Versatile and Multivalent Nanobodies Efficiently Neutralize SARS-CoV-2

Yufei Xiang[1], Sham Nambulli # [2,3], Zhengyun Xiao # [1], Heng Liu # [5], Zhe Sang [1,4], W. Paul Duprex [2,3], Dina Schneidman-Duhovny [6], Cheng Zhang [5], and Yi Shi [1,4]

1. Department of Cell Biology
2. Center for Vaccine Research
3. Department of Microbiology and Molecular Genetics, University of Pittsburgh
4. Pitt/CMU Program for Computational Biology
5. Department of Pharmacology and Chemical Biology
   University of Pittsburgh, Pittsburgh, PA, USA
6. School of Computer Science and Engineering, Institute of Life Sciences, The Hebrew University of Jerusalem, Israel

# Equal contributions
Correspondence: Dina Schneidman-Duhovny (dina.schneidman@mail.huji.ac.il), Cheng Zhang (chengzh@pitt.edu) or Yi Shi (yi.shi@pitt.edu)

Abstract

Cost-effective, efficacious therapeutics are urgently needed against the COVID-19 pandemic. Here, we used camelid immunization and proteomics to identify a large repertoire of highly potent neutralizing nanobodies (Nbs) to the SARS-CoV-2 spike (S) protein receptor-binding domain (RBD). We discovered Nbs with picomolar to femtomolar affinities that inhibit viral infection at sub-ng/ml concentration and determined a structure of one of the most potent in complex with RBD. Structural proteomics and integrative modeling revealed multiple distinct and non-overlapping epitopes and indicated an array of potential neutralization mechanisms. We constructed multivalent Nb constructs that achieved ultrahigh neutralization potency (IC50s as low as 0.058 ng/ml) and may prevent mutational escape. These thermostable Nbs can be rapidly produced in bulk from microbes and resist lyophilization, and aerosolization.

Proteomics rapidly identified a large repertoire of anti-SARS-CoV-2 nanobodies, enabling ultrapotent and nebulizable forms.

Globally a novel, highly transmissible coronavirus severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (*1, 2*) has infected more than 30 million people and has claimed almost 1,000,000 lives, with the numbers still rising as of September 2020. Despite preventive measures, such as quarantines and lock-downs that help curb viral transmission, the virus rebounds after lifting social restrictions. Safe and effective therapeutics and vaccines remain in dire need.

Like other zoonotic coronaviruses, SARS-CoV-2 expresses a surface spike (S) glycoprotein, which consists of S1 and S2 subunits forming a homotrimeric viral spike to interact with host cells. The interaction is mediated by the S1 receptor-binding domain (RBD), which binds the peptidase domain (PD) of angiotensin-

converting enzyme-2 (hACE2) as a host receptor (*3*). Structural studies have revealed different conformations of the spike (*4, 5*). In the pre-fusion stage, the RBD switches between a closed conformation, and an open conformation for hACE2 interaction. In the post-fusion stage, the S2 undergoes a dramatic conformational change to trigger host membrane fusion (*6*). Investigations into COVID-19 convalescence individuals' sera have led to the identification ofpotent neutralizing antibodies (NAbs) primarily targeting the RBD but also non-RBD epitopes (*7-13*). High-quality NAbs may overcome the risks of Fc-associated antibody-dependent enhancement (ADE) and are promising therapeutic candidates (*14, 15*).

V_HH antibodies or nanobodies (Nbs) are minimal, monomeric antigen-binding domains derived from camelid single-chain antibodies (*16*). Unlike IgG antibodies, Nbs are small (~15 kDa), highly soluble and stable, readily bioengineered into bi/multivalent forms, and are amenable to low-cost, efficient microbial production. Due to their robust physicochemical properties, Nbs can be administered by inhalation, making their use against the respiratory viruses very appealing (*17, 18*). Recently, several SARS-CoV-2 neutralizing Nbs have been identified, by screening SARS-CoV or Middle East respiratory syndrome (MERS) cross-reacting Nbs or using synthetic Nb libraries for RBD binding. However, these synthetic Nbs generally neutralize the virus at µg to sub-µg/ml concentration (*12, 19-22*), which are hundreds of times less potent than the best NAbs, likely due to monovalency and lack of affinity maturation (*23 , 24*). The development of highly potent anti-SARS-CoV-2 Nbs may provide a novel means for versatile, cost-effective therapeutics and point-of-care diagnosis.

To produce high-quality SARS-CoV-2 neutralizing Nbs, we immunized a llama with the recombinant RBD. Compared to the pre-bleed, the post-immunized serum showed potent and specific serologic activities towards RBD binding with a titer of 1.75×10^6 (**Fig S1A**). The serum efficiently neutralized the pseudotyped SARS-CoV-2 at the half-maximal neutralization titer (NT50) of ~310,000 (**Fig S1B**), orders of magnitude higher than the convalescent sera obtained from recovered COVID-19 patients (*7, 8*). To further characterize these activities, we separated the single-chain V_HH antibodies from the IgGs. We confirmed that the single-chain antibodies achieve specific, high-affinity binding to the RBD and possess sub-nM half-maximal inhibitory concentration (IC50 = 509 pM) against the pseudotyped virus (**Fig S1C**).

We identified thousands of high-affinity V_HH Nbs from the RBD-immunized llama serum using a robust proteomic strategy that we have recently developed (*25*) (**Fig S2A**). This repertoire includes ~350 unique CDR3s (complementarity-determining regions). For *E.coli* expression, we selected 109 highly diverse Nb sequences from the repertoire with unique CDR3s to cover various biophysical, structural, and potentially different antiviral properties. 94 Nbs were purified and tested for RBD binding by ELISA, from which we confirmed 71 RBD-specific binders (**Fig S2B-C, Table S1, S4**). Of these RBD-specific binders, 49 Nbs presented high solubility and high-affinity (ELISA IC50 below 30 nM, **Fig 1A**), and were promising candidates for functional characterizations. We used a SARS-CoV-2-GFP pseudovirus neutralization assay to screen and characterize the antiviral activities of these high-affinity Nbs. 94% of the tested Nbs neutralize the pseudotype

virus below 3 μM (**Fig 1B**), with 90% of them below 500 nM. Only 20-40% of high-affinity RBD-specific mAbs identified from patient sera have been reported to possess comparable potency (*7, 8*). Over three quarters (76%) of the Nbs efficiently neutralized the pseudovirus below 50 nM, and 6% had neutralization activities below 0.5 nM. We selected the most potent 18 based on the pseudovirus GFP reporter screen and measured their potency accurately using the pseudovirus-luciferase reporter assay. Finally, we tested the potential of 14 to neutralize the SARS-CoV-2 Munich strain using the PRNT50 assay (*26*). All the Nbs reached 100% neutralization and neutralized the virus in a dose-dependent manner. The IC50s span from single-digit ng/ml to sub- ng/ml, of the 3 most potent Nbs 89, 20, and 21, showed neutralization of 2.1 ng/ml (0.133 nM), 1.6 ng/ml (0.102 nM), and 0.7 ng/ml (0.045 nM), respectively, in the pseudovirus assay (**Fig 1C**) and 0.154 nM, 0.048 nM, and 0.022 nM, in the SARS-CoV-2 assay (**Fig 1D-E**). Overall, there was an excellent correlation between the two neutralization assays (R^2= 0.92, **Fig S3**).

