-2 plaque reduction neutralization test

The potency of *Nanosota-1* in neutralizing live SARS-CoV-2 was evaluated using a SARS-CoV-2 plaque reduction neutralization test (PRNT) assay. Specifically, individual drug candidate was serially diluted in DMEM and mixed with SARS-CoV-2 (at a titer of 80 plaque-forming unit (PFU)) at 37°C for 1 hour. The mixtures were then added into Vero E6 cells at 37°C for an additional 45 minutes. After removing the culture medium, cells were overlaid with

0.6% agarose and cultured for 3 days. Plaques were visualized by 0.1% crystal violet staining. The efficacy of each drug candidate was calculated and expressed as the concentration capable of reducing the number of virus plaques by 50% (i.e., ND₅₀) compared to control serum-exposed virus.

5 *SARS-CoV-2 micro-neutralization assay*

The potency of *Nanosota-1* drugs in neutralizing authentic SARS-CoV-2 infections was evaluated using a SARS-CoV-2 micro-neutralization assay. Specifically, confluent Vero E6 cells grown in 96-well microtiter plates were pre-treated with serially diluted individual drug for one hour before infection with 100 infectious SARS-CoV-2 particles in 100 µl EMEM
10 supplemented with 2% FBS. Vero E6 cells treated with parallelly diluted dimethyl sulfoxide (DMSO) with or without virus were included as positive and negative controls, respectively. After cultivation at 37°C for 3 days, individual wells were observed under the microcopy for the status of virus-induced formation of cytopathic effect (CPE). The efficacy of individual drugs was calculated and expressed as the lowest concentration capable of completely preventing
15 virus-induced CPE in 100% of the wells (i.e., ND₁₀₀) or 50% of the wells (i.e., ND₅₀).

SARS-CoV-2 challenge of hamsters

Equal sex hamsters (n=30) were obtained from Envigo (IN), randomly assigned into different groups (with gender balance), and challenged via intranasal inoculation with 1x10⁶ Median Tissue Culture Infectious Dose (TCID₅₀) of SARS-CoV-2 in 100 µL (50 µL per nare) of
20 DMEM. Four groups of hamsters (n=6 each) were treated with *Nanosota-1C-Fc* via intraperitoneal injection at one of the following time points and dosages: (1) 24 hours pre-challenge at 20 mg/kg body weight of hamsters; (2) 4 hours post-challenge at 20 mg/kg body weight of hamsters; (3) 4 hours post-challenge at 10 mg/kg body weight of hamsters; (4) 24 hours post-challenge at 20 mg/kg body weight of hamsters. Hamsters in the control (negative)
25 group were administered PBS buffer 24 hours pre-challenge. Body weights were collected daily beginning prior to challenge. Nasal swabs were collected prior to challenge and additionally 1 day, 2 days, 3 days, 5 days and 10 days post-challenge for quantitative real-time RT-PCR (nasal swabs collected on day 2 and day 3 were lost due to Hurricane Laura). Hamsters were humanely euthanized 10 days post-challenge via overexposure to CO₂. The lungs and bronchial tubes were
30 collected and fixed in formalin for histopathological analysis. At a sample size of 6 animals per group, G*Power analysis indicates that an effect size of 1.6 is detectable with a power of .80 (alpha = .05 one-tailed). This experiment was performed in accordance with the guidelines set by the Institutional Animal Care and Use Committee at the University of Texas Medical Branch (UTMB).

SARS-CoV-2 challenge of human ACE2-transgenic mice

Human ACE2-transgenic mice (K18-hACE2-transgenic mice) (J. Zheng et al., Nature 589, 603-607 (2021) and P. B. McCray et al., Journal of Virology 81, 813-821 (2007)) (n=35; males and females; 7-8 months old) were obtained from the Jackson Laboratories. Mice were challenged via intranasal inoculation with SARS-CoV-2 (5×10^3 PFU) in a volume of 50 μ l DMEM. Sample size was constrained by the availability of resources. Five groups of mice (n=7 in each group) were treated with *Nanosota-1C-Fc* via intraperitoneal injection at one of the following time points and dosages: (1) 24 hours pre-challenge at 20 mg/kg body weight; (2) 24 hours pre-challenge at 10 mg/kg body weight; (3) 4 hours post-challenge at 20 mg/kg body weight; (4) 4 hours post-challenge at 10 mg/kg body weight. Mice in the negative control group were administered PBS buffer 24 hours pre-challenge.

Viral titers in the lungs of mice were measured by a plaque assay. To this end, 5 mice from each group were euthanized on day 2 post-challenge. Lung homogenate supernatants were collected and then serially diluted in DMEM. 12-well plates of VeroE6 cells were inoculated and then incubated at 37 °C in 5% CO₂ for 1 hour and gently rocked every 15 minutes. After removing the inocula, the plates were overlaid with 0.6% agarose containing 2% FBS. After 3 days, overlays were removed, and plaques were visualized via staining with 0.1% crystal violet. Viral titers were quantified as PFU per ml tissue. At a sample size of 5 animals per group, G*Power analysis indicates that we can detect an effect size of 1.72 with a power of .80 (alpha = .05 one-tailed).

Histological examination of lungs was performed. For this purpose, the remaining 2 mice from each group were euthanized on day 5 post-challenge and then perfused transcardially with PBS. Mouse lungs were fixed in formalin. Sections (approximately 4 μ m each) were stained with hematoxylin and eosin.

Half-life of Nanosota-1 drugs in mice

Male C57BL/6 mice (3 to 4 weeks old) (Envigo) were intravenously injected (tail-vein) with *Nanosota-1C* or *Nanosota-1C-Fc* (100 μ g in 100 μ l PBS buffer). At varying time points, mice were euthanized and whole blood was collected. Then sera were prepared through centrifugation of the whole blood at 1500xg for 10 min. The sera were then subjected to ELISA for evaluation of their SARS-CoV-2 RBD-binding capability.

Biodistribution of Nanosota-1 drugs in mice

To evaluate the *in vivo* biodistribution of *Nanosota-1C-Fc* and *Nanosota-1C*, the nanobodies were labeled with Zirconium-89 [⁸⁹Zr] and injected into mice. Briefly, the nanobodies were first conjugated to the bifunctional chelator p-SCN-Bn-Deferoxamine (DFO,

Macrocyclic) as previously described (Zeglis BM & Lewis JS (2015) The bioconjugation and radiosynthesis of ^{89}Zr -DFO-labeled antibodies. *Journal of visualized experiments : JoVE* (96)), and [^{89}Zr] (University of Wisconsin Medical Physics Department) was then conjugated as previously described Hintz HM, *et al.* (2020) Imaging Fibroblast Activation Protein Alpha Improves Diagnosis of Metastatic Prostate Cancer with Positron Emission Tomography. *Clinical cancer research : an official journal of the American Association for Cancer Research*). [^{89}Zr]-labeled nanobodies (1.05 MBq, 1-2 μg nanobody, 100 μl PBS) were intravenously injected (tail-vein). Mice were euthanized at different time points. Organs were collected and counted on an automatic gamma-counter (Hidex). The total number of counts per minute (cpm) for each organ or tissue was compared with a standard sample of known activity and mass. Count data were corrected to both background and decay. The percent injected dose per gram (%ID/g) was calculated by normalization to the total amount of activity injected into each mouse.

Table B. Binding affinities between Nanosota-1 drugs and SARS-CoV-2 RBD as measured using surface plasmon resonance. The previously determined binding affinity between human ACE2 and RBD is shown as a comparison (Shang J, *et al.* (2020) *Nature* 581(7807):221-224).

	K_d with SARS-CoV-2 RBD (M)	k_{off} (s^{-1})	k_{on} ($M^{-1}s^{-1}$)
Nanosota-1A (before affinity maturation)	2.28×10^{-7}	9.35×10^{-3}	4.10×10^4
Nanosota-1B (after 1 st round of affinity maturation)	6.08×10^{-8}	7.19×10^{-3}	1.18×10^5
Nanosota-1C (after 2 nd round of affinity maturation)	1.42×10^{-8}	2.96×10^{-3}	2.09×10^5
Nanosota-1C-Fc (after 2 nd round of affinity maturation; containing a C-terminal human Fc tag)	1.57×10^{-11}	9.68×10^{-5}	6.15×10^6
ACE2	4.42×10^{-8}	7.75×10^{-3}	1.75×10^5

Table 1.

SEQ ID NO:	Sequences	Comment
1	MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYR QAPGQGLEWVTSIYSDGSTVYADSVKGRFTVSRDNPVKSTV SLQMNSLKPEDTGVYHGMAGSKSGQELDHWGQGTQVTVSS	Nanosota-1A VHH sequence

2	GFTFKNAD	Nanosota-1A: CDR-H1
3	IYSDGST	Nanosota-1A: CDR-H2
4	MAGSKSGQELDH	Nanosota-1A: CDR-H3
5	MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYR QAPGQGLEWVTSIYSDGSTVYADSVKGRFTVSRDNPVKSTV SLQMNSLKPEDTGVYHGMAGSKSGQELDHWGQGTQVTVSS GPGGQHHHHHHGAYPYDVPDYAS	Nanosota-1A VHH sequence with linker (bold) and His ₆ and HA tags (italics)
6	MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYR QAPGQGLEWVTSIYSDGSTVYADSVKGRFTVSRDNPVKSTV SLQMNSLKPEDTGVYHGMAGSKSGQELDHWGQGTQVTVSS EPKSCDKHTCPPCPAPPELLGGPSVFLFPPKPKDTLMISR TPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQ YNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT ISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPS DIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKS RWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK	Nanosota-1A VHH sequence with IgG1 Fc domain sequence
29	ATGGCTCAGGTGCAGCTCGTGGAGTCTGGGGGAGGCTTGG TGCAGCCTGGGGGGTCTCTGAGACTCTCCTGTGCAGCCTC TGGCTTCACCTTCAAAAACGCGGACATGAACTGGTATCGC CAGGCACCAGGTGAGGGGCTCGAGTGGGTGACTAGTATTT <u>ATAGTGACGGTAGTACAGTCTATGCAGACTCCGTGAAGGG</u> <u>CCGATTACCGTCTCCAGAGACAACCCCAAGAGCACGGTG</u> <u>TCTCTGCAAATGAACAGCCTGAAACCTGAGGACACAGGTG</u> <u>TCTATCATGGTATGGCAGGTTCAAAGTCGGGCCAAGAGTT</u> <u>GGATCATTGGGGCCAGGGGACCCAGGTACCGTCTCCAGC</u>	DNA sequence for Nanosota-1A VHH, CDR (CDR-H1, CDR-H2, CDR- H3) sequences are underlined
23	ATGGCTCAGGTGCAGCTCGTGGAGTCTGGGGGAGGCTTGG TGCAGCCTGGGGGGTCTCTGAGACTCTCCTGTGCAGCCTC TGGCTTCACCTTCAAAAACGCGGACATGAACTGGTATCGC CAGGCACCAGGTGAGGGGCTCGAGTGGGTGACTAGTATTT <u>ATAGTGACGGTAGTACAGTCTATGCAGACTCCGTGAAGGG</u> <u>CCGATTACCGTCTCCAGAGACAACCCCAAGAGCACGGTG</u> <u>TCTCTGCAAATGAACAGCCTGAAACCTGAGGACACAGGTG</u> <u>TCTATCATGGTATGGCAGGTTCAAAGTCGGGCCAAGAGTT</u> <u>GGATCATTGGGGCCAGGGGACCCAGGTACCGTCTCCAGC</u> GGCCCGGAGGCCAA <i>CACCATCACCACCATCATGGCGCAT</i> <i>ATCCGTATGATGTGCCGGACTATGCTTCT</i>	DNA sequence for Nanosota-1A VHH, CDR (CDR-H1, CDR-H2, CDR- H3) sequences are underlined, with linker (bold) and, His ₆ and HA tags (italics).
24	ATGGCTCAGGTGCAGCTCGTGGAGTCTGGGGGAGGCTTGG TGCAGCCTGGGGGGTCTCTGAGACTCTCCTGTGCAGCCTC TGGCTTCACCTTCAAAAACGCGGACATGAACTGGTATCGC CAGGCACCAGGTGAGGGGCTCGAGTGGGTGACTAGTATTT ATAGTGACGGTAGTACAGTCTATGCAGACTCCGTGAAGGG CCGATTACCGTCTCCAGAGACAACCCCAAGAGCACGGTG TCTCTGCAAATGAACAGCCTGAAACCTGAGGACACAGGTG TCTATCATGGTATGGCAGGTTCAAAGTCGGGCCAAGAGTT GGATCATTGGGGCCAGGGGACCCAGGTACCGTCTCCAGC GAGCCCAAATCTTGTGACAAA <i>ACTCACACATGCCACCGT</i> <i>GCCCAGCACCTGAACTCCTGGGGGGACCGTCAGTCTTCCT</i> <i>CTTCCCCCAAACCCCAAGGACACCCTCATGATCTCCCGG</i> <i>ACCCCTGAGGTACATGCGTGGTGGTGGACGTGAGCCACG</i>	DNA sequence for Nanosota-1A VHH sequence with IgG1 Fc domain sequence (bold)

	AAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGT GGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAG TACAACAGCACGTACCGTGTGGTCAGCGTCCTCACCGTCC TGCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCAA GGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACC ATCTCCAAAGCCAAAGGGCAGCCCCGAGAACCACAGGTGT ACACCCTGCCCCCATCCCGGGAGGAGATGACCAAGAACCA GGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGC GACATCGCCGTGGAGTGGGAGAGCAATGGGCAGCCGGAGA ACAAC TACAAGACCACGCC TCCCGTGCTGGACTCCGACGG CTCCTTCTTCTCTACAGCAAGCTCACCGTGGACAAGAGC AGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGC ATGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTC CCTGTCTCCGGGTAAA	
7	MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYR QAPGQGLEWVTSIYSDGSTVYADSVKGRFTVSRDNPSTV SLQMNSLKPEDTGVY <u>Y</u> CMAGSKSG <u>H</u> ELDHWGQGTQVTVSS	Nanosota-1B VHH sequence, which was generated after the 1 st round of affinity maturation. Mutated residues compared to Nanosota-1A are shown in bold/underline.
2	GFTFKNAD	Nanosota-1B: CDR-H1
3	IYSDGST	Nanosota-1B: CDR-H2
8	MAGSKSGHELDH	Nanosota-1B: CDR-H3
9	MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYR QAPGQGLEWVTSIYSDGSTVYADSVKGRFTVSRDNPSTV SLQMNSLKPEDTGVYYCMAGSKSGHELDHWGQGTQVTVSS GP GGQ <i>HHHHHHGAYPYDVPDYAS</i>	Nanosota-1B VHH sequence with linker (bold)and His ₆ and HA tags (italics)
10	MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYR QAPGQGLEWVTSIYSDGSTVYADSVKGRFTVSRDNPSTV SLQMNSLKPEDTGVYYCMAGSKSGHELDHWGQGTQVTVSS EPKSCDKHTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISR TPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQ YNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT ISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPS DIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKS RWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK	Nanosota-1B VHH sequence with IgG1 Fc domain sequence
30	ATGGCTCAGGTGCAGCTCGTGGAGTCTGGGGGAGGCTTGG TGCAGCCTGGGGGGTCTCTGAGACTCTCCTGTGCAGCCTC <u>TGGCTTCACCTTCAAAAACGCGGACATGAACTGGTATCGC</u> <u>CAGGCACCAGGTCAGGGGCTCGAGTGGGTGACTAGTATTT</u> <u>ATAGTGACGGTAGTACAGTCTATGCAGACTCCGTGAAGGG</u> <u>CCGATTACCGTCTCCAGAGACAACCCCAAGAGCACGGTG</u> <u>TCTCTGCAAATGAACAGCCTGAAACCTGAGGACACAGGTG</u>	DNA sequence for Nanosota-1B VHH, CDR (CDR-H1, CDR-H2, CDR- H3) sequences are underlined

	TCTATTATTGTATGGCAGGTTCAAAGTCGGGCCATGAGTT GGATCATTGGGGCCAGGGGACCCAGGTCACCGTCTCCAGC	
25	ATGGCTCAGGTGCAGCTCGTGGAGTCTGGGGGAGGCTTGG TGCAGCCTGGGGGGTCTCTGAGACTCTCCTGTGCAGCCTC TGGCTTCACCTTCAAAAACGCGGACATGAACTGGTATCGC CAGGCACCAGGTCAGGGGCTCGAGTGGGTGACTAGTATTT ATAGTGACGGTAGTACAGTCTATGCAGACTCCGTGAAGGG CCGATTACCGTCTCCAGAGACAACCCCAAGAGCACGGTG TCTCTGCAAATGAACAGCCTGAAACCTGAGGACACAGGTG TCTATTATTGTATGGCAGGTTCAAAGTCGGGCCATGAGTT GGATCATTGGGGCCAGGGGACCCAGGTCACCGTCTCCAGC GGCCCGGGAGGCCAA CACCATCACCAACCATCATGGCGCAT ATCCGTATGATGTGCCGGACTATGCTTCT	DNA sequence for Nanosota-1B VHH, CDR (CDR-H1, CDR-H2, CDR-H3) sequences are underlined, with linker (bold)and, His ₆ and HA tags (italics).
26	ATGGCTCAGGTGCAGCTCGTGGAGTCTGGGGGAGGCTTGG TGCAGCCTGGGGGGTCTCTGAGACTCTCCTGTGCAGCCTC TGGCTTCACCTTCAAAAACGCGGACATGAACTGGTATCGC CAGGCACCAGGTCAGGGGCTCGAGTGGGTGACTAGTATTT ATAGTGACGGTAGTACAGTCTATGCAGACTCCGTGAAGGG CCGATTACCGTCTCCAGAGACAACCCCAAGAGCACGGTG TCTCTGCAAATGAACAGCCTGAAACCTGAGGACACAGGTG TCTATTATTGTATGGCAGGTTCAAAGTCGGGCCATGAGTT GGATCATTGGGGCCAGGGGACCCAGGTCACCGTCTCCAGC GAGCCCAAATCTTGTGACAAACTCACACATGCCACCGT GCCCAGCACCTGAACTCCTGGGGGGACCGTCAGTCTTCCT CTTCCCCCAAACCCAAAGGACACCCTCATGATCTCCCGG ACCCCTGAGGTACATGCGTGGTGGTGGACGTGAGCCACG AAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGT GGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAG TACAACAGCACGTACCGTGTGGTCAGCGTCCTCACCGTCC TGCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCAA GGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACC ATCTCAAAGCCAAAGGGCAGCCCCGAGAACCACAGGTGT ACACCCTGCCCCCATCCCGGGAGGAGATGACCAAGAACCA GGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGC GACATCGCCGTGGAGTGGGAGAGCAATGGGCAGCCGGAGA ACAACACAAGACCACGCCCTCCCGTGCTGGACTCCGACGG CTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGC AGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGC ATGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTC CCTGTCTCCGGGTAAA	DNA sequence for Nanosota-1B VHH sequence with IgG1 Fc domain sequence (bold)
11	MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYR QVPGQGLEWVTSIYSDG R TVYADSVKGRFTVSRDNPVKSTV SLQMNSLKPEDTGVYYCMAGSKSGHELDHWGQGTQVTVSS	Nanosota-1C VHH sequence, which was generated after the 2 nd round of affinity maturation. Mutated residues compared to Nanosota-1B are shown in bold/underline.