We measured the binding kinetics of Nbs 89, 20, and 21 by surface plasmon resonance (SPR) (**Fig S4A-B**). Nbs 89 and 20 have an affinity of 108 pM and 10.4 pM, and the most potent Nb21 did not show detectable dissociation from the RBD during 20 min SPR analysis. The sub-picomolar affinity of Nb21 potentially explains its unusual neutralization potency (**Fig 1F**). We determined the thermostability of Nbs 89, 20, and 21 from the *E.coli* periplasmic preparations to be 65.9, 71.8, and 72.8°C, respectively (**Fig S4C**). Finally, we tested the on-shelf stability of Nb21, which remained soluble after ~ 6 weeks of storage at room temperature after purification. No multimeric forms or aggregations were detected by size-exclusion chromatography (SEC) (**Fig S4D**). Together these results suggest that these neutralizing Nbs have excelelnt physicochemical properties for advanced therapeutic applications.

We employed an integrative approach by SEC, cross-linking/mass spectrometry, and structural modeling for epitope mapping.(*27-30*). First, we performed SEC experiments to distinguish between Nbs that share the same RBD epitope as Nb21 and those that bind to non-overlapping epitopes. Nbs 9, 16, 17, 20, 64, 82, 89, 99, and 107 competed with Nb21 for RBD binding based on SEC profiles (**Fig 2A, S5**), indicating that their epitopes significantly overlap. In contrast, higher mass species (from early elution volumes) corresponding to the trimeric complexes composed of Nb21, RBD, and one of the Nbs (34, 36, 93, 105, and 95) were evident (**Fig 2B, S6A-H**). Moreover, Nb105 competed with Nb34 and Nb95, which did not compete for RBD interaction, suggesting the presence of two distinct and non-overlapping epitopes. Second, we cross-linked Nb-RBD complexes by DSS (disuccinimidyl suberate) and identified on average, four intermolecular cross-links by MS for Nbs 20, 93, 34, 95, and 105. The cross-links were used to map the RBD epitopes derived from the SEC data (**Methods**). Our cross-linking models identified five epitopes (I, II, III, IV, and V corresponding to Nbs 20, 93, 34, 95, and 105) (**Fig 2C**). The models satisfied 90% of the cross-links with an average precision of 7.8 Å (**Fig 2D**, **Table S2**). Our analysis confirmed the presence of a dominant Epitope I (e.g., epitopes of Nbs 20 and 21) overlapping with the hACE2 binding site. Epitope II also co-localized with the non-conserved hACE2 binding site. Both epitopes I and II Nbs can compete with hACE2 binding to RBD at very low concentrations *in vitro* (**Fig S7A**). Epitopes III-V co-localized with conserved sites **(Fig S7B-C**).

Interestingly, epitope I Nbs had significantly shorter CDR3 (four amino acids shorter, p = 0.005) than other epitope binders (**Fig S6I**). Despite this, the vast majority of the selected Nbs potently inhibited the virus with an IC50 below 30 ng/ml (2 nM) (**Table S1**).

To explore the molecular mechanisms that underlie the potent neutralization activities of Epitope I Nbs, we determined a crystal structure of the RBD-Nb20 complex at a resolution of 3.3 Å by molecular replacement (**Methods, Table S3, Fig S13**). Most of the residues in RBD (N334-G526) and the entire Nb20, particularly those at the protein interaction interface, are well resolved. There are two copies of RBD-Nb20 complexes in one asymmetric unit, which are almost identical with an RMSD of 0.277 Å over 287 Cα atoms. In the structure, all three CDRs of Nb20 interact with the RBD by binding to its large extended external loop with two short β-strands (**Fig 3A**) (*31*). E484 of RBD forms hydrogen bonding and ionic interactions with the side chains of R31 (CDR1) and Y104 (CDR3) of Nb20, while Q493 of RBD forms hydrogen bonds with the main chain carbonyl of A29 (CDR1) and the side chain of R97 (CDR3) of Nb20. These interactions constitute a major polar interaction network at the RBD and Nb20 interface. R31 of Nb20 also engages in a cation-π interaction with the side chain of F490 of the RBD (**Fig 3B**). In addition, M55 from the CDR2 of Nb20 packs against residues L452, F490, and L492 of RBD to form hydrophobic interactions at the interface. Another small patch of hydrophobic interactions is formed among residues V483 of RBD and F45 and L59 from the framework β-sheet of Nb20 (**Fig 3C**).

The binding mode of Nb20 to RBD is distinct from other reported neutralizing Nbs, which generally recognize similar epitopes in the RBD external loop region (*32-34*) (**Fig S8**). The extensive hydrophobic and polar interactions (**Fig 3B-C**) between RBD and Nb20 stem from the remarkable shape complementarity (**Fig 3D**) between the CDRs and the external RBD loop, leading to ultrahigh-affinity (~ 10 pM). We further modeled the structure of the best neutralizer Nb21 with RBD based on our crystal structure (**Methods**). Only four residues vary between Nb20 and Nb21 (**Fig S9A**), all of which are on CDRs. Two substitutions are at the RBD binding interface. S52 and M55 in the CDR2 of Nb20 are replaced by two asparagine residues N52 and N55 in Nb21. In our superimposed structure, N52 forms a new H-bond with N450 of RBD (**Fig S9B**). While N55 does not engage in additional interactions with RBD, it creates a salt bridge with the side chain of R31, which stabilizes the polar interaction network among R31 and Y104 of Nb21 and Q484 of RBD (**Fig S9B**). All of those likely contribute to a slower off-rate of Nb21 (**Fig 1F, S4A**) and stronger neutralization potency. Structural comparison of RBD-Nb20/21 and RBD-hACE2 (PDB 6LZG) (*31*) clearly showed that the interfaces for Nb20/21 and hACE2 partially overlap (**Fig 3d, S9C**). Notably, the CDR1 and CDR3 of Nb20/21 would clash with the first helix of hACE2, the primary binding site for RBD (**Fig S9D**).