2	GFTFKNAD	Nanosota-1C: CDR-H1
12	IYSDGRT	Nanosota-1C: CDR-H2
8	MAGSKSGHELDH	Nanosota-1C: CDR-H3
13	MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYR QVPGQGLEWVTSIYSDGRTVYADSVKGRFTVSRDNPSTV SLQMNSLKPEDTGVYYCMAGSKSGHELDHWGQGTQVTVSS GPGGQHHHHHHGAYPYDVPDYAS	Nanosota-1C VHH sequence with linker (bold) and His ₆ and HA tags (italics)
14	MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYR QVPGQGLEWVTSIYSDGRTVYADSVKGRFTVSRDNPSTV SLQMNSLKPEDTGVYYCMAGSKSGHELDHWGQGTQVTVSS EPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISR TPEVTCVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQ YNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT ISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPS DIAVEWESNGQPENNYKTTPVLDSDGSFFLYSKLTVDKS RWQQGNVFCFSVMHEALHNHYTQKSLSLSPGK	Nanosota-1C VHH sequence with IgG1 Fc domain sequence
31	ATGGCTCAGGTGCAGCTCGTGGAGTCTGGGGGAGGCTTGG TGCAGCCTGGGGGGTCTCTGAGACTCTCCTGTGCAGCCTC TGGCTTCACCTTCAAAAACGCGGACATGAACTGGTATCGC CAGGTACCAGGTCAAGGGCTCGAGTGGGTGACTAGTATTT <u>ATAGTGACGGTAGAACAGTCTATGCAGACTCCGTGAAGGG</u> <u>CCGATTACCGTCTCCAGAGACAACCCCAAGAGCACGGTG</u> <u>TCTCTGCAAATGAACAGCCTGAAACCTGAGGACACAGGTG</u> <u>TCTATTATTGTATGGCAGGTTCAAAGTCGGGCCATGAGTT</u> <u>GGATCATTGGGGCCAGGGGACCCAGGTCACCGTCTCCAGC</u>	DNA sequence for Nanosota-1C VHH, CDR (CDR-H1, CDR-H2, CDR-H3) sequences are underlined
27	ATGGCTCAGGTGCAGCTCGTGGAGTCTGGGGGAGGCTTGG TGCAGCCTGGGGGGTCTCTGAGACTCTCCTGTGCAGCCTC TGGCTTCACCTTCAAAAACGCGGACATGAACTGGTATCGC CAGGTACCAGGTCAAGGGCTCGAGTGGGTGACTAGTATTT <u>ATAGTGACGGTAGAACAGTCTATGCAGACTCCGTGAAGGG</u> <u>CCGATTACCGTCTCCAGAGACAACCCCAAGAGCACGGTG</u> <u>TCTCTGCAAATGAACAGCCTGAAACCTGAGGACACAGGTG</u> <u>TCTATTATTGTATGGCAGGTTCAAAGTCGGGCCATGAGTT</u> <u>GGATCATTGGGGCCAGGGGACCCAGGTCACCGTCTCCAGC</u> GGCCCGGGAGGCCAA <i>CACCATCACCAACCATCATGGCGCAT</i> <i>ATCCGTATGATGTGCCGGACTATGCTTCT</i>	DNA sequence for Nanosota-1C VHH, CDR (CDR-H1, CDR-H2, CDR-H3) sequences are underlined, with linker (bold) and, His ₆ and HA tags (italics).
28	ATGGCTCAGGTGCAGCTCGTGGAGTCTGGGGGAGGCTTGG TGCAGCCTGGGGGGTCTCTGAGACTCTCCTGTGCAGCCTC TGGCTTCACCTTCAAAAACGCGGACATGAACTGGTATCGC CAGGTACCAGGTCAAGGGCTCGAGTGGGTGACTAGTATTT ATAGTGACGGTAGAACAGTCTATGCAGACTCCGTGAAGGG CCGATTACCGTCTCCAGAGACAACCCCAAGAGCACGGTG TCTCTGCAAATGAACAGCCTGAAACCTGAGGACACAGGTG TCTATTATTGTATGGCAGGTTCAAAGTCGGGCCATGAGTT GGATCATTGGGGCCAGGGGACCCAGGTCACCGTCTCCAGC GAGCCCAAATCTTGTGACAAACTCACACATGCCACCGT GCCCAGCACCTGAACTCCTGGGGGGACCGTCAGTCTTCCT CTTCCCCCAAACCCCAAGGACACCCTCATGATCTCCCGG ACCCCTGAGGTACATGCGTGGTGGTGGACGTGAGCCACG AAGACCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGT	DNA sequence for Nanosota-1C VHH sequence with IgG1 Fc domain sequence (bold)

	GGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAG TACAACAGCACGTACCGTGTGGTCAGCGTCCTCACCGTCC TGCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCAA GGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACC ATCTCCAAAGCCAAAGGGCAGCCCCGAGAACCACAGGTGT ACACCCTGCCCCCATCCCGGGAGGAGATGACCAAGAACCA GGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGC GACATCGCCGTGGAGTGGGAGAGCAATGGGCAGCCGGAGA ACAATAACAAGACCACGCCTCCCGTGCTGGACTCCGACGG CTCCTTCTTCTCTACAGCAAGCTCACCGTGGACAAGAGC AGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGC ATGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTC CCTGTCTCCGGGTAAA	
15	GPGGQ	Linker sequence
16	HHHHHH	His6 tag sequence
17	GAYPYDVPDYAS	Embodiment of sequence comprising HA tag
22	YPYDVPDYA	Embodiment of HA tag sequence
18	HHHHHHGAYPYDVPDYAS	His6 and HA tag sequence
19	EPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISR TPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQ YNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT ISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPS DIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKS RWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK	IgG1 Fc sequence
20	IYSDG (S/R) T	Embodiment of CDR-H2 consensus sequence
21	MAGSKSG (Q/H) ELDH	Embodiment of CDR-H3 consensus sequence

All publications, patents and patent applications cited herein are incorporated herein by reference. While in the foregoing specification this invention has been described in relation to certain embodiments thereof, and many details have been set forth for purposes of illustration, it will be apparent to those skilled in the art that the invention is susceptible to additional embodiments and that certain of the details described herein may be varied considerably without departing from the basic principles of the invention.

The use of the terms "a" and "an" and "the" and "or" and similar referents in the context of describing the invention are to be construed to cover both the singular and the plural, unless otherwise indicated herein or clearly contradicted by context. Thus, for example, reference to "a subject polypeptide" includes a plurality of such polypeptides and reference to "the agent"

includes reference to one or more agents and equivalents thereof known to those skilled in the art, and so forth.

The terms “comprising,” “having,” “including,” and “containing” are to be construed as open-ended terms (*i.e.*, meaning “including, but not limited to”) unless otherwise noted.

5 Recitation of ranges of values herein are merely intended to serve as a shorthand method of referring individually to each separate value falling within the range, unless otherwise indicated herein, and each separate value is incorporated into the specification as if it were individually recited herein. All methods described herein can be performed in any suitable order unless otherwise indicated herein or otherwise clearly contradicted by context. The use of any and all
10 examples, or exemplary language (*e.g.*, “such as”) provided herein, is intended merely to better illuminate the invention and does not pose a limitation on the scope of the invention unless otherwise claimed. No language in the specification should be construed as indicating any non-claimed element as essential to the practice of the invention.

Embodiments of this invention are described herein, including the best mode known to
15 the inventor for carrying out the invention. Variations of those embodiments may become apparent to those of ordinary skill in the art upon reading the foregoing description. The inventor expects skilled artisans to employ such variations as appropriate, and the inventor intends for the invention to be practiced otherwise than as specifically described herein. Accordingly, this invention includes all modifications and equivalents of the subject matter
20 recited in the claims appended hereto as permitted by applicable law. Moreover, any combination of the above-described elements in all possible variations thereof is encompassed by the invention unless otherwise indicated herein or otherwise clearly contradicted by context.

With respect to ranges of values, the invention encompasses each intervening value between the upper and lower limits of the range to at least a tenth of the lower limit's unit, unless
25 the context clearly indicates otherwise. Further, the invention encompasses any other stated intervening values. Moreover, the invention also encompasses ranges excluding either or both of the upper and lower limits of the range, unless specifically excluded from the stated range.

Further, all numbers expressing quantities of ingredients, reaction conditions, % purity, polypeptide and polynucleotide lengths, and so forth, used in the specification and claims, are
30 modified by the term “about,” unless otherwise indicated. Accordingly, the numerical parameters set forth in the specification and claims are approximations that may vary depending upon the desired properties of the present invention. At the very least, and not as an attempt to limit the application of the doctrine of equivalents to the scope of the claims, each numerical parameter should at least be construed in light of the number of reported significant digits,

applying ordinary rounding techniques. Nonetheless, the numerical values set forth in the specific examples are reported as precisely as possible. Any numerical value, however, inherently contains certain errors from the standard deviation of its experimental measurement.

5 Unless defined otherwise, the meanings of all technical and scientific terms used herein are those commonly understood by one of skill in the art to which this invention belongs. One of skill in the art will also appreciate that any methods and materials similar or equivalent to those described herein can also be used to practice or test the invention. Further, all publications mentioned herein are incorporated by reference in their entireties.

CLAIMS

What is claimed is:

1. An isolated anti-SARS-CoV-2 nanobody comprising one or more complementarity determining regions (CDRs) selected from the group consisting of:
 - (a) a CDR1 comprising an amino acid sequence having at least 80% sequence identity to an amino acid sequence of GFTFKNAD (SEQ ID NO:2);
 - (b) a CDR2 comprising an amino acid sequence having at least 80% sequence identity to an amino acid sequence of IYSDG(S/R)T (SEQ ID NO:20); and
 - (c) a CDR3 comprising an amino acid sequence having at least 80% sequence identity to an amino acid sequence of MAGSKSG(Q/H)ELDH (SEQ ID NO:21).
2. The isolated anti-SARS-CoV-2 nanobody of claim 1, comprising one or more CDRs selected from the group consisting of:
 - (a) a CDR1 comprising an amino acid sequence having at least 90% sequence identity to an amino acid sequence of GFTFKNAD (SEQ ID NO:2);
 - (b) a CDR2 comprising an amino acid sequence having at least 90% sequence identity to an amino acid sequence of IYSDG(S/R)T (SEQ ID NO:20); and
 - (c) a CDR3 comprising an amino acid sequence having at least 90% sequence identity to an amino acid sequence of MAGSKSG(Q/H)ELDH (SEQ ID NO:21).
3. The isolated anti-SARS-CoV-2 nanobody of claim 1, comprising:
 - (a) a CDR1 comprising the amino acid sequence of GFTFKNAD (SEQ ID NO:2);
 - (b) a CDR2 comprising the amino acid sequence of IYSDG(S/R)T (SEQ ID NO:20);and
 - (c) a CDR3 comprising the amino acid sequence of MAGSKSG(Q/H)ELDH (SEQ ID NO:21).
4. The isolated anti-SARS-CoV-2 nanobody of claim 3, comprising:
 - (a) a CDR1 comprising the amino acid sequence of GFTFKNAD (SEQ ID NO:2);
 - (b) a CDR2 comprising the amino acid sequence of IYSDGRT (SEQ ID NO:12); and
 - (c) a CDR3 comprising the amino acid sequence of MAGSKSGHELDH (SEQ ID NO:8).

5. The isolated anti-SARS-CoV-2 nanobody of claim 3, comprising:
 - (a) a CDR1 comprising the amino acid sequence of GFTFKNAD (SEQ ID NO:2);
 - (b) a CDR2 comprising the amino acid sequence of IYSDGST (SEQ ID NO:3); and
 - (c) a CDR3 comprising the amino acid sequence of MAGSKSGQELDH (SEQ ID NO:4).

6. The isolated anti-SARS-CoV-2 nanobody of claim 3, comprising:
 - (a) a CDR1 comprising the amino acid sequence of GFTFKNAD (SEQ ID NO:2);
 - (b) a CDR2 comprising the amino acid sequence of IYSDGST (SEQ ID NO:3); and
 - (c) a CDR3 comprising the amino acid sequence of MAGSKSGHELDH (SEQ ID NO:8).

7. The isolated anti-SARS-CoV-2 nanobody of claim 1, comprising an amino acid sequence that has at least 80% sequence identity to any one of:
 - (a) MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYRQVPGQGL
EWVTSIYSDGRTVYADSVKGRFTVSRDNPSTVSLQMNSLKPEDTGVYYCMAGSKSG
HELDHWGQGTQVTVSS (SEQ ID NO:11);
 - (b) MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYRQAPGQGL
EWVTSIYSDGSTVYADSVKGRFTVSRDNPSTVSLQMNSLKPEDTGVYHGMAGSKSGQ
ELDHWGQGTQVTVSS (SEQ ID NO:1); and
 - (c) MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYRQAPGQGL
EWVTSIYSDGSTVYADSVKGRFTVSRDNPSTVSLQMNSLKPEDTGVYYCMAGSKSGH
ELDHWGQGTQVTVSS (SEQ ID NO:7).

8. The isolated anti-SARS-CoV-2 nanobody of claim 1, comprising an amino acid sequence that has at least 90% sequence identity to any one of:
 - (a) MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYRQVPGQGL
EWVTSIYSDGRTVYADSVKGRFTVSRDNPSTVSLQMNSLKPEDTGVYYCMAGSKSG
HELDHWGQGTQVTVSS (SEQ ID NO:11);
 - (b) MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYRQAPGQGL
EWVTSIYSDGSTVYADSVKGRFTVSRDNPSTVSLQMNSLKPEDTGVYHGMAGSKSGQ
ELDHWGQGTQVTVSS (SEQ ID NO:1); and

(c) MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYRQAPGQGL
EWVTSIYSDGSTVYADSVKGRFTVSRDNPSTVSLQMNSLKPEDTGVYYCMAGSKSGH
ELDHWGQGTQVTVSS (SEQ ID NO:7).

9. The isolated anti-SARS-CoV-2 nanobody of claim 1, comprising an amino acid sequence that has at least 95% sequence identity to any one of:

(a) MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYRQVPGQGL
EWVTSIYSDGRTVYADSVKGRFTVSRDNPSTVSLQMNSLKPEDTGVYYCMAGSKSG
HELDHWGQGTQVTVSS (SEQ ID NO:11);

(b) MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYRQAPGQGL
EWVTSIYSDGSTVYADSVKGRFTVSRDNPSTVSLQMNSLKPEDTGVYHGMAGSKSGQ
ELDHWGQGTQVTVSS (SEQ ID NO:1); and

(c) MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYRQAPGQGL
EWVTSIYSDGSTVYADSVKGRFTVSRDNPSTVSLQMNSLKPEDTGVYYCMAGSKSGH
ELDHWGQGTQVTVSS (SEQ ID NO:7).

10. The isolated anti-SARS-CoV-2 nanobody of claim 1, comprising an amino acid sequence that has at least 97% sequence identity to any one of:

(a) MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYRQVPGQGL
EWVTSIYSDGRTVYADSVKGRFTVSRDNPSTVSLQMNSLKPEDTGVYYCMAGSKSG
HELDHWGQGTQVTVSS (SEQ ID NO:11);

(b) MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYRQAPGQGL
EWVTSIYSDGSTVYADSVKGRFTVSRDNPSTVSLQMNSLKPEDTGVYHGMAGSKSGQ
ELDHWGQGTQVTVSS (SEQ ID NO:1); and

(c) MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYRQAPGQGL
EWVTSIYSDGSTVYADSVKGRFTVSRDNPSTVSLQMNSLKPEDTGVYYCMAGSKSGH
ELDHWGQGTQVTVSS (SEQ ID NO:7).

11. The isolated anti-SARS-CoV-2 nanobody of claim 1, comprising an amino acid sequence selected from the group consisting of:

(a) MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYRQVPGQGL
EWVTSIYSDGRTVYADSVKGRFTVSRDNPSTVSLQMNSLKPEDTGVYYCMAGSKSG
HELDHWGQGTQVTVSS (SEQ ID NO:11);

(b) MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYRQAPGQGL
EWVTSIYSDGSTVYADSVKGRFTVSRDNPSTVSLQMNSLKPEDTGVYHGMAGSKSG
QELDHWGQGTQVTVSS (SEQ ID NO:1); and

(c) MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYRQAPGQGL
EWVTSIYSDGSTVYADSVKGRFTVSRDNPSTVSLQMNSLKPEDTGVYYCMAGSKSGH
ELDHWGQGTQVTVSS (SEQ ID NO:7).

12. The isolated anti-SARS-CoV-2 nanobody of claim 1, comprising SEQ ID NO:11.
13. The isolated anti-SARS-CoV-2 nanobody of claim 1, comprising SEQ ID NO:1.
14. The isolated anti-SARS-CoV-2 nanobody of claim 1, comprising SEQ ID NO:7.
15. The isolated anti-SARS-CoV-2 nanobody as described in any one of claims 1-14, wherein the nanobody is operably linked to at least one detectable agent.
16. The isolated anti-SARS-CoV-2 nanobody of claim 15, wherein the at least one detectable agent is a polypeptide tag.
17. The isolated anti-SARS-CoV-2 nanobody of claim 16, wherein the at least one polypeptide tag is operably linked to the N-terminus of the nanobody.
18. The isolated anti-SARS-CoV-2 nanobody of claim 16, wherein the at least one polypeptide tag is operably linked to the C-terminus of the nanobody.
19. The isolated anti-SARS-CoV-2 nanobody of claim 16, comprising an amino acid sequence that has at least 90% sequence identity to any one of:
 - (a) MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYRQVPGQGL
EWVTSIYSDGRTVYADSVKGRFTVSRDNPSTVSLQMNSLKPEDTGVYYCMAGSKSG
HELDHWGQGTQVTVSSGPGGQH HHHHHHGA YPYDVPDYAS (SEQ ID NO:13);
 - (b) MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYRQAPGQGL
EWVTSIYSDGSTVYADSVKGRFTVSRDNPSTVSLQMNSLKPEDTGVYHGMAGSKSGQ
ELDHWGQGTQVTVSSGPGGQH HHHHHHGA YPYDVPDYAS (SEQ ID NO:5); and

(c) MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYRQAPGQGL
EWVTSIYSDGSTVYADSVKGRFTVSRDNPSTVSLQMNSLKPEDTGVYYCMAGSKSGH
ELDHWGQGTQVTVSSGPGGQH HHHHHHGA YPYDVPDYAS (SEQ ID NO:9).

20. The isolated anti-SARS-CoV-2 nanobody of claim 16, comprising an amino acid sequence selected from the group consisting of:

(a) MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYRQVPGQGL
EWVTSIYSDGRTVYADSVKGRFTVSRDNPSTVSLQMNSLKPEDTGVYYCMAGSKSG
HELDHWGQGTQVTVSSGPGGQH HHHHHHGA YPYDVPDYAS (SEQ ID NO:13);

(b) MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYRQAPGQGL
EWVTSIYSDGSTVYADSVKGRFTVSRDNPSTVSLQMNSLKPEDTGVYHGMAGSKSGQ
ELDHWGQGTQVTVSSGPGGQH HHHHHHGA YPYDVPDYAS (SEQ ID NO:5); and

(c) MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYRQAPGQGL
EWVTSIYSDGSTVYADSVKGRFTVSRDNPSTVSLQMNSLKPEDTGVYYCMAGSKSGH
ELDHWGQGTQVTVSSGPGGQH HHHHHHGA YPYDVPDYAS (SEQ ID NO:9).

21. A nanobody-Fc fusion polypeptide comprising an isolated anti-SARS-CoV-2 nanobody as described in any one of claims 1-20 operably linked to an Fc domain amino acid sequence.

22. The nanobody-Fc fusion polypeptide of claim 21, comprising an amino acid sequence that has at least 90% sequence identity to any one of:

(a) MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYRQVPGQGL
EWVTSIYSDGRTVYADSVKGRFTVSRDNPSTVSLQMNSLKPEDTGVYYCMAGSKSG
HELDHWGQGTQVTVSSEPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEV
TCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLN
GKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYP
SDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFCFSVMHEAL
HNHYTQKSLSLSPGK (SEQ ID NO:14);

(b) MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYRQAPGQGL
EWVTSIYSDGSTVYADSVKGRFTVSRDNPSTVSLQMNSLKPEDTGVYHGMAGSKSGQ
ELDHWGQGTQVTVSSEPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVT
CVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNG
KEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPS

DIAVEWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALH
NHYTQKSLSLSPGK (SEQ ID NO:6); and

(c) MAQVQLVESGGGLVQPGGSLRLSCAASGFTFKNADMNWYRQAPGQGL
EWVTSIYSDGSTVYADSVKGRFTVSRDNPSTVSLQMNSLKPEDTGVIYCMAGSKSGH
ELDHWGQGTQVTVSSEPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVT
CVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNG
KEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPS
DIAVEWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALH
NHYTQKSLSLSPGK (SEQ ID NO:10).

23. The nanobody-Fc fusion polypeptide of claim 21, comprising an amino acid sequence that has at least 95% sequence identity to any one of SEQ ID NOs:14, 6, and 10.

24. The nanobody-Fc fusion polypeptide of claim 21, comprising an amino acid sequence that has at least 97% sequence identity to any one of SEQ ID NOs:14, 6, and 10.

25. The nanobody-Fc fusion polypeptide of claim 21, comprising any one of SEQ ID NOs:14, 6, and 10.

26. A protein molecule comprising two independently selected nanobody-Fc fusion polypeptides as described in any one of claims 21-25, wherein the two Fc polypeptides are linked to form a dimer.

27. The protein molecule of claim 26, wherein the two nanobody-Fc fusion polypeptides are the same.

28. The protein molecule of claim 26, wherein the two nanobody-Fc fusion polypeptides are different.

29. The isolated anti-SARS-CoV-2 nanobody of any one of claims 1-20, the nanobody-Fc fusion polypeptide of any one of claims 21-25, or the protein molecule of any one of claims 26-28, which is an inhibitor of SARS-CoV-2.

30. The isolated anti-SARS-CoV-2 nanobody of any one of claims 1-20, the polypeptide of any one of claims 21-25, or the protein molecule of any one of claims 26-28, which is humanized.
31. A composition comprising an isolated anti-SARS-CoV-2 nanobody of any one of claims 1-20, a nanobody-Fc fusion polypeptide of any one of claims 21-25, or a protein molecule of any one of claims 26-28, and a carrier.
32. The composition of claim 31, which is a pharmaceutical composition comprising a pharmaceutically acceptable carrier.
33. An isolated polynucleotide comprising a nucleotide sequence encoding an isolated anti-SARS-CoV-2 nanobody of any one of claims 1-20 or a nanobody-Fc fusion polypeptide of any one of claims 21-25.
34. A vector comprising the polynucleotide of claim 33.
35. A cell comprising the polynucleotide of claim 33 or the vector of claim 34.
36. The cell of claim 35, which is a bacterial cell.
37. A method of detecting the presence of SARS-CoV-2 in a cell, the method comprising contacting the cell with an isolated anti-SARS-CoV-2 nanobody of any one of claims 1-20, a nanobody-Fc fusion polypeptide of any one of claims 21-25, or a protein molecule of any one of claims 26-28, and detecting whether a complex is formed between 1) the nanobody, polypeptide or protein molecule; and 2) SARS-CoV-2.
38. The method of claim 37, wherein the cell is contacted *in vitro*.
39. The method of claim 37, wherein the cell is contacted *in vivo*.
40. A method of inhibiting the activity of SARS-CoV-2, comprising contacting SARS-CoV-2 with an isolated anti-SARS-CoV-2 nanobody of any one of claims 1-20, a nanobody-Fc fusion polypeptide of any one of claims 21-25, or the protein molecule of any one of claims 26-28.

41. The method of claim 40, wherein the SARS-CoV-2 is contacted *in vitro*.
42. The method of claim 40, wherein the SARS-CoV-2 is contacted *in vivo*.
43. The method of any one of claims 40-42, wherein the activity of the SARS-CoV-2 is inhibited by at least about 25% as compared to a control.
44. A method for treating or preventing a SARS-CoV-2 infection in a mammal, comprising administering an effective amount of an isolated anti-SARS-CoV-2 nanobody of any one of claims 1-20, a nanobody-Fc fusion polypeptide of any one of claims 21-25, or a protein molecule of any one of claims 26-28, to the mammal.
45. The method of claim 44, further comprising administering at least one additional therapeutic agent to the mammal.
46. An isolated anti-SARS-CoV-2 nanobody of any one of claims 1-20, a nanobody-Fc fusion polypeptide of any one of claims 21-25, or a protein molecule of any one of claims 26-28, for the prophylactic or therapeutic treatment of a SARS-CoV-2 infection.
47. The use of an isolated anti-SARS-CoV-2 nanobody of any one of claims 1-20, a nanobody-Fc fusion polypeptide of any one of claims 21-25, or a protein molecule of any one of claims 26-28, to prepare a medicament for the treatment of a SARS-CoV-2 infection in a mammal.
48. An isolated anti-SARS-CoV-2 nanobody of any one of claims 1-20, a nanobody-Fc fusion polypeptide of any one of claims 21-25, or a protein molecule of any one of claims 26-28, for use in medical therapy.
49. A kit comprising:
 - 1) an isolated anti-SARS-CoV-2 nanobody of any one of claims 1-20, a nanobody-Fc fusion polypeptide of any one of claims 21-25, or a protein molecule of any one of claims 26-28;
 - 2) packaging material; and
 - 3) instructions for administering the nanobody, polypeptide or protein molecule, to a mammal to treat or prevent a SARS-CoV-2 infection.

50. The kit of claim 49, further comprising at least one other therapeutic agent.

1/26

Biacore measurement: SARS-CoV-2 RBD (His tagged on the C terminus) covalently immobilized (via amine group) to sensor chip	K_d (M)	K_{off} (s ⁻¹)	K_{on} (M ⁻¹ s ⁻¹)
ACE2 (Shang et al., Nature, 2020)	4.42×10^{-8}	7.75×10^{-3}	1.75×10^5
Nanosota-1A (before affinity maturation)	2.28×10^{-7}	9.35×10^{-3}	4.10×10^4
Nanosota-1B (after 1 st round of affinity maturation)	6.08×10^{-8}	7.19×10^{-3}	1.18×10^5
Nanosota-1C (after 2 nd round of affinity maturation)	1.42×10^{-8}	2.96×10^{-3}	2.09×10^5
Nanosota-1C-Fc (after 2 nd round of affinity maturation)	1.57×10^{-11}	9.68×10^{-5}	6.15×10^6