To understand the antiviral efficacy of our Nbs, we superimposed RBD-Nb complexes to different spike conformations based on cryoEM structures. We found that three copies of Nb20/21 can simultaneously bind all three RBDs in their “down” conformations (PDB 6VXX) (*4*) that correspond to the inactive spike (**Fig 4B**). Our analysis indicates a potential mechanism by which Nbs 20 and 21 (Epitope I) lock RBDs in their down conformation with ultrahigh affinity. Combined with the steric interference with hACE2 binding in the RBD open

conformation (**Fig 4A**), these mechanisms may explain the exceptional neutralization potencies of Epitope I Nbs.

Other epitope-binders do not fit into this inactive conformation without steric clashes and appear to use different neutralization strategies (**Fig 4C**). For example, Epitope II: Nb 93 co-localizes with the hACE2 binding site and can bind the spike in the one RBD "up" conformation (**Fig 4D**, PDB 6VSB) (*3*). It may neutralize the virus by blocking the hACE2 binding site. Epitope III and IV Nbs can only bind when two or three RBDs are in their "up" conformations (PDB 6XCN) (*24*) where the epitopes are exposed. In the all RBDs "up" conformation, three copies of Nbs can directly interact with the trimeric spike. Interestingly, through RBD binding, Epitope III: Nb34 can be accommodated on top of the trimer to lock the helices of S2 in the prefusion stage, preventing their large conformational changes for membrane fusion (**Fig 4E**). When superimposed onto the all "up" conformation, Epitope IV: Nb95 is proximal to the rigid NTD of the trimer, presumably restricting the flexibility of the spike domains (**Fig 4F**).

Epitope mapping enabled us to bioengineer homo- and hetero-dimeric and homo-trimeric Nbs. Homodimers/trimers based on Nb20 or Nb21 were designed to increase the antiviral activities through avidity binding to the trimeric spike. Heterodimers pairing Nb21 with Nbs that bind a different epitope were designed to prevent viral escape. The homodimers/trimers used flexible linker sequences of 25 (GS) or 31 (EK) amino acids (**Methods**). The heterodimers used flexible linkers of 12 amino acids.

We found up to ~30 fold improvement for the homotrimeric constructs of $Nb21_3$ (IC50 = 1.3 pM) and $Nb20_3$ (IC50 = 4.1 pM) compared to the respective monomeric form by the pseudovirus luciferase assay (**Fig 1C, 1E, 5A, 5C**). Similar results were obtained from the SARS-CoV-2 PRNT (**Fig 5B-C, S11A)**. The improvements are likely greater than these values indicate, as the measured values may reflect the assay's lower detection limits. For the heterodimeric constructs, up to a 4-fold increase of potency (i.e., Nb21-Nb34) was observed. The multivalent constructs retained similar physicochemical properties to the monomeric Nbs, including high solubility, yield, thermostability, and remained intact (non-proteolyzed) under the neutralization assay condition (**Fig S10**). They remained highly potent for pseudovirus neutralization after lyophilization and aerosolization (**Methods**, **Fig S11B-G**), indicating the outstanding stability and potential flexibility of administration. The majority of the RBD mutations observed in GISAID (*35*) are very low in frequency (<0.0025) which may increase under Nb selection. Therefore, a cocktail consisting of ultrapotent, multivalent constructs that bind simultaneously a variety of epitopes with potentially different neutralization mechanisms will likely efficiently block virus mutational escape (**Fig 5E, S12**) (*9, 36-38*).

Here, *in vivo* antibody affinity maturation followed by advanced proteomics (*25*) enabled the rapid discovery of a diverse repertoire of high-affinity RBD Nbs, including an ultrapotent neutralizer with sub-picomolar affinity, which is unprecedented for natural, single-domain antibodies. We demonstrated the simplicity and versatility of Nb bioengineering and the outstanding physicochemical properties of the

monomeric Nbs and their multivalent forms. To our knowledge, the multivalent constructs represent the most potent SARS-CoV-2 neutralizers to date. Flexible and efficient administration, such as inhalation may further improve their antiviral efficacy while minimizing the dose, cost, and potential toxicity for clinical applications. The high sequence similarity between Nbs and human IgGs may restrain the immunogenicity (*39*). It is possible to fuse the antiviral Nbs with highly stable, albumin-Nb constructs (*40*) to improve pharmacokinetics. These high-quality Nbs can also be applied as rapid and economic point-of-care diagnostics. We envision that the Nb technology described here will contribute to curbing the current pandemic and possibly a future event.

References

1. N. Zhu *et al.*, A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med* **382**, 727-733 (2020).
2. P. Zhou *et al.*, A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270-273 (2020).
3. D. Wrapp *et al.*, Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**, 1260-1263 (2020).
4. A. C. Walls *et al.*, Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **181**, 281-292 e286 (2020).
5. Y. Cai *et al.*, Distinct conformational states of SARS-CoV-2 spike protein. *Science*, (2020).
6. X. Fan, D. Cao, L. Kong, X. Zhang, Cryo-EM analysis of the post-fusion structure of the SARS-CoV spike glycoprotein. *Nat Commun* **11**, 3618 (2020).
7. Y. Cao *et al.*, Potent Neutralizing Antibodies against SARS-CoV-2 Identified by High-Throughput Single-Cell Sequencing of Convalescent Patients' B Cells. *Cell* **182**, 73-84 e16 (2020).
8. D. F. Robbiani *et al.*, Convergent antibody responses to SARS-CoV-2 in convalescent individuals. *Nature*, (2020).
9. J. Hansen *et al.*, Studies in humanized mice and convalescent humans yield a SARS-CoV-2 antibody cocktail. *Science*, (2020).
10. L. Liu *et al.*, Potent neutralizing antibodies against multiple epitopes on SARS-CoV-2 spike. *Nature*, (2020).
11. P. J. M. Brouwer *et al.*, Potent neutralizing antibodies from COVID-19 patients define multiple targets of vulnerability. *Science* **369**, 643-650 (2020).
12. D. Wrapp *et al.*, Structural Basis for Potent Neutralization of Betacoronaviruses by Single-Domain Camelid Antibodies. *Cell* **181**, 1436-1441 (2020).
13. A. Z. Wec *et al.*, Broad neutralization of SARS-related viruses by human monoclonal antibodies. *Science* **369**, 731-736 (2020).
14. T. Zohar, G. Alter, Dissecting antibody-mediated protection against SARS-CoV-2. *Nat Rev Immunol* **20**, 392-394 (2020).
15. N. Eroshenko *et al.*, Implications of antibody-dependent enhancement of infection for SARS-CoV-2 countermeasures. *Nat Biotechnol* **38**, 789-791 (2020).
16. S. Muyldermans, Nanobodies: Natural Single-Domain Antibodies. *Annu Rev Biochem* **82**, 775-797 (2013).
17. P. Vanlandschoot *et al.*, Nanobodies(R): new ammunition to battle viruses. *Antiviral Res* **92**, 389-407 (2011).
18. L. Detalle *et al.*, Generation and Characterization of ALX-0171, a Potent Novel Therapeutic Nanobody for the Treatment of Respiratory Syncytial Virus Infection. *Antimicrob Agents Chemother* **60**, 6-13 (2016).
19. R. Konwarh, Nanobodies: Prospects of Expanding the Gamut of Neutralizing Antibodies Against the Novel Coronavirus, SARS-CoV-2. *Front Immunol* **11**, 1531 (2020).
20. T. F. Custódio *et al.*, https://www.biorxiv.org/content/10.1101/2020.06.23.165415v1 (2020).
21. C. Liu *et al.*, https://www.biorxiv.org/content/10.1101/2020.03.02.972927v1 (2020).
22. J. Gai *et al.*, https://www.biorxiv.org/content/10.1101/2020.08.09.242867v1 (2020).
23. C. J. Bracken *et al.*, https://www.biorxiv.org/content/10.1101/2020.08.08.242511v1 (2020).

24. C. O. Barnes *et al.*, Structures of Human Antibodies Bound to SARS-CoV-2 Spike Reveal Common Epitopes and Recurrent Features of Antibodies. *Cell*, (2020).
25. Y. Xiang *et al.*, https://www.biorxiv.org/content/10.1101/2020.08.21.261917v1 (2020).
26. W. B. Klimstra *et al.*, SARS-CoV-2 growth, furin-cleavage-site adaptation and neutralization using serum from acutely infected hospitalized COVID-19 patients. *J Gen Virol*, (2020).
27. M. P. Rout, A. Sali, Principles for Integrative Structural Biology Studies. *Cell* **177**, 1384-1403 (2019).
28. C. Yu *et al.*, Charting protein-protein interactions using Cross-Linking Mass Spectrometry (XL-MS). *Abstr Pap Am Chem S* **253**, (2017).
29. A. Leitner, M. Faini, F. Stengel, R. Aebersold, Crosslinking and Mass Spectrometry: An Integrated Technology to Understand the Structure and Function of Molecular Machines. *Trends Biochem Sci* **41**, 20-32 (2016).
30. B. T. Chait, M. Cadene, P. D. Olinares, M. P. Rout, Y. Shi, Revealing Higher Order Protein Structure Using Mass Spectrometry. *J Am Soc Mass Spectr* **27**, 952-965 (2016).
31. Q. H. Wang *et al.*, Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell* **181**, 894-+ (2020).
32. T. Li *et al.*, https://www.biorxiv.org/content/10.1101/2020.06.09.143438v1 (2020).
33. J. D. Walter *et al.*, https://www.biorxiv.org/content/10.1101/2020.04.16.045419v1 (2020).
34. J. Huo *et al.*, Neutralizing nanobodies bind SARS-CoV-2 spike RBD and block interaction with ACE2. *Nat Struct Mol Biol*, (2020).
35. Y. L. Shu, J. McCauley, GISAID: Global initiative on sharing all influenza data - from vision to reality. *Eurosurveillance* **22**, 2-4 (2017).
36. A. Baum *et al.*, Antibody cocktail to SARS-CoV-2 spike protein prevents rapid mutational escape seen with individual antibodies. *Science* **369**, 1014-1018 (2020).
37. Y. Bar-On *et al.*, Safety and antiviral activity of combination HIV-1 broadly neutralizing antibodies in viremic individuals. *Nat Med* **24**, 1701-1707 (2018).
38. M. Marovich, J. R. Mascola, M. S. Cohen, Monoclonal Antibodies for Prevention and Treatment of COVID-19. *JAMA*, (2020).
39. I. Jovcevska, S. Muyldermans, The Therapeutic Potential of Nanobodies. *BioDrugs* **34**, 11-26 (2020).
40. Z. Shen *et al.*, https://www.biorxiv.org/content/10.1101/2020.08.19.257725v1 (2020).
41. P. C. Fridy *et al.*, A robust pipeline for rapid production of versatile nanobody repertoires. *Nat Methods* **11**, 1253-1260 (2014).
42. J. Dunbar, C. M. Deane, ANARCI: antigen receptor numbering and receptor classification. *Bioinformatics* **32**, 298-300 (2016).
43. S. Kumar, G. Stecher, M. Li, C. Knyaz, K. Tamura, MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol* **35**, 1547-1549 (2018).
44. A. Tareen, J. B. Kinney, Logomaker: beautiful sequence logos in Python. *Bioinformatics* **36**, 2272-2274 (2020).
45. Y. Shi *et al.*, Structural characterization by cross-linking reveals the detailed architecture of a coatomer-related heptameric module from the nuclear pore complex. *Mol Cell Proteomics* **13**, 2927-2943 (2014).
46. Y. Shi *et al.*, A strategy for dissecting the architectures of native macromolecular assemblies. *Nat Methods* **12**, 1135-1138 (2015).
47. Y. Xiang, Z. Shen, Y. Shi, Chemical Cross-Linking and Mass Spectrometric Analysis of the Endogenous Yeast Exosome Complexes. *Methods Mol Biol* **2062**, 383-400 (2020).
48. S. J. Kim *et al.*, Integrative structure and functional anatomy of a nuclear pore complex. *Nature* **555**, 475-482 (2018).
49. A. Sali, T. L. Blundell, Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* **234**, 779-815 (1993).
50. A. Fiser, A. Sali, ModLoop: automated modeling of loops in protein structures. *Bioinformatics* **19**, 2500-2501 (2003).
51. D. Schneidman-Duhovny *et al.*, A method for integrative structure determination of protein-protein complexes. *Bioinformatics* **28**, 3282-3289 (2012).
52. D. Schneidman-Duhovny, H. J. Wolfson, Modeling of Multimolecular Complexes. *Methods Mol Biol* **2112**, 163-174 (2020).