FIGURE 1A

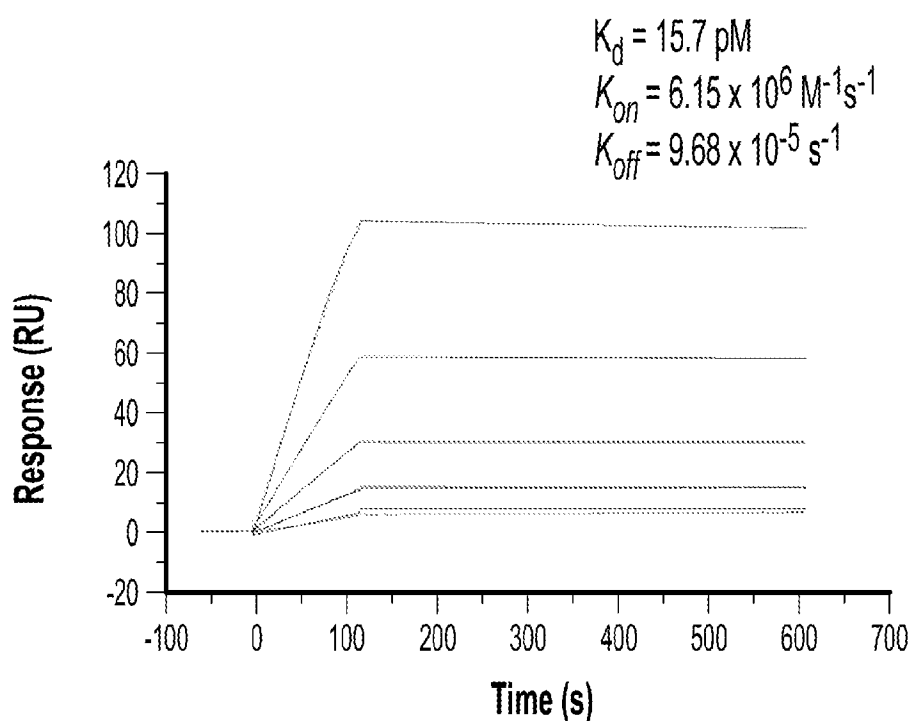


FIGURE 1B

2/26

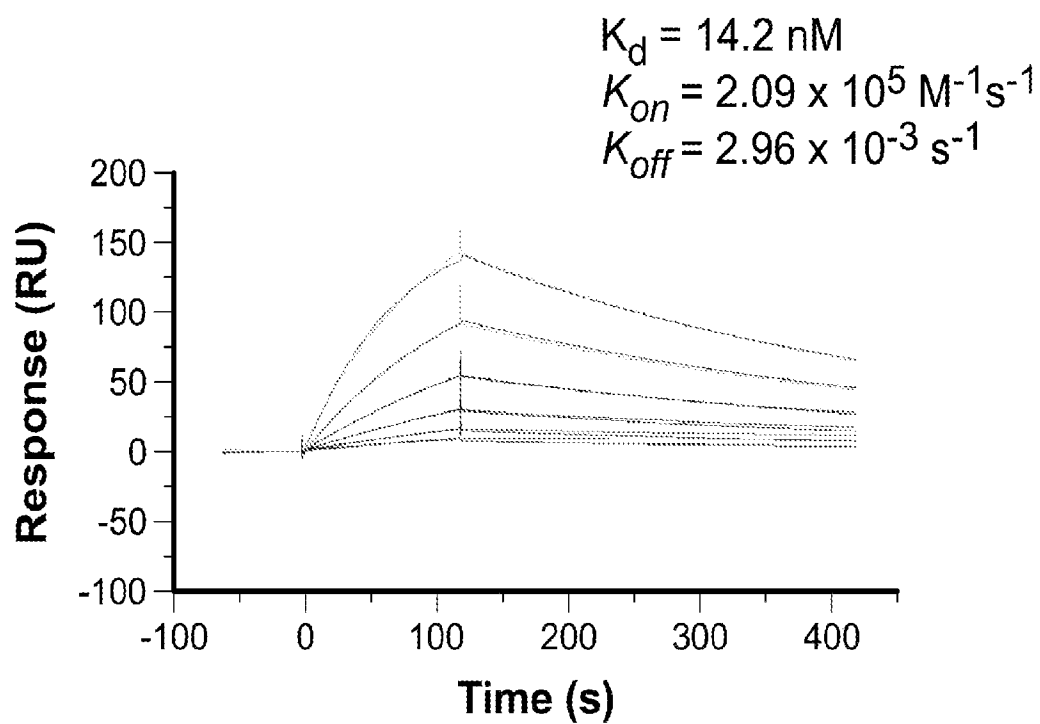


FIGURE 1C

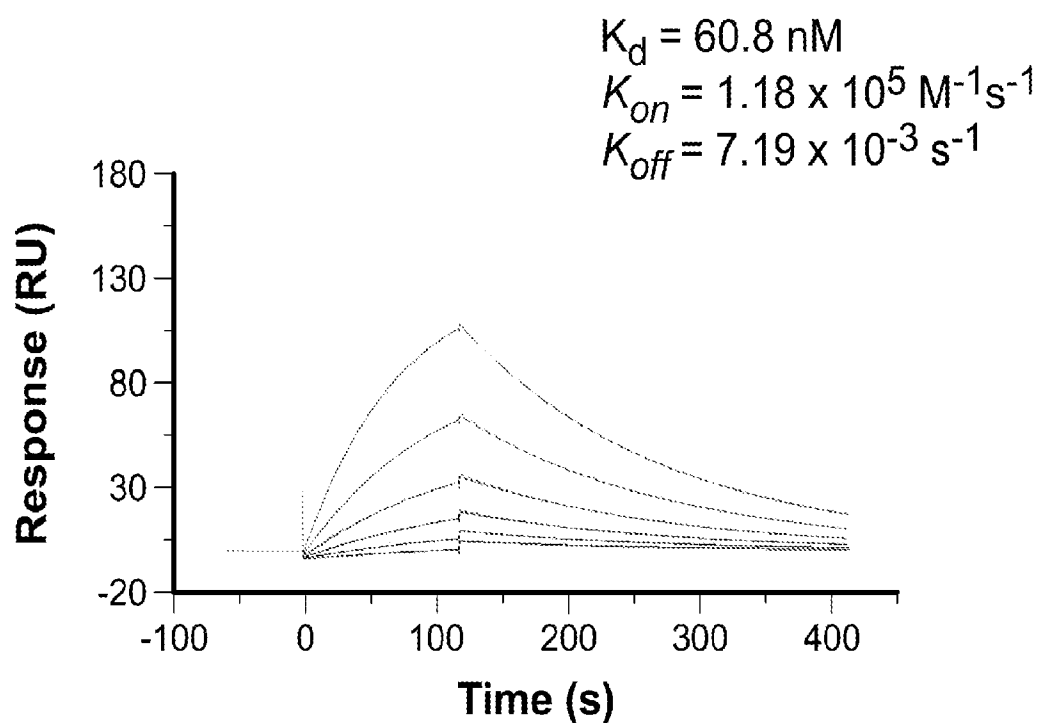


FIGURE 1D

3/26

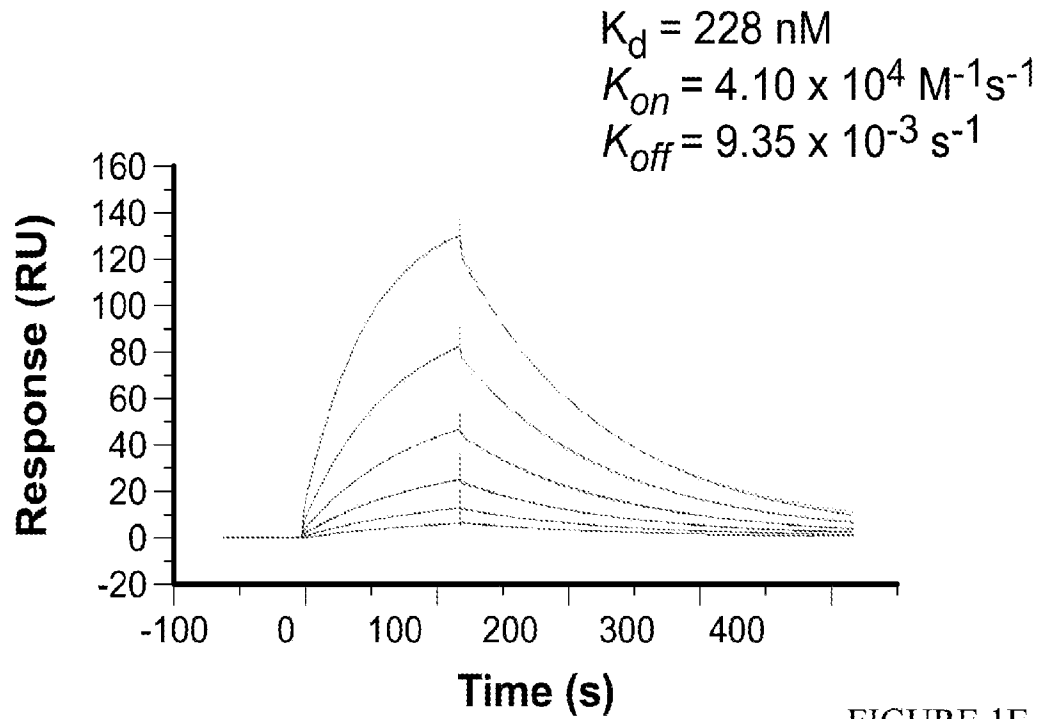


FIGURE 1E

Inhibition of SARS-CoV-2 pseudovirus entry by Nanosota-1

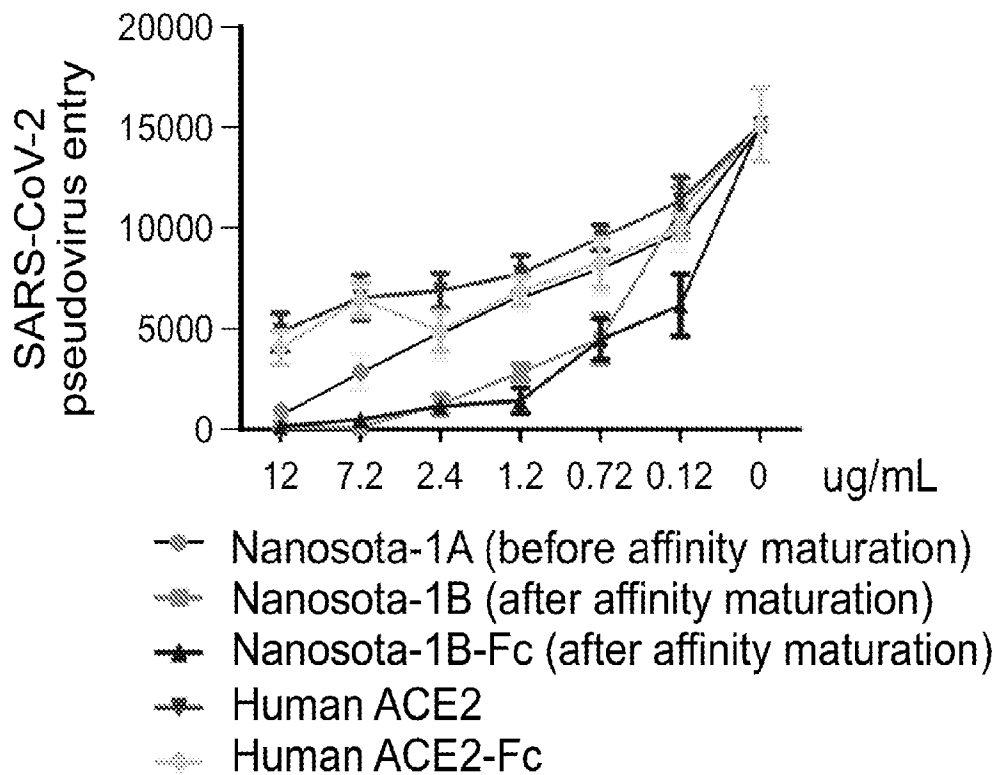


FIGURE 2A

4/26

Inhibition of SARS-CoV-2 pseudovirus entry by Nanosota-1

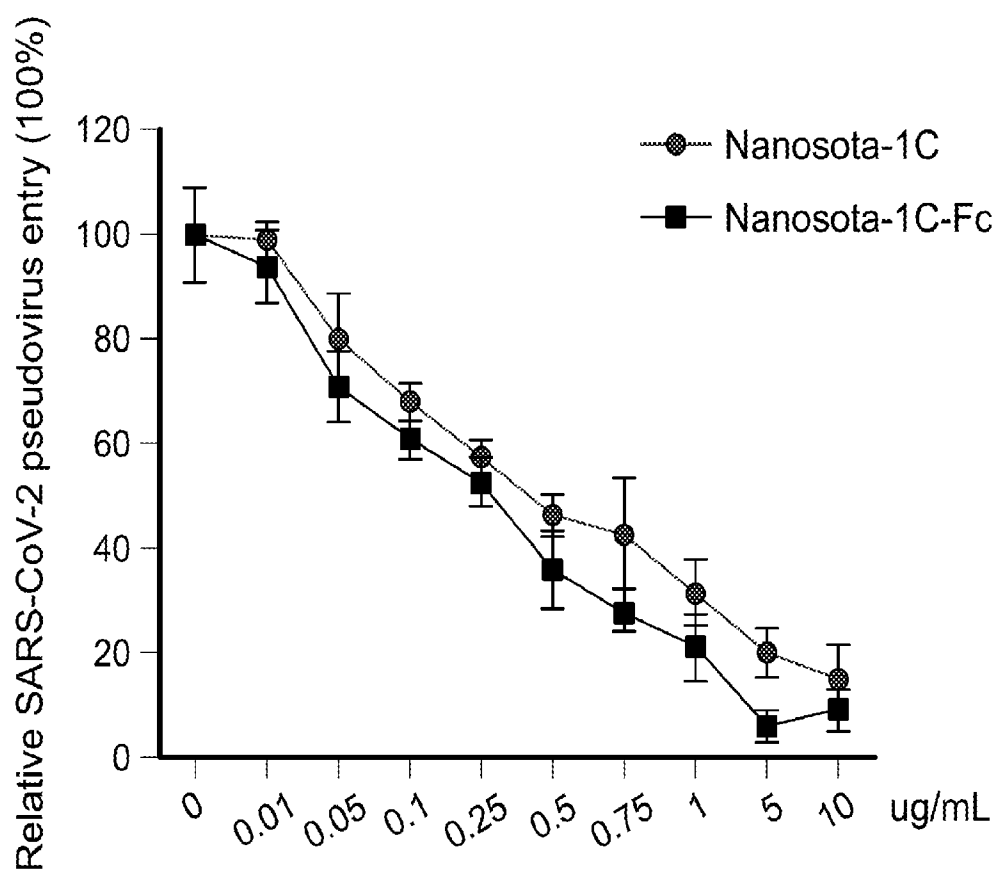


FIGURE 2B

5/26

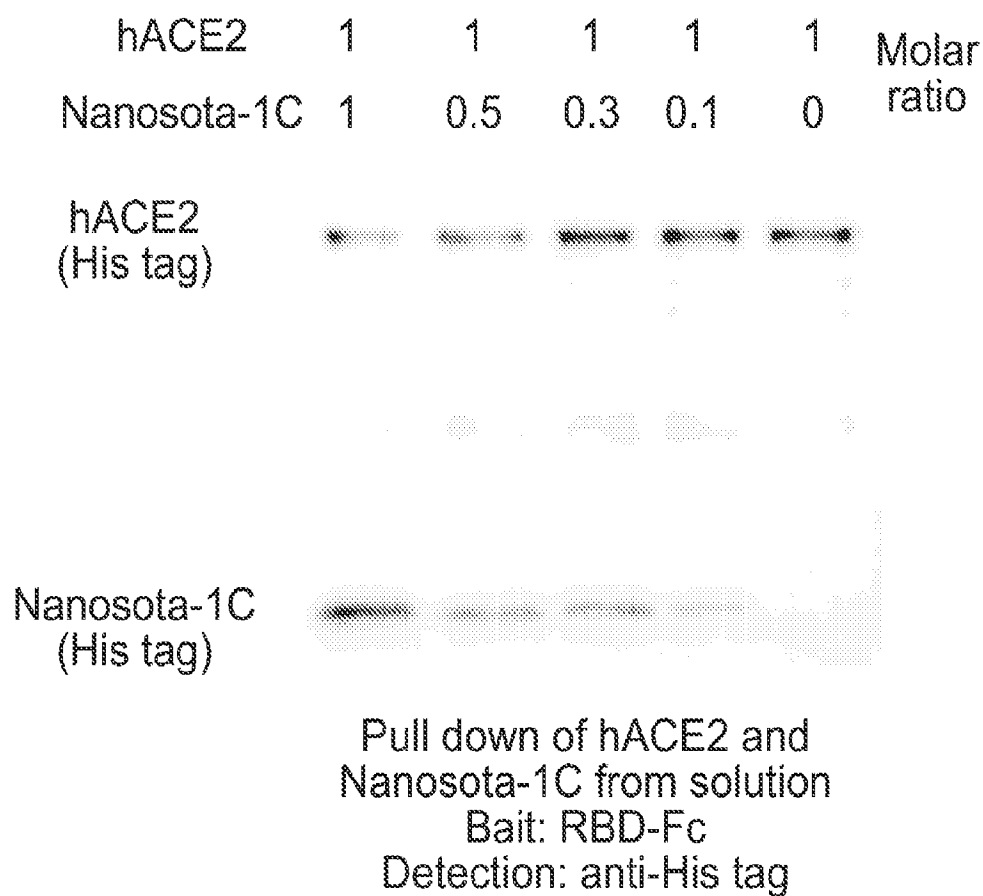


FIGURE 3A

6/26

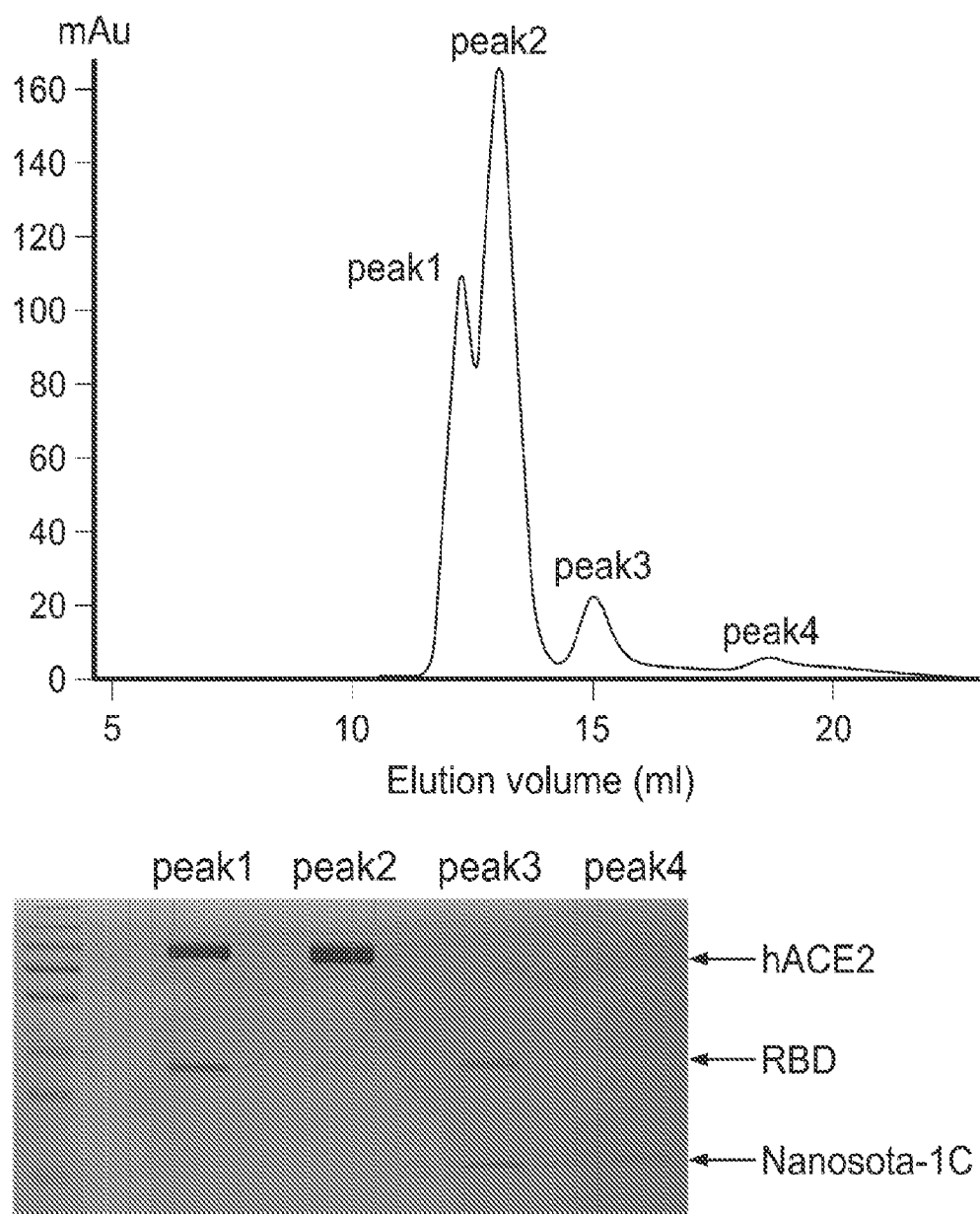


FIGURE 3B

7/26

Bioavailability of Nanosota-1C-Fc in mice

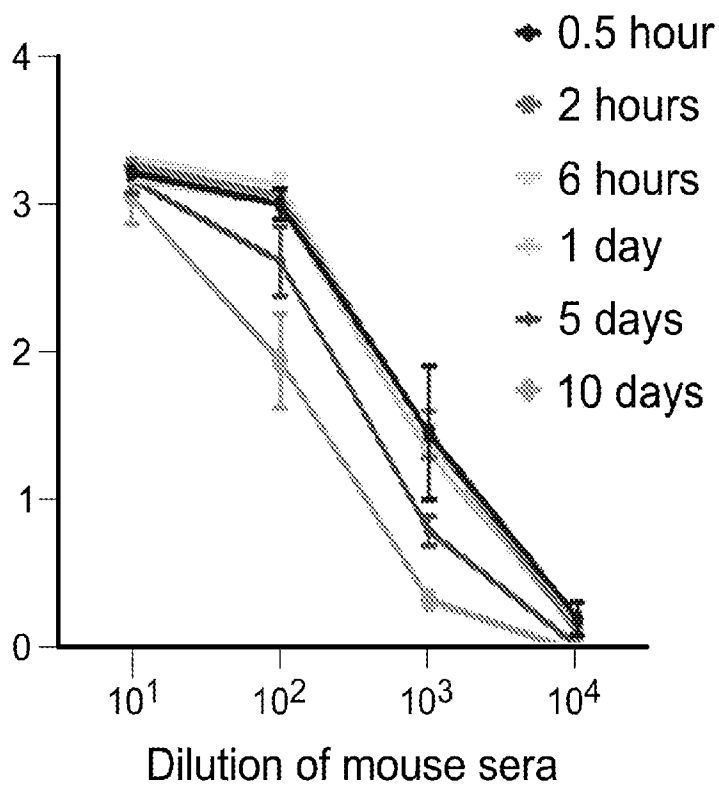


FIGURE 4A

Bioavailability of Nanosota-1C in mice

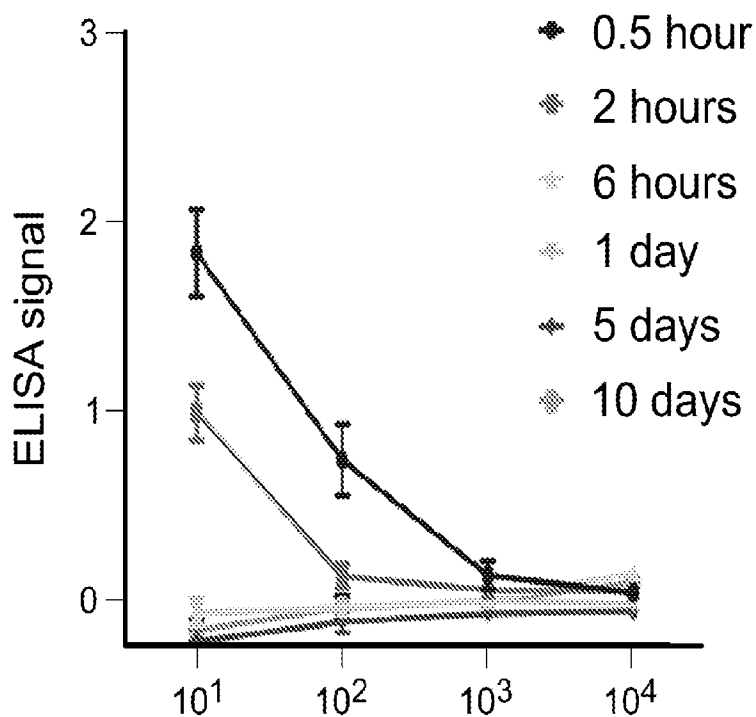


FIGURE 4B

8/26