53. G. Q. Dong, H. Fan, D. Schneidman-Duhovny, B. Webb, A. Sali, Optimized atomic statistical potentials: assessment of protein interfaces and loops. *Bioinformatics* **29**, 3158-3166 (2013).
54. Z. Otwinowski, W. Minor, Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**, 307-326 (1997).
55. A. J. McCoy *et al.*, Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674 (2007).
56. P. D. Adams *et al.*, PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221 (2010).
57. P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132 (2004).
58. C. J. Williams *et al.*, MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci* **27**, 293-315 (2018).
59. T. D. Goddard *et al.*, UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci* **27**, 14-25 (2018).
60. F. H. Niesen, H. Berglund, M. Vedadi, The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protoc* **2**, 2212-2221 (2007).

Acknowledgments

We thank the staff at the GM/CA of APS in the Argonne National Laboratory (US) for their assistance with X-ray diffraction data collection. We thank the UPMC genome center for Illumina MiSeq, Zhiyi Wei (Southern University of Science and Technology) for the help with crystal structure determination, and Yang Liu for critical reading of the manuscript. This work was supported by The University of Pittsburgh School of Medicine (Y.S.), a CTSI pilot fund (Y.S.), NIH grant R35GM137905 (Y.S.), The University of Pittsburgh and the Center for Vaccine Research (WPD), NIH grant R35GM128641 (C.Z.), ISF 1466/18 (D.S.), and Israeli Ministry of Science and Technology (D.S.).

Contributions

Y.S. and D.S conceived the study. Y.X. performed most of the experiments. S.N. performed the PRNT SARS-CoV-2 neutralization assay. Z.X. produced the multivalent Nbs and performed thermostability measurements. C.Z. determined the X-ray structure with the help of H. L.. Y.X., Y.S., D.S., C.Z., Z.S., S.N., and P.D. analyzed the data. Y.S. cheerled the study and drafted the manuscript. All authors edited the manuscript.

Competing interests

Y.X. and Y.S. are co-inventors on a provisional patent filed by The University of Pittsburgh covering the Nbs described in this manuscript.

Data availability

The coordinates and structure factors for SARS-CoV-2 RBD with Nb20 have been deposited in the Protein Data Bank under the accession codes PDB 7JVB.

The proteomics data of chemical crosslink and mass spectrometric analysis (CX-MS) analysis has been deposited into the MassIVE data repository and the accession code is MSV000086198. Reviewers can download the data using the following account:
Username: MSV000086198_reviewer; Password: SARS2

Supplementary Materials

Materials and Methods

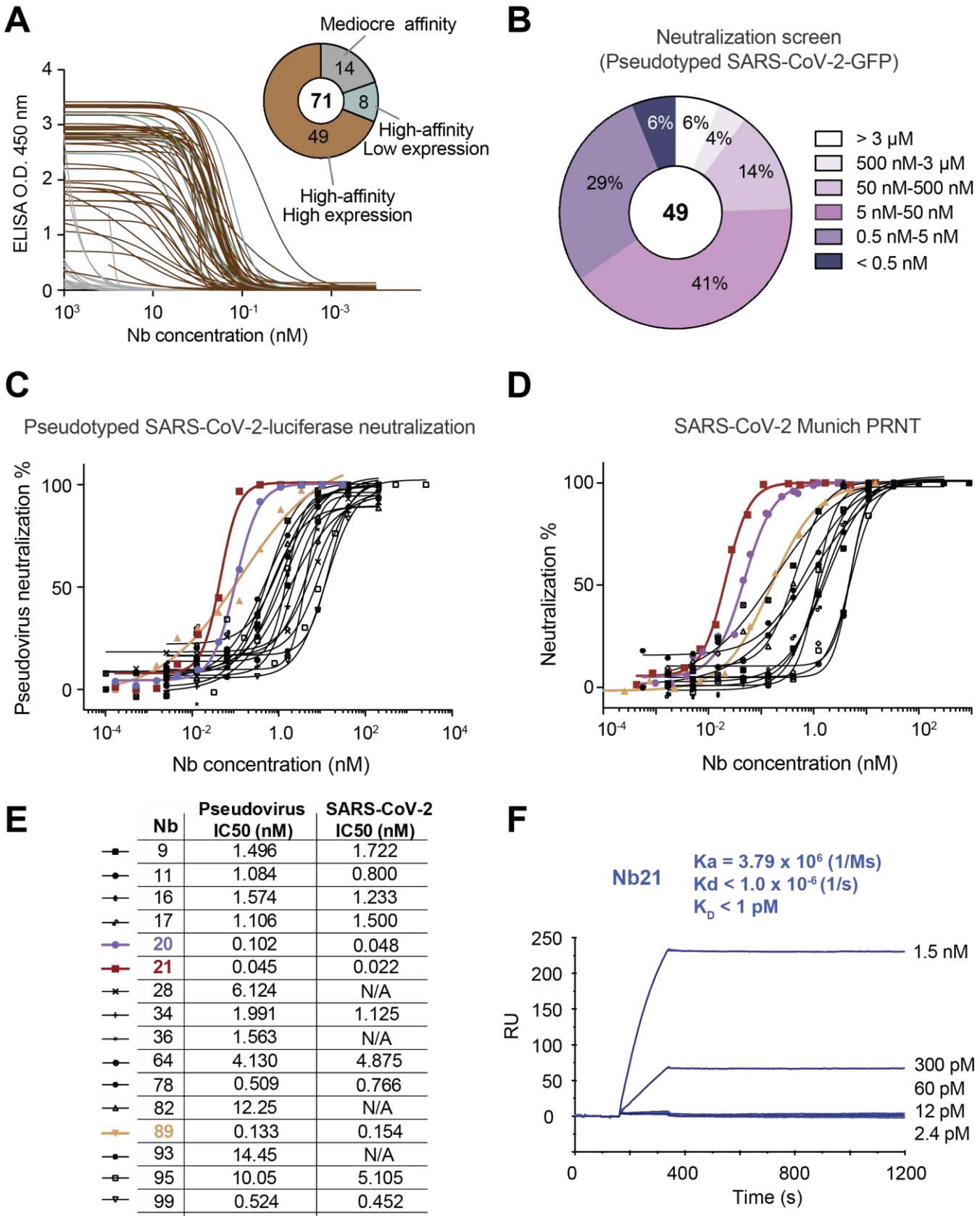
Figures S1 to S13

Tables S1 to S4

References (*41*–*60*)

Figure Legends

Figure 1



E

Nb	Pseudovirus IC50 (nM)	SARS-CoV-2 IC50 (nM)
9	1.496	1.722
11	1.084	0.800
16	1.574	1.233
17	1.106	1.500
20	0.102	0.048
21	0.045	0.022
28	6.124	N/A
34	1.991	1.125
36	1.563	N/A
64	4.130	4.875
78	0.509	0.766
82	12.25	N/A
89	0.133	0.154
93	14.45	N/A
95	10.05	5.105
99	0.524	0.452
105	5.105	5.070
107	0.569	0.181

Figure 1. Production and characterizations of high-affinity RBD Nbs for SARS-CoV-2 neutralization

1A: The binding affinities of 71 Nbs towards RBD by ELISA. The pie chart shows the number of Nbs according to affinity and solubility.

1B: Screening of 49 high-affinity Nbs with high-expression level by SARS-CoV-2-GFP pseudovirus neutralization assay. n = 1 for Nbs with neutralization potency IC50 <= 50 nM, n = 2 for Nbs with neutralization potency IC50 > 50 nM.

1C: The neutralization potency of 18 highly potent Nbs was calculated based on the pseudotyped SARS-CoV-2 neutralization assay (luciferase). Purple, red, and yellow lines denote Nbs 20, 21, and 89 with IC50 < 0.2 nM. Two different purifications of the pseudovirus were used. The average neutralization percentage was shown for each data point (n = 5 for Nbs 20, 21; n = 2 for all other Nbs).

1D: The neutralization potency of 14 neutralizing Nbs by SARS-CoV-2 plaque reduction neutralization test (PRNT). The average neutralization percentage was shown for each data point (n = 4 for Nbs 20, 21, and 89; n = 2 for other Nbs).

1E. A table summary of pseudovirus and SARS-CoV-2 neutralization potencies of 18 Nbs. N/A: not tested.

1F. The SPR binding kinetics measurement of Nb21.

Figure 2

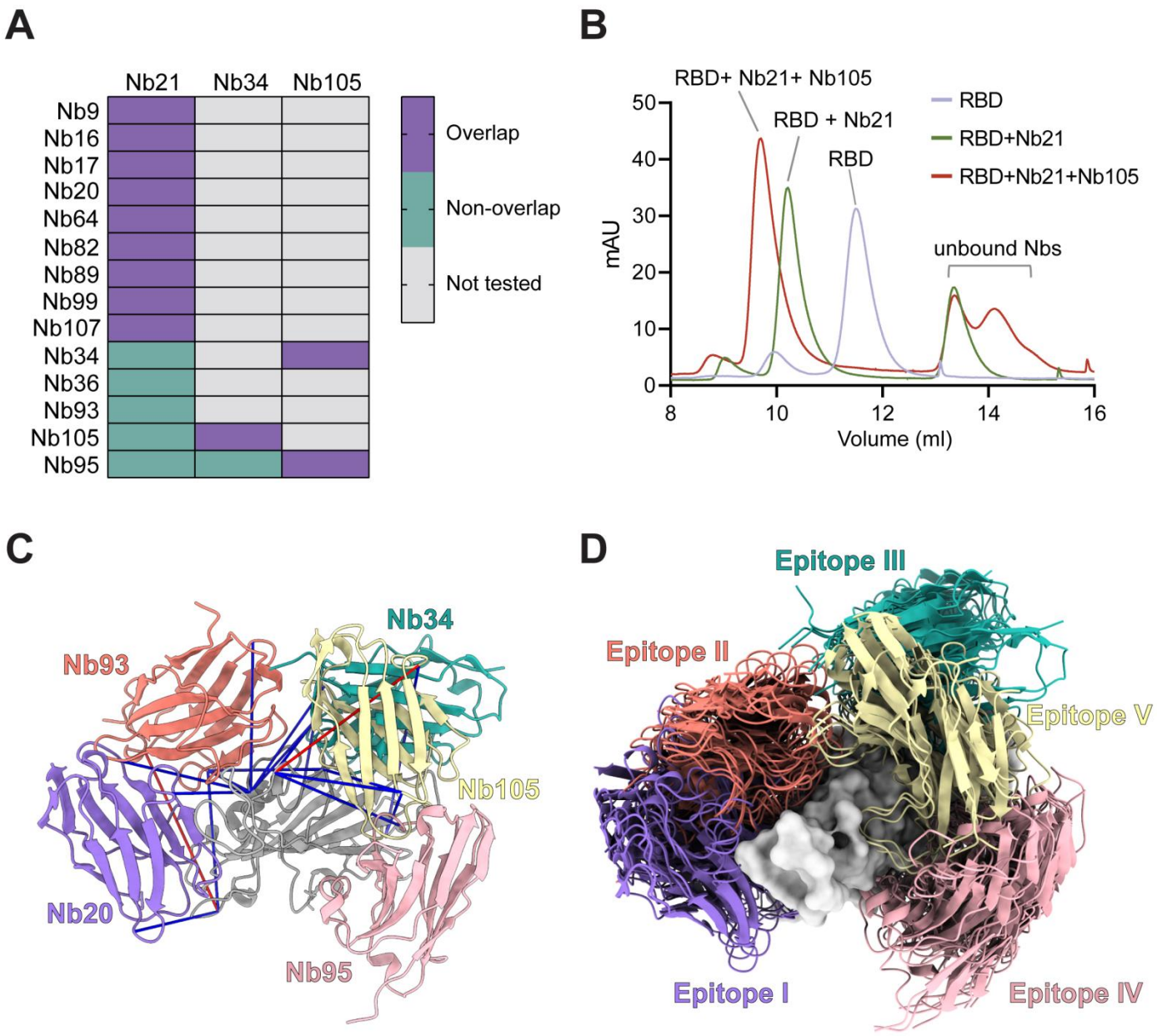


Figure 2. Nb epitope mapping by integrative structural proteomics

2A: A summary of Nb epitopes based on size exclusion chromatography (SEC) analysis. Light salmon color: Nbs that bind the same RBD epitope. Sea green: Nbs of different epitopes.

2B: A representation of SEC profiling of RBD, RBD-Nb21 complex, and RBD-Nb21-Nb105 complex. The y-axis represents UV 280 nm absorbance units (mAu).

2C: A cartoon model showing the localization of five Nbs that bind different epitopes: Nb20 (medium purple), Nb34 (light sea green), Nb93 (salmon), Nb105 (pale goldenrod), and Nb95 (light pink) in complex with the RBD (gray). The Blue and red lines represent DSS cross-links shorter or longer than 28Å, respectively.

2D: Top 10 scoring cross-linking based models for each Nb (cartoons) on top of the RBD surface.