Bioavailability of PBS control in mice

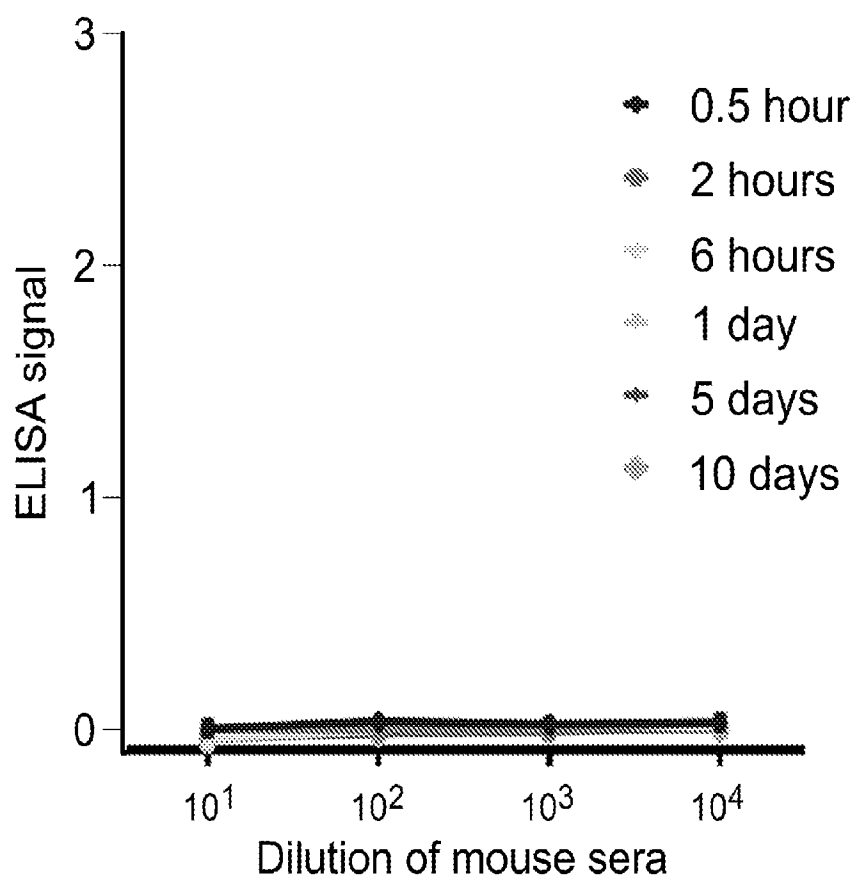


FIGURE 4C

9/26

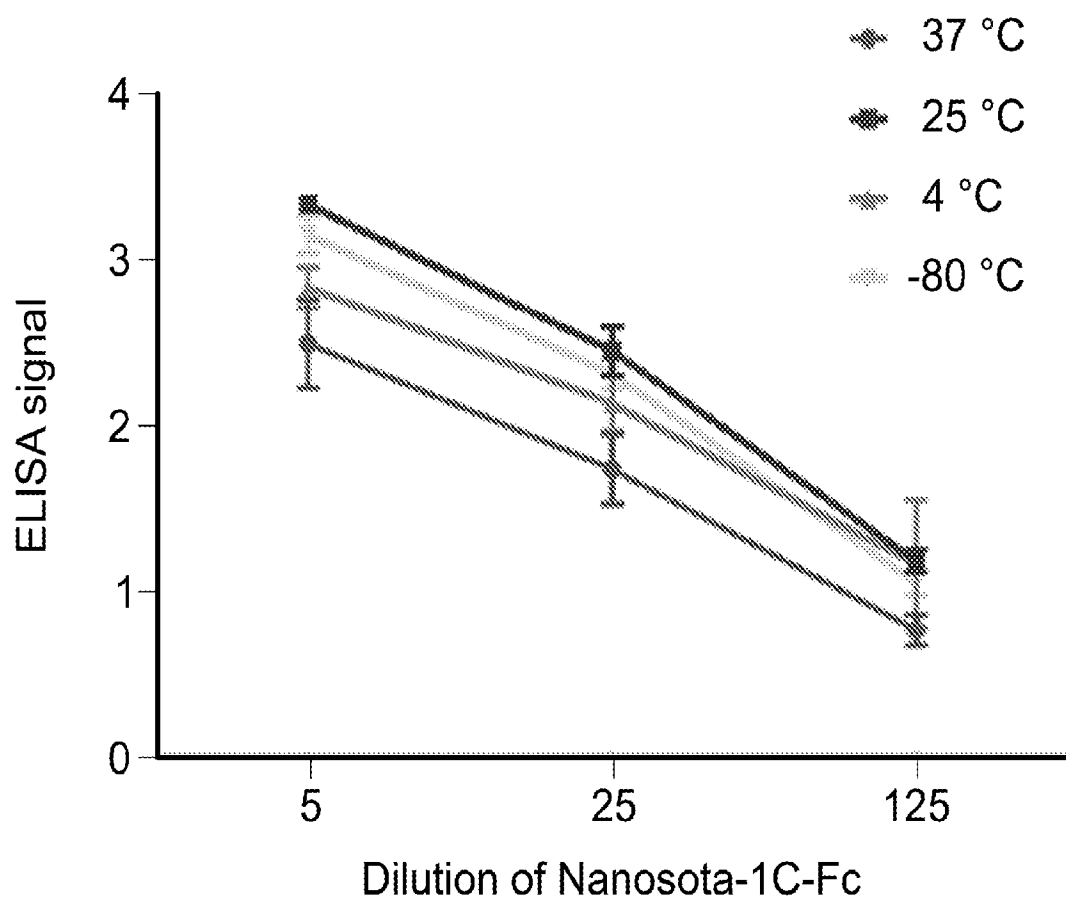


FIGURE 5

10/26

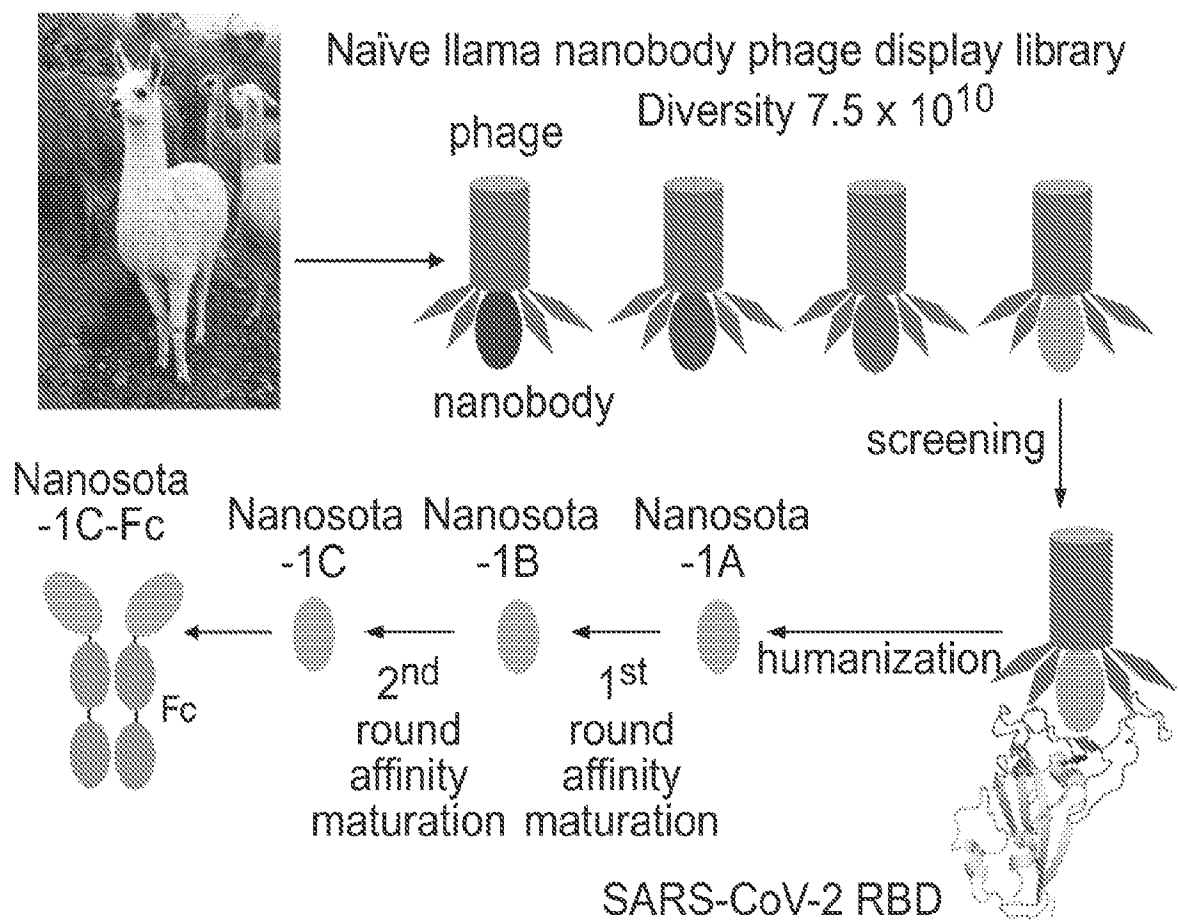


FIGURE 6

11/26

Molar ratio	1	1	1	1	1	1	1	ACE2
	8	4	2	1	0.5	0.25	0	Nanosota-1C



Pull down of ACE2-His and Nanosota-1C from solution
Bait: RBD-Fc; Detection: anti-His tag

FIGURE 7A

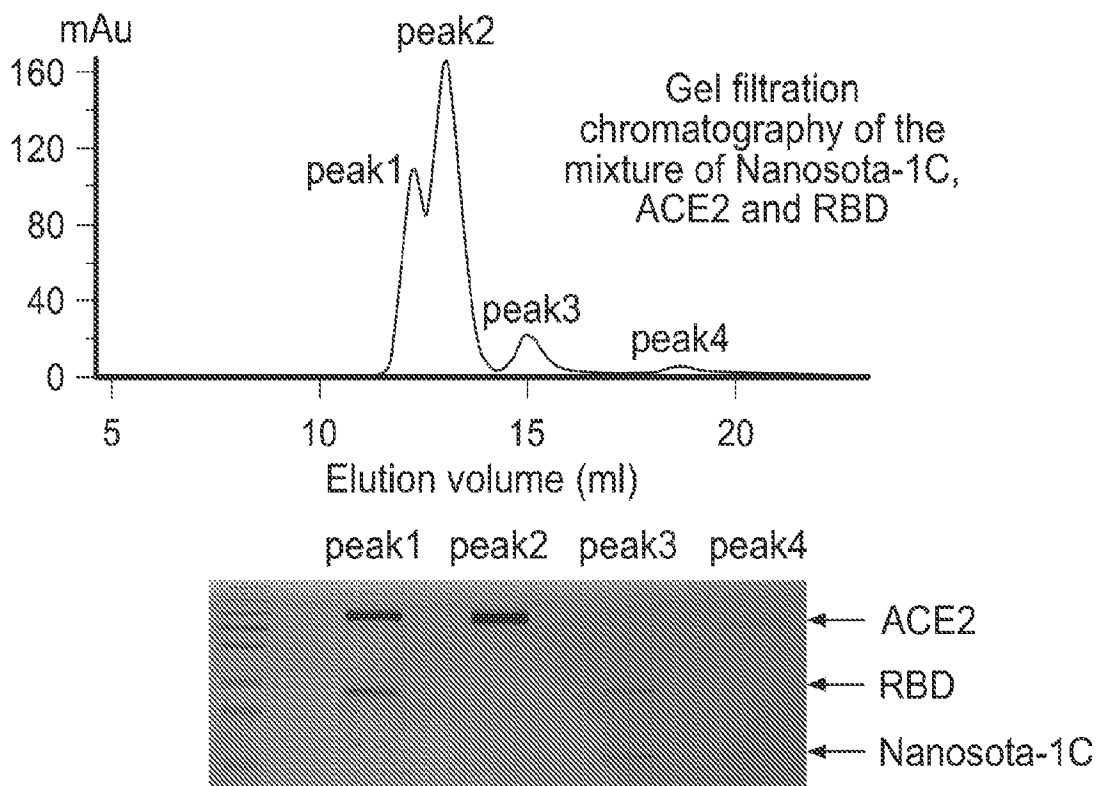


FIGURE 7B

12/26

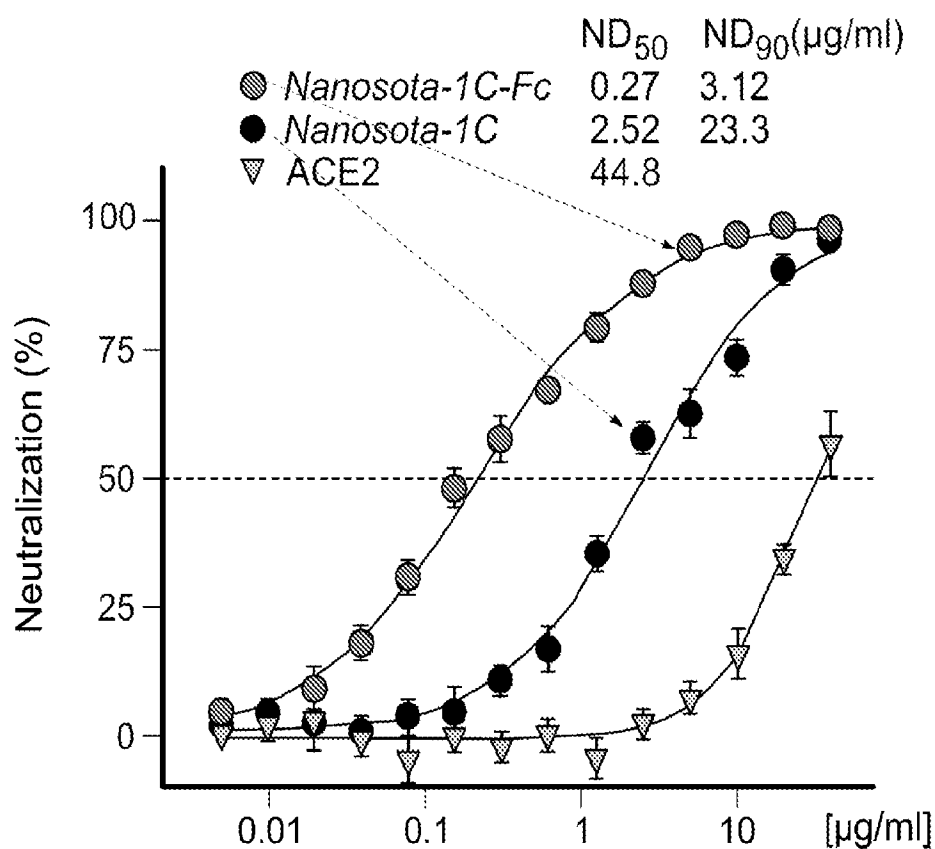


FIGURE 8A

Authentic SARS-CoV-2 neutralization

[μg/ml]	Nanosota-1C		Nanosota-1C-Fc		Virus control	Mock control	ACE2	
2000					+	-	+	-
1000	-	-	-	-	+	-	+	+
333	-	-	-	-	+	-	+ Infection - No infection	
111	-	-	-	-	+	-		
37	+	-	-	-	+	-		
12	+	+	-	-	+	-		
4	+	+	+	+	+	-		
1	+	+	+	+	+	-		
ND ₁₀₀			12 μg/ml					
ND ₅₀	37 μg/ml						2 mg/ml	

FIGURE 8B

13/26

Clinical outcome: body weight

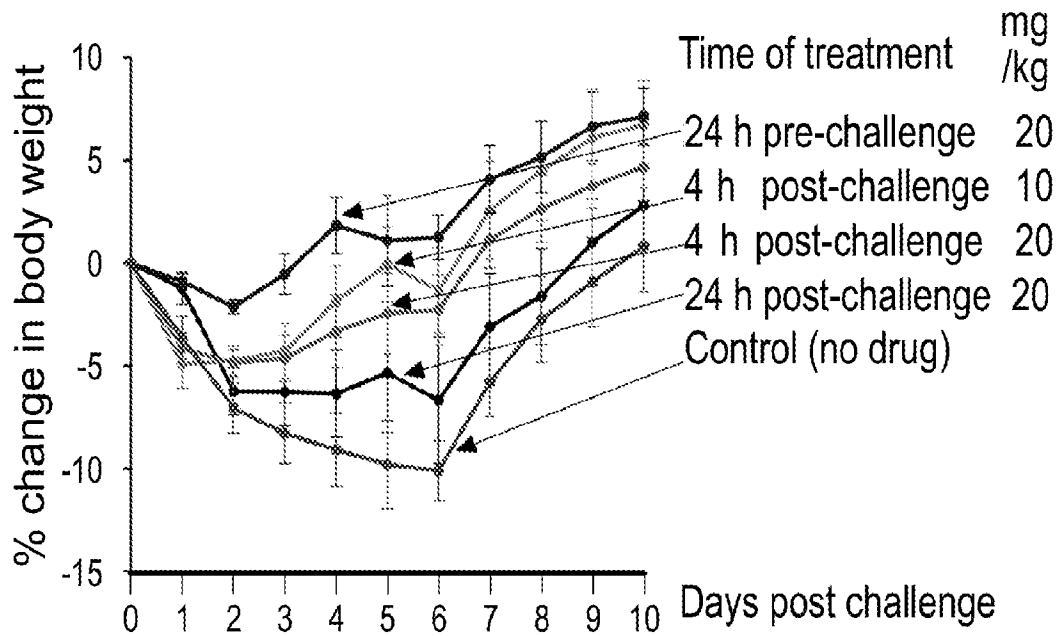


FIGURE 9A

Pathology: bronchiolo-alveolar hyperplasia

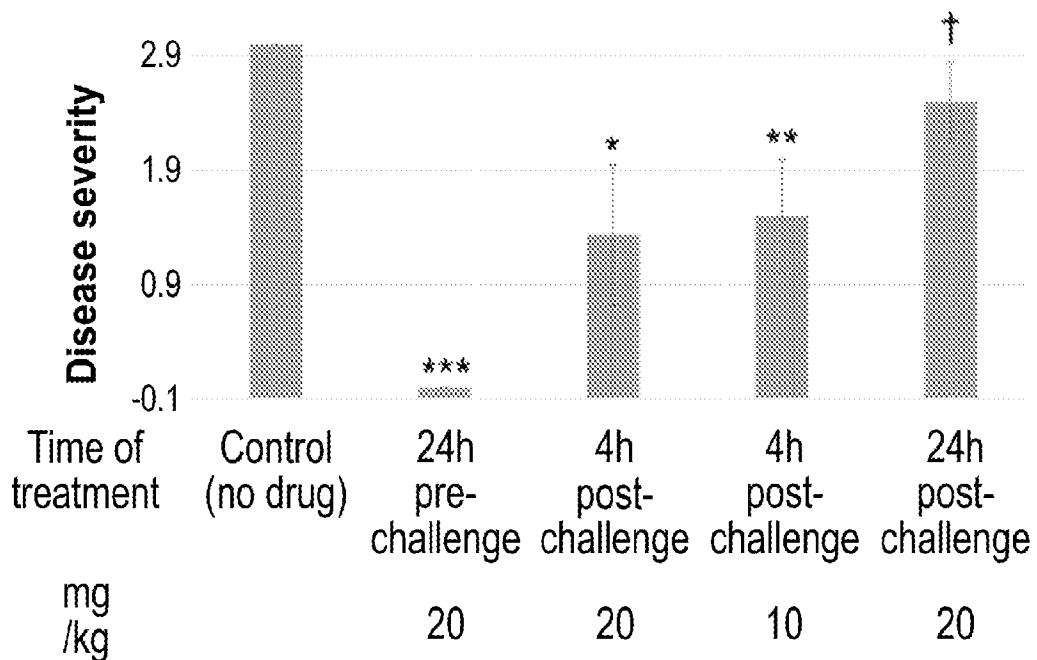


FIGURE 9B

14/26

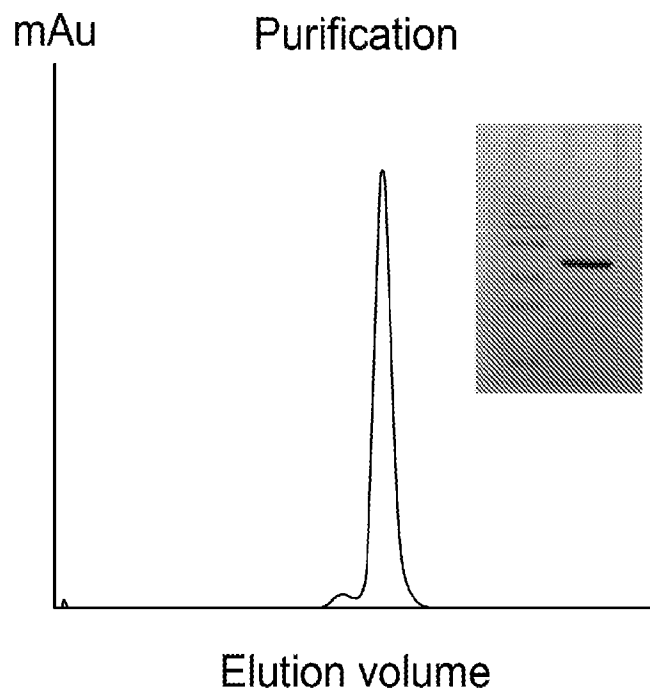


FIGURE 10A

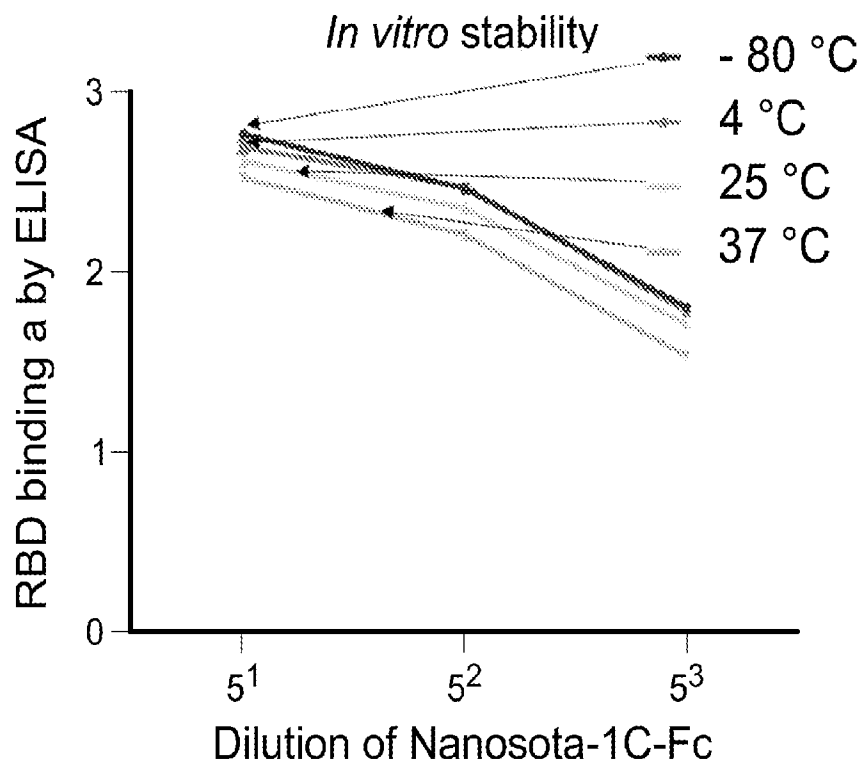


FIGURE 10B

15/26

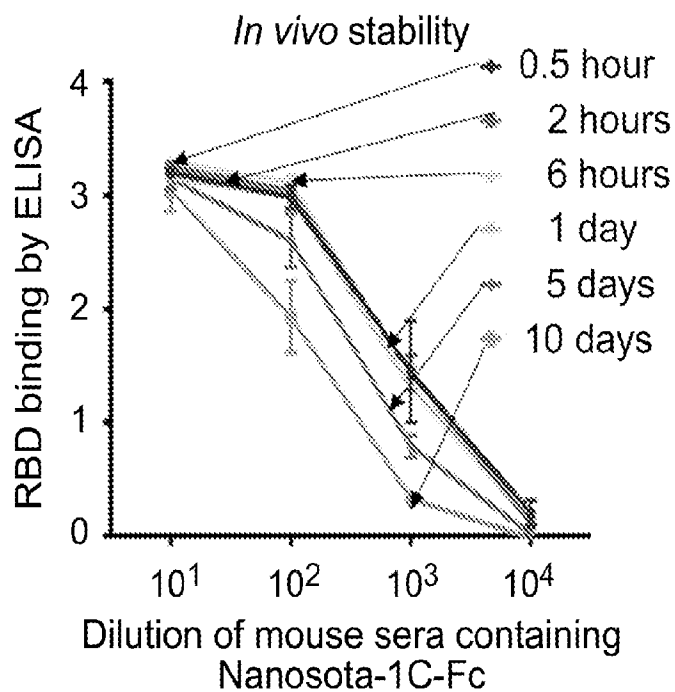


FIGURE 10C

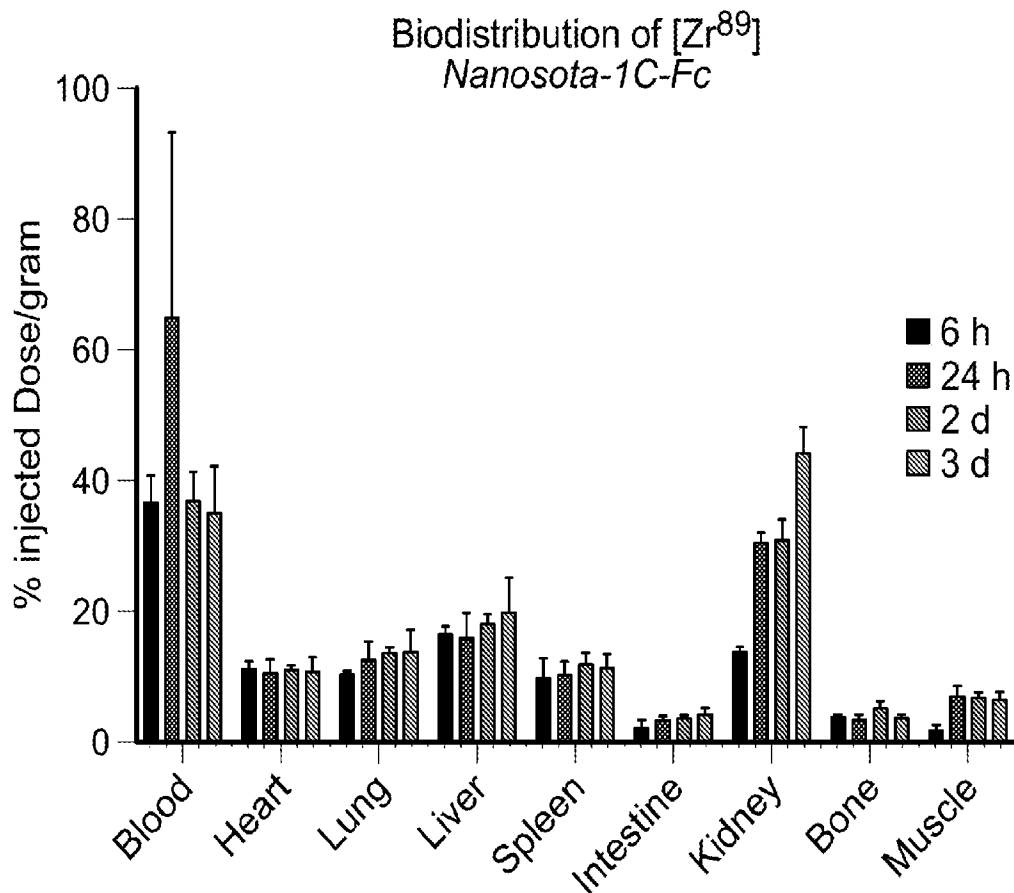
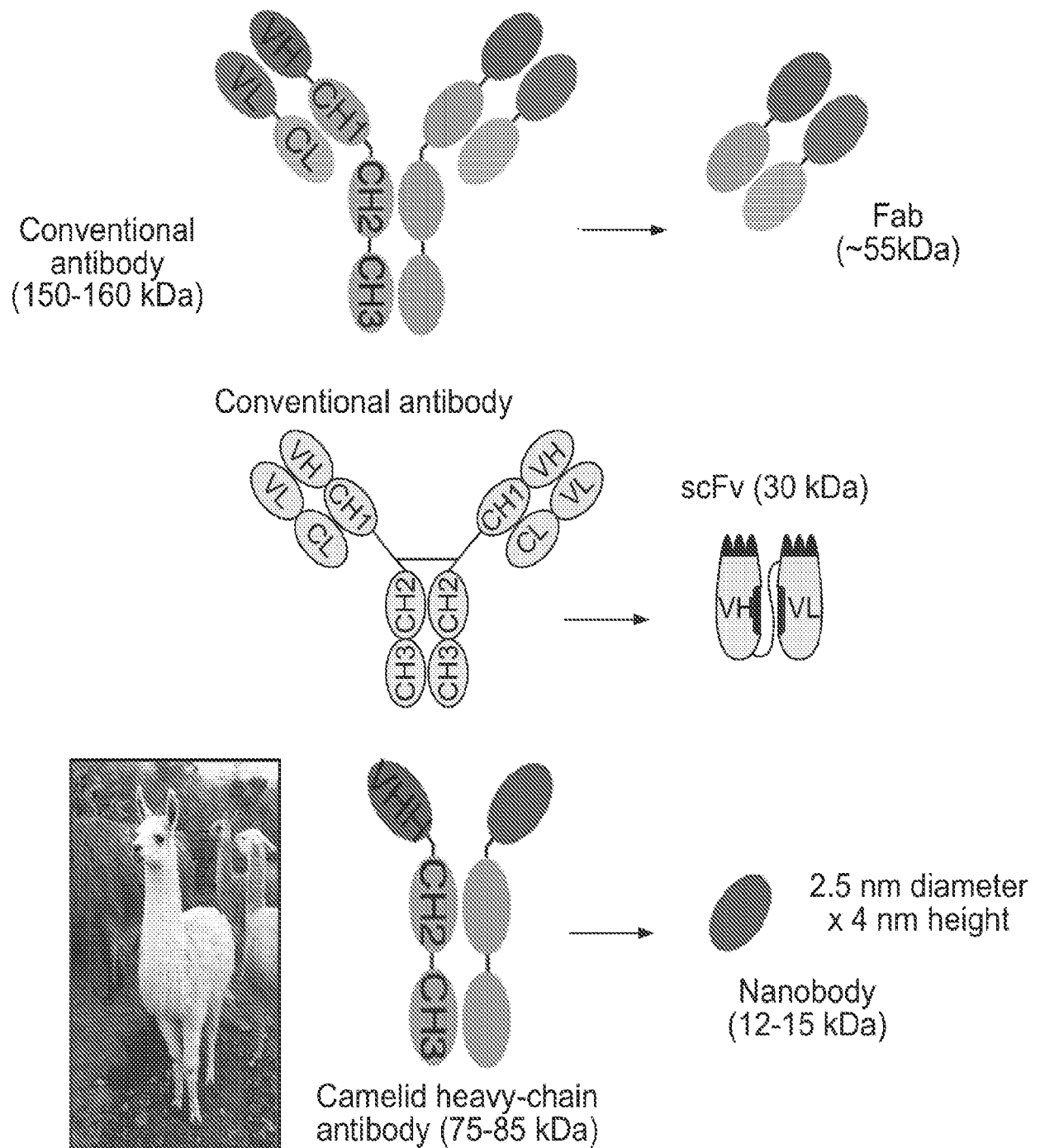