Figure 3

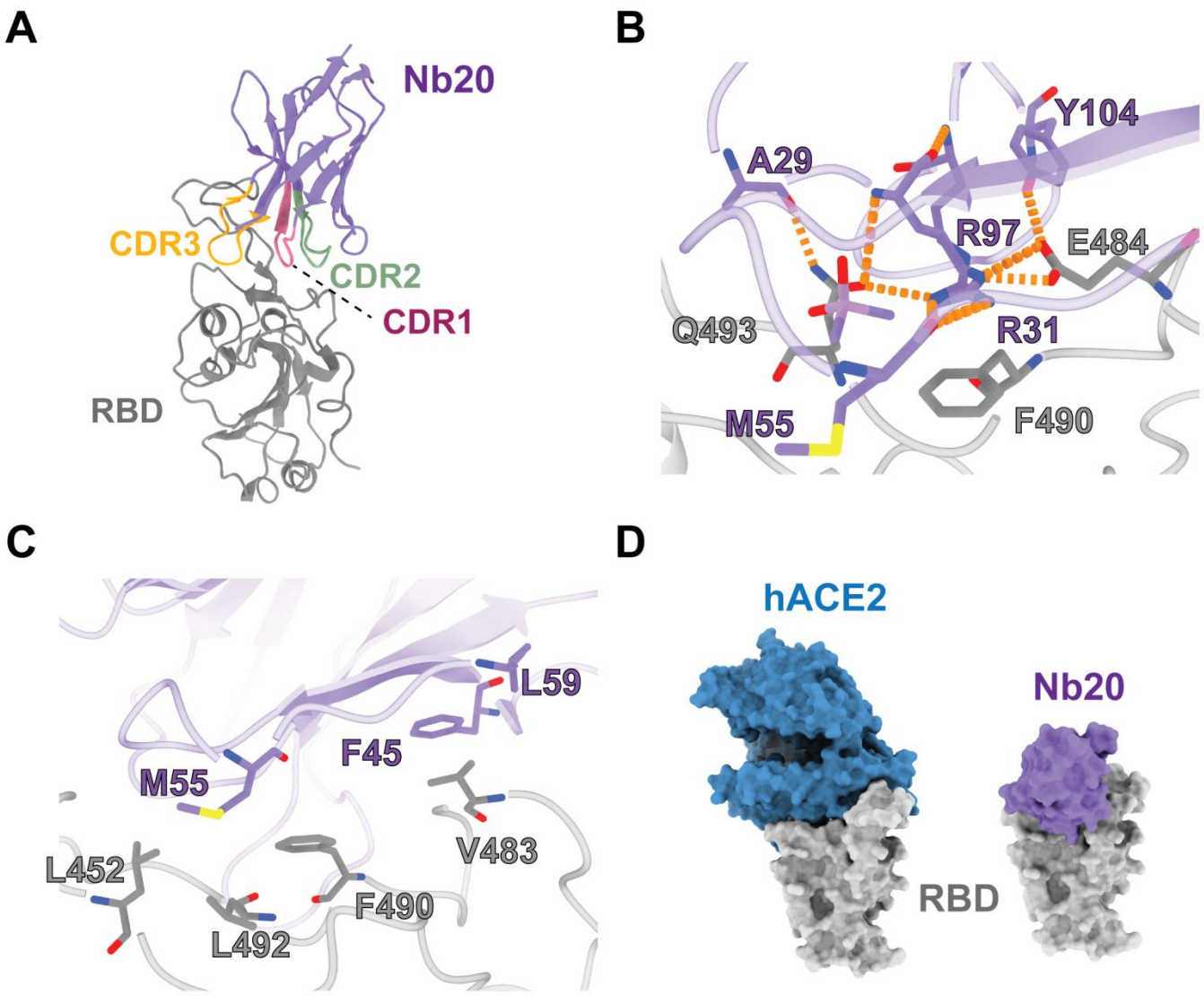


Figure 3. Crystal structure analysis of an ultrahigh affinity Nb in complex with the RBD

3A: Cartoon presentation of Nb20 in complex with the RBD. CDR1, 2, and 3 are in red, green, and orange, respectively.

3B: Zoomed-in view of an extensive polar interaction network that centers on R35 of Nb20.

3C: Zoomed-in view of hydrophobic interactions.

3D: Surface presentation of the Nb20-RBD and hACE2-RBD complex (PDB: 6M0J)

Figure 4

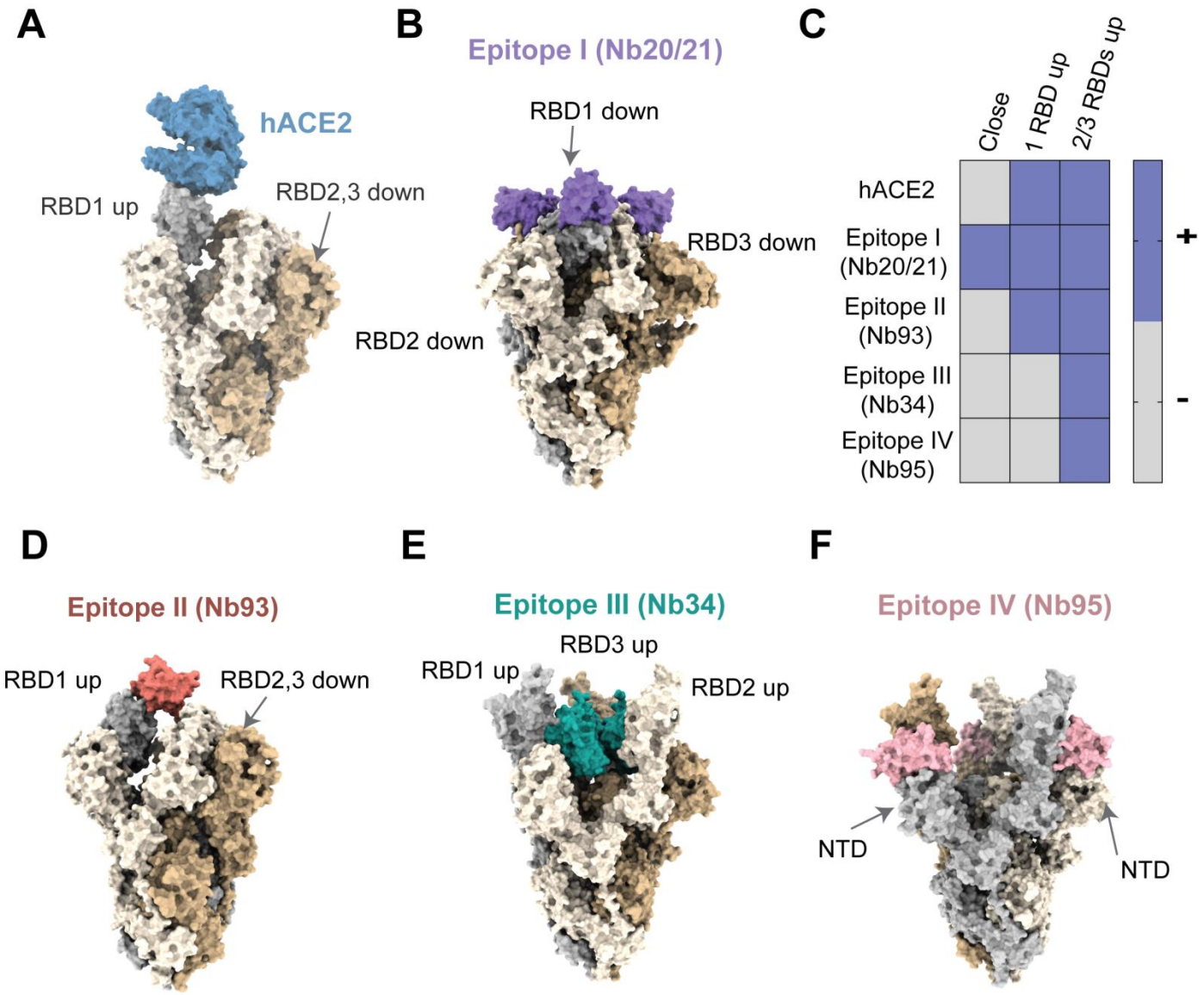


Figure 4. Potential mechanisms of SARS-CoV-2 neutralization by Nbs.

4A: hACE2 (blue) binding to spike trimer conformation (wheat, beige, and gray colors) with one RBD up (PDBs 6VSB, 6LZG).

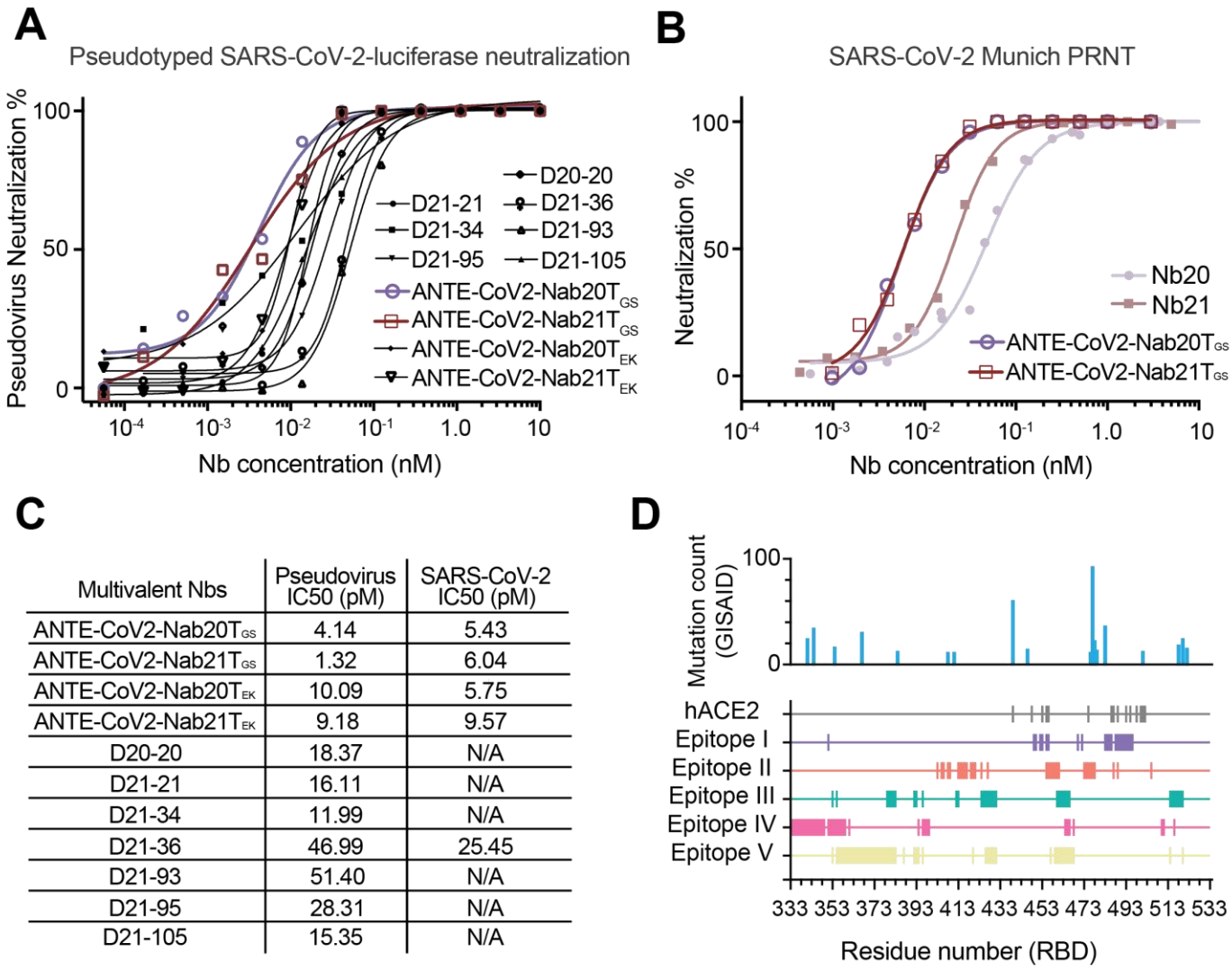
4B: Nb20 (Epitope I, medium purple) partially overlaps with the hACE2 binding site and can bind the closed spike conformation with all RBDs down (PDB 6VXX).

4C: A summary of spike conformations accessible (+) to the Nbs of different epitopes.

4D: Nb93 (Epitope II, salmon) partially overlaps with the hACE2 binding site and can bind to spike conformations with at least one RBD up (PDB 6VSB).

4E-F: Nb34 (Epitope III, light sea blue) and Nb95 (Epitope IV, light pink) do not overlap with the hACE2 binding site and bind to spike conformations with at least two open RBDs (PDB 6XCN).

Figure 5



Multivalent Nbs	Pseudovirus IC50 (pM)	SARS-CoV-2 IC50 (pM)
ANTE-CoV2-Nab20T$_{GS}$	4.14	5.43
ANTE-CoV2-Nab21T$_{GS}$	1.32	6.04
ANTE-CoV2-Nab20T$_{EK}$	10.09	5.75
ANTE-CoV2-Nab