FIGURE 10D

16/26



2019: nanobody drug approved by FDA to treat blood clotting disorder

FIGURE 11

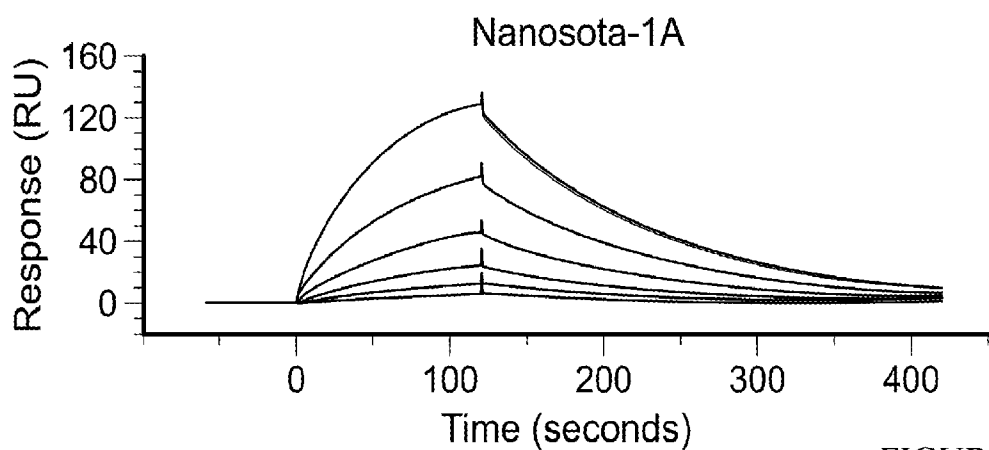
17/26

FIGURE 12A

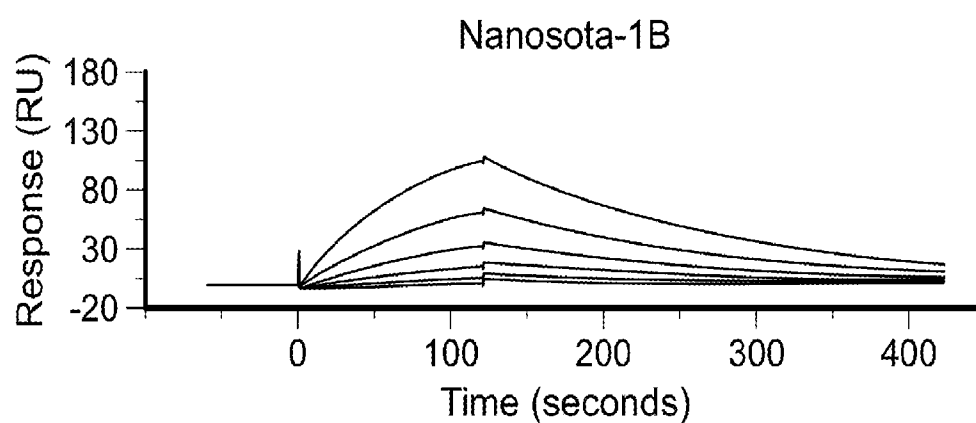


FIGURE 12B

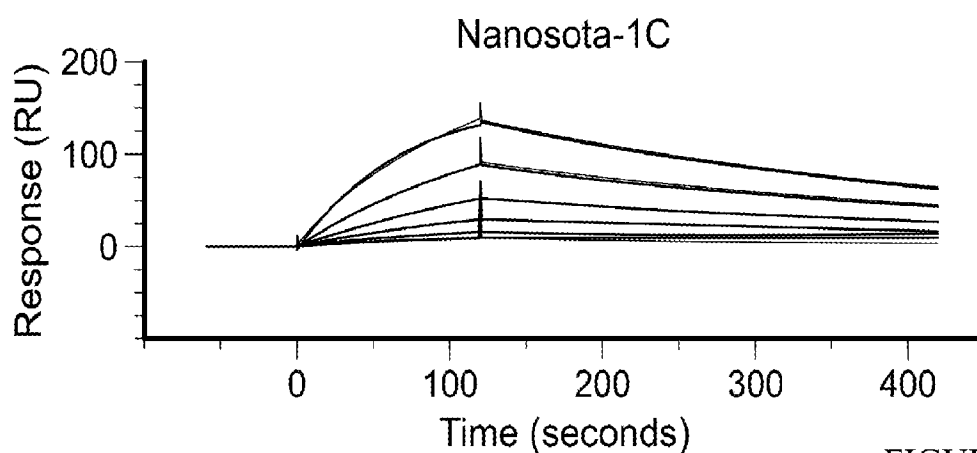


FIGURE 12C

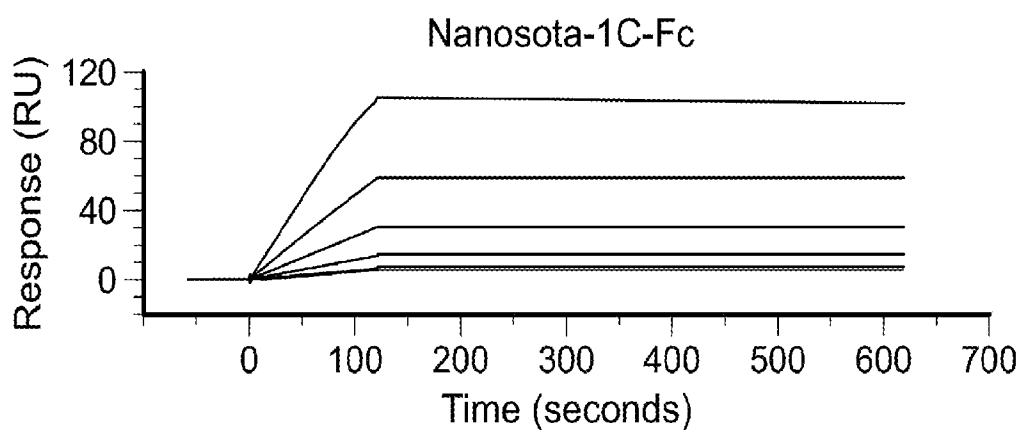


FIGURE 12D

18/26

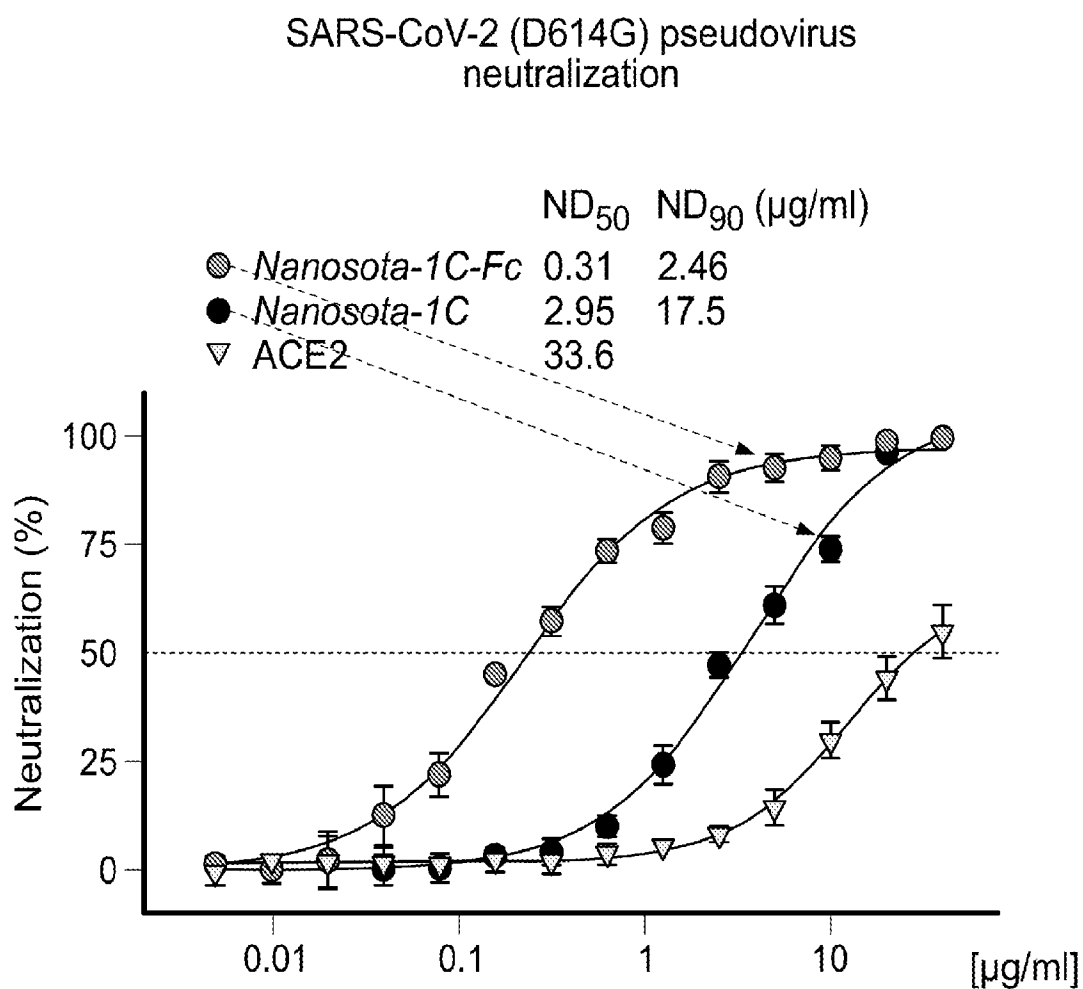


FIGURE 13

19/26

Virus loads in nasal swabs as measures by qRT-PCR

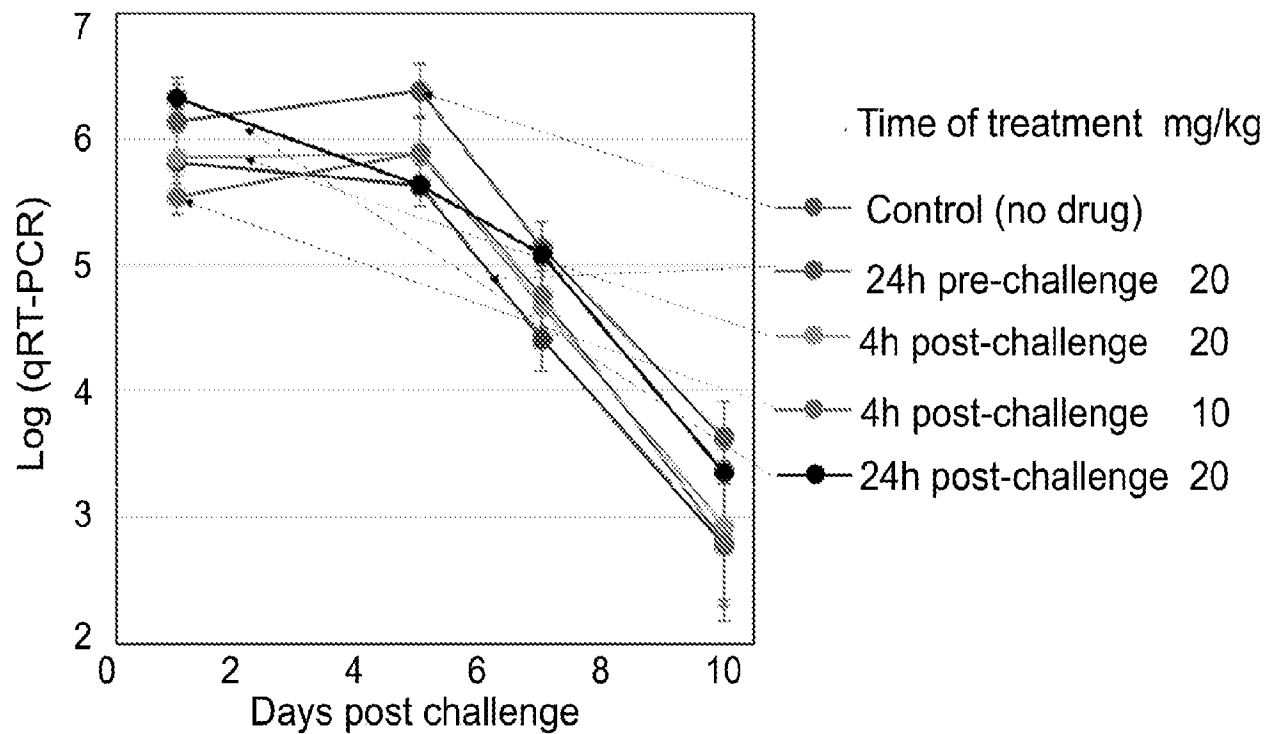


FIGURE 14

20/26

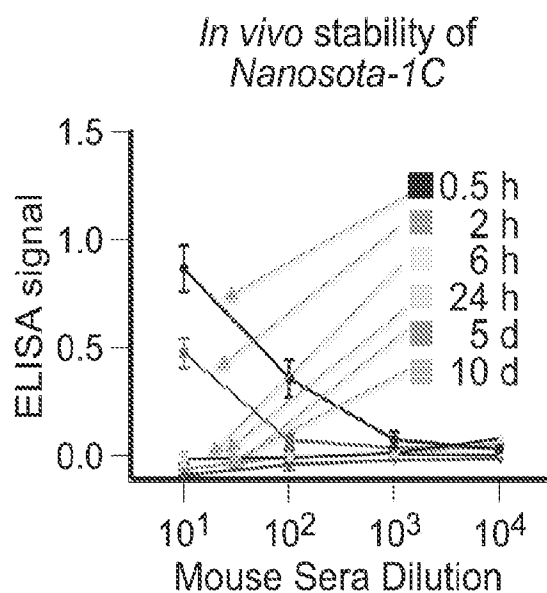


FIGURE 15A

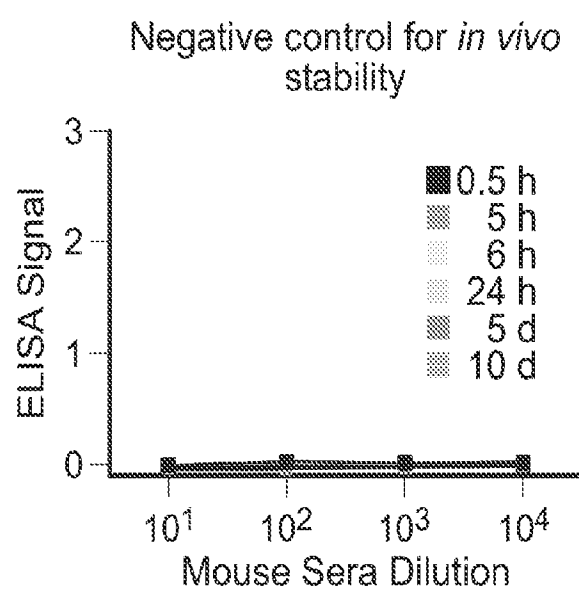


FIGURE 15B

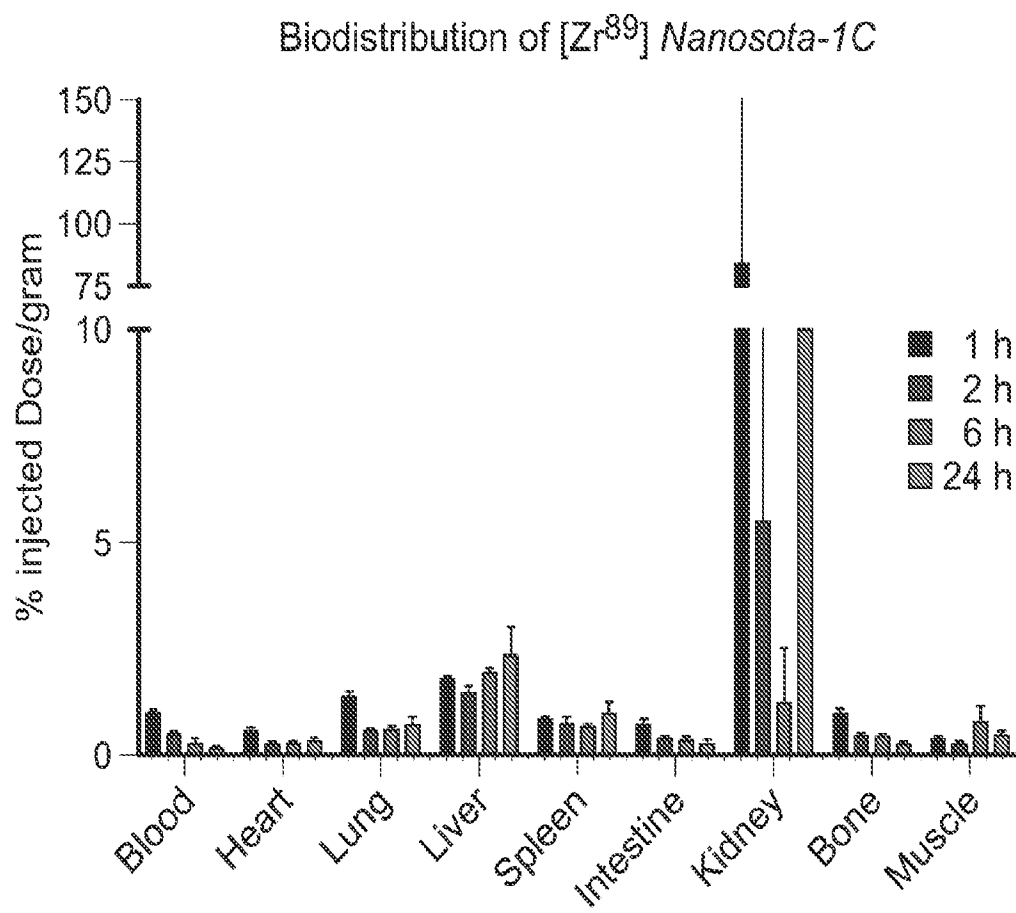


FIGURE 15C

21/26

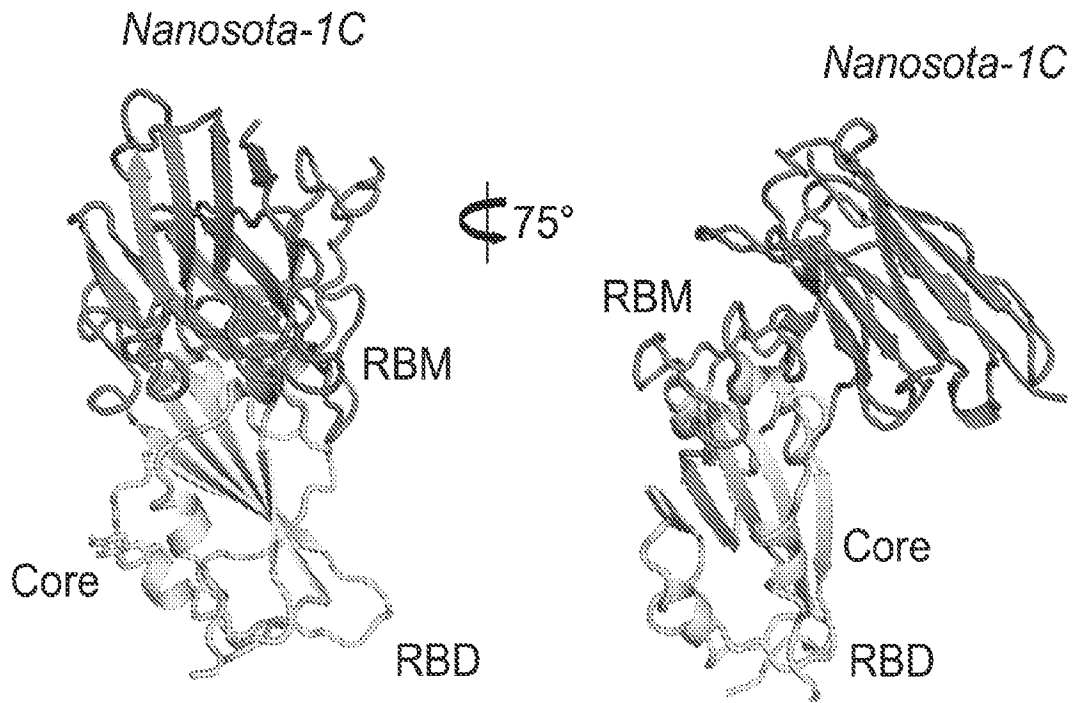


FIGURE 16A

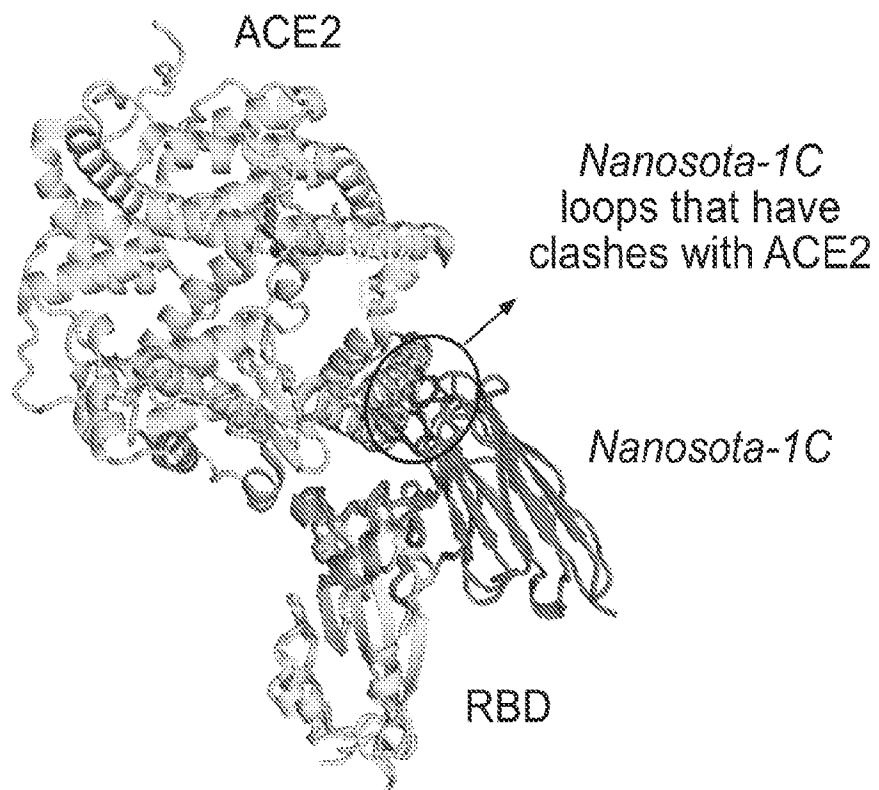


FIGURE 16B

22/26

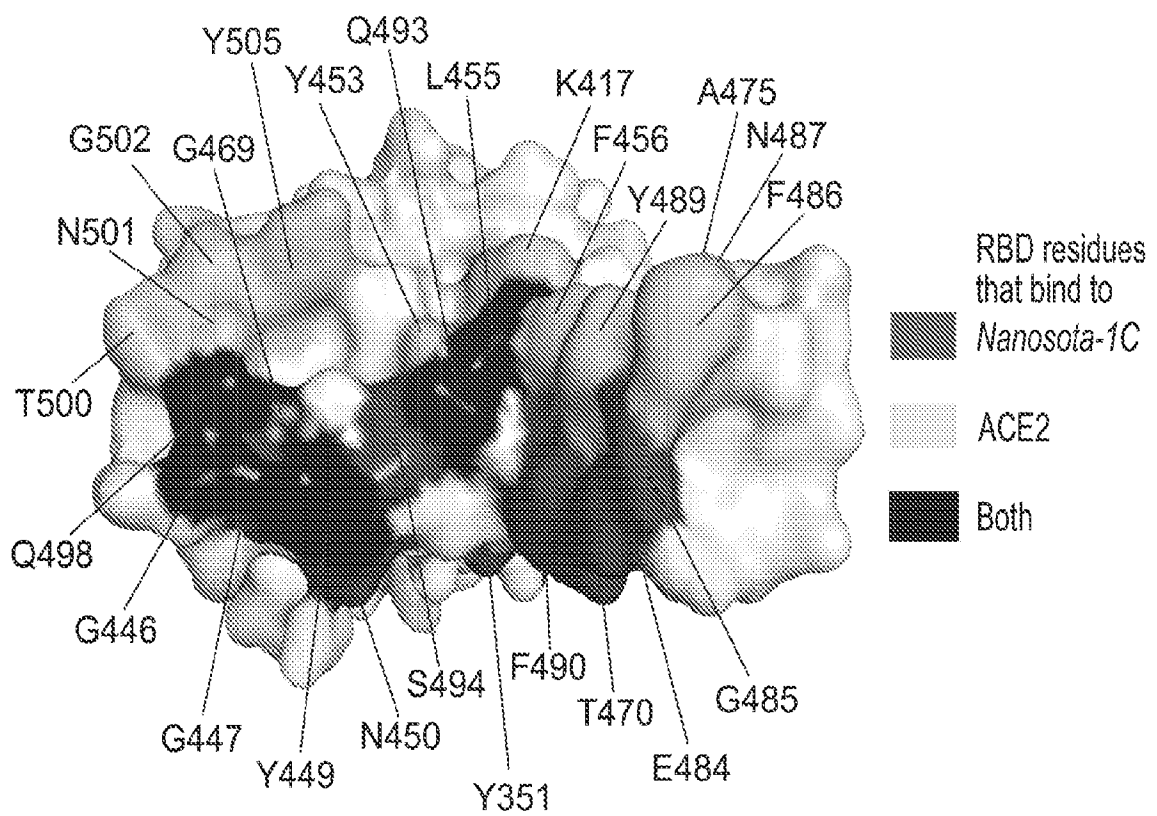
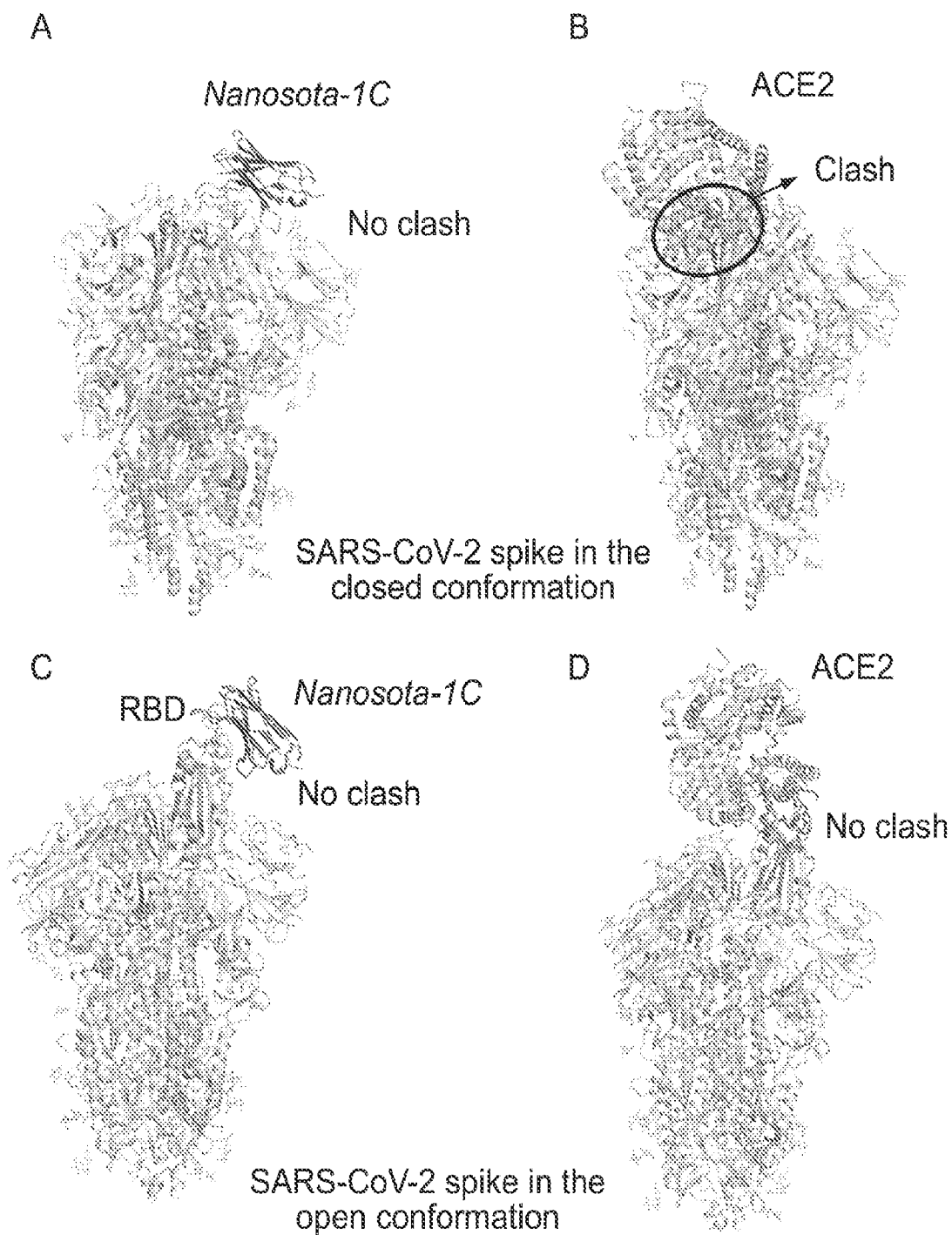


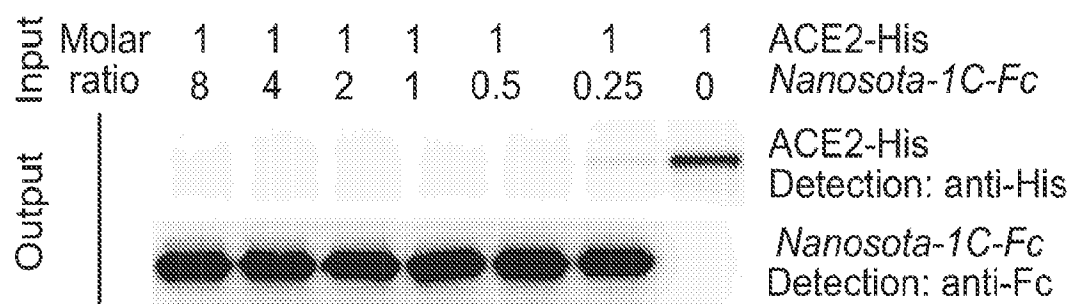
FIGURE 17

23/26



FIGURES 18A-18D

24/26



Pull down of ACE2-His and *Nanosota-1C-Fc*
Bait: biotinylated RBD-His

FIGURE 19

Live SARS-CoV-2 neutralization

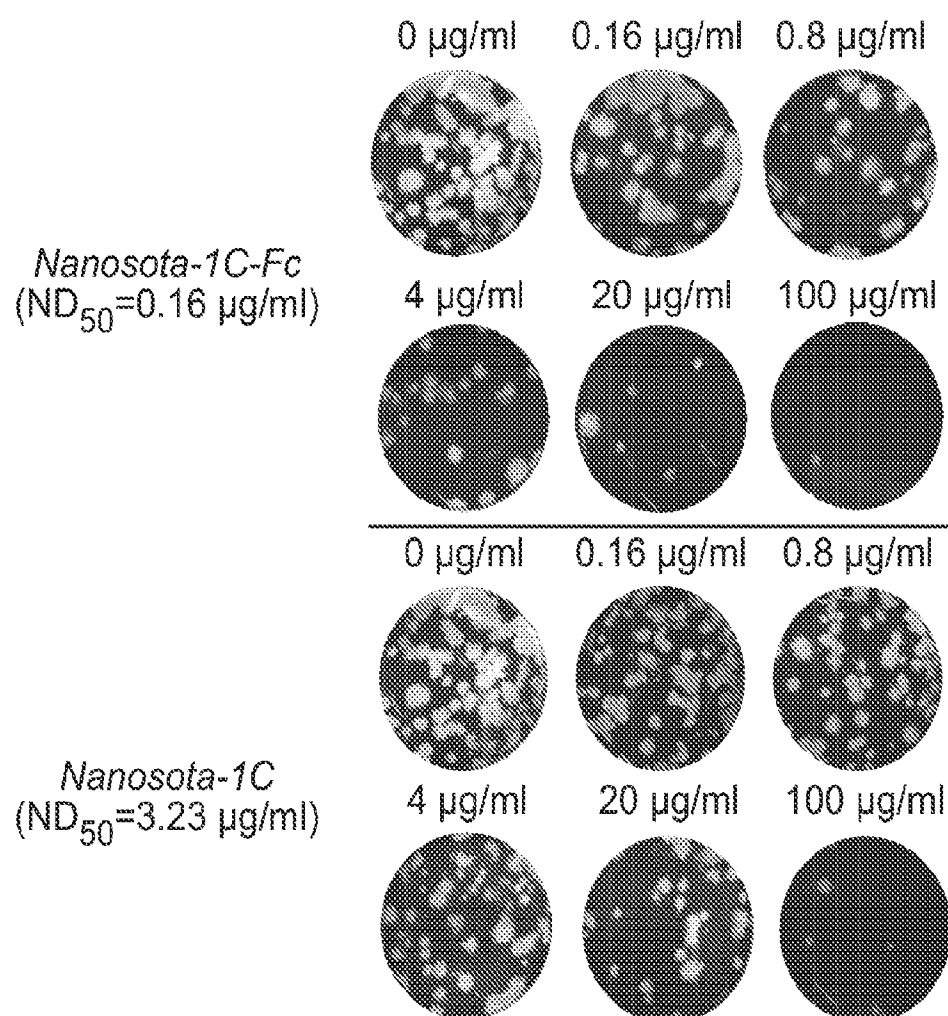


FIGURE 20

25/26

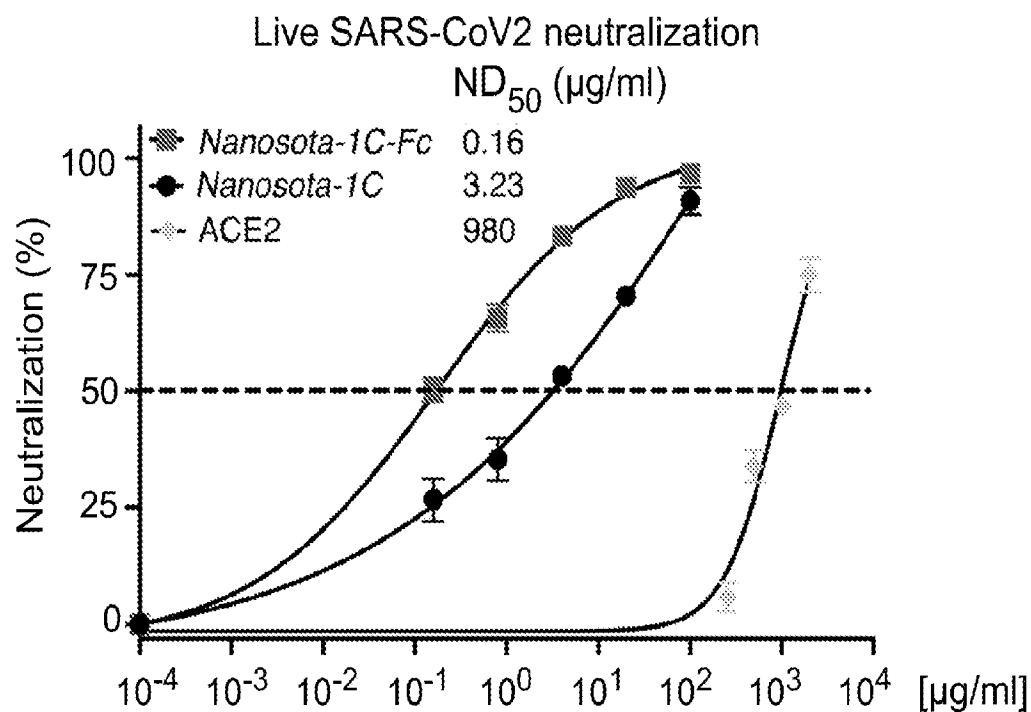


FIGURE 21A

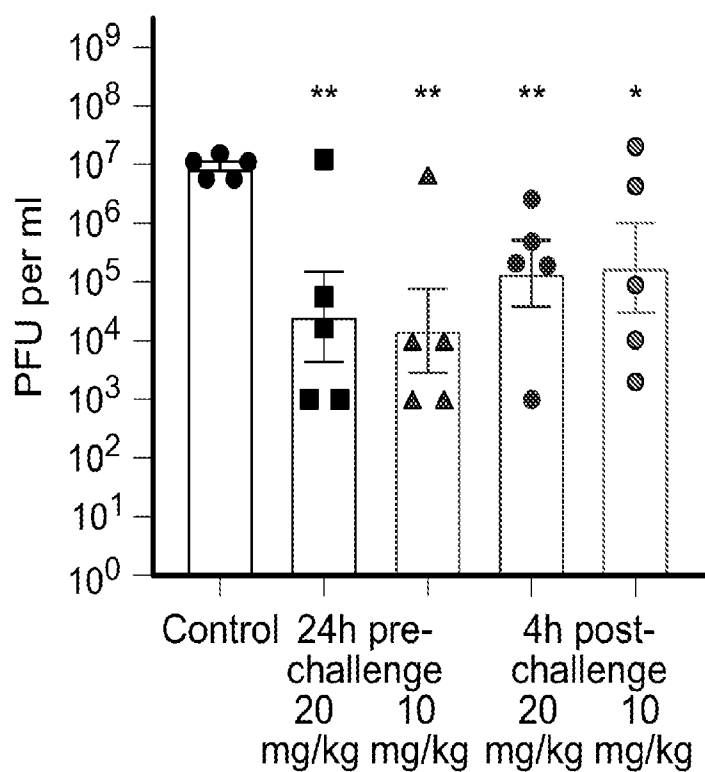


FIGURE 22A

Histology: lungs (mice)

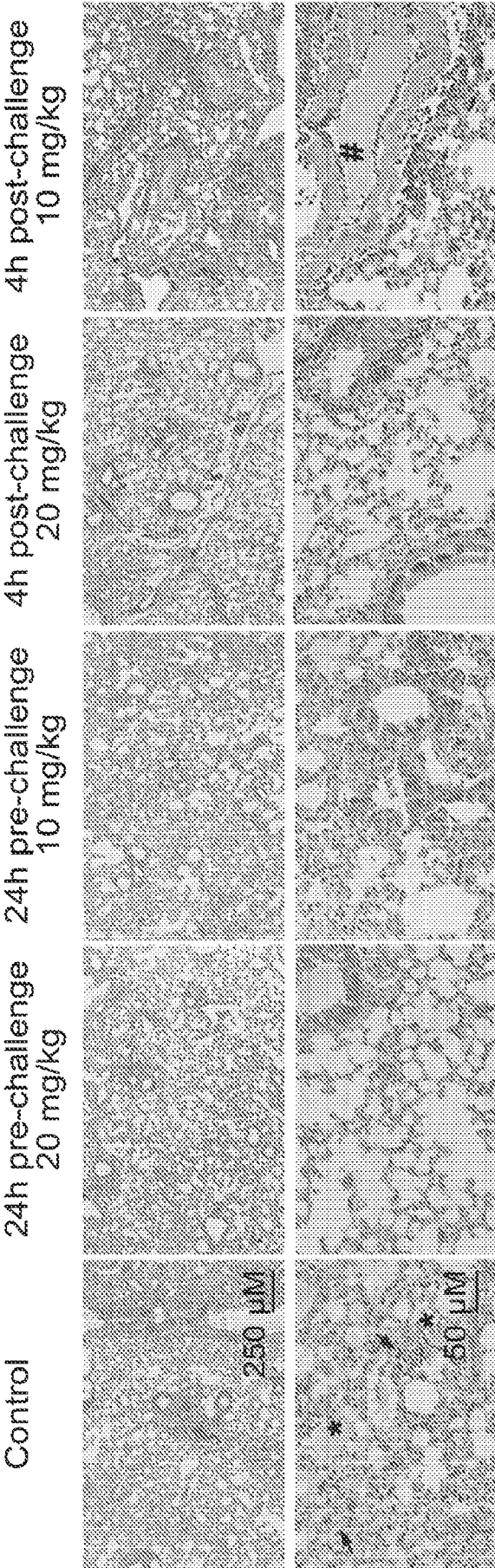


FIGURE 22B

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US21/41328

A. CLASSIFICATION OF SUBJECT MATTER

IPC - C07K 16/10; A61K 39/42; A61P 31/14 (2021.01)

CPC - C07K 16/10; A61K 39/42; A61P 31/14; C07K 2317/569; C07K 2317/52

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

See Search History document

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

See Search History document

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2018/027124 A1 (MEDIMMUNE, LLC) 8 February 2018; paragraph [0031]	1-20
A	US 2014/0079691 A1 (ANAPTYSBIO, INC.) 20 March 2014; Claim 1	1-20
A	WO 2017/117331 A1 (THE BRIGHAM AND WOMEN'S HOSPITAL, INC., et al.) 6 July 2017; paragraph [0012]	1-20
P,X	(YE, G et al.) Chain C, nanobody. Genbank entry (online). National Center for Biotechnology Information. 2 November 2020 [retrieved on 21 October 2021]. Retrieved from the Internet: [URL: https://www.ncbi.nlm.nih.gov/protein/7KM5_C]	1-20

☐ Further documents are listed in the continuation of Box C.☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"D" document cited by the applicant in the international application

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

21 October 2021 (21.10.2021)

Date of mailing of the international search report

NOV 23 2021

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents

P.O. Box 1450, Alexandria, Virginia 22313-1450

Facsimile No. 571-273-8300

Authorized officer

Shane Thomas

Telephone No. PCT Helpdesk: 571-272-4300

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US21/41328

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
- a. ☐ forming part of the international application as filed:
 - ☐ in the form of an Annex C/ST.25 text file.
 - ☐ on paper or in the form of an image file.
 - b. ☐ furnished together with the international application under PCT Rule 13~~ter~~.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
 - c. ☒ furnished subsequent to the international filing date for the purposes of international search only:
 - ☒ in the form of an Annex C/ST.25 text file (Rule 13~~ter~~.1(a)).
 - ☐ on paper or in the form of an image file (Rule 13~~ter~~.1(b) and Administrative Instructions, Section 713).
2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.
3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US21/41328

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claims Nos.: 21-50
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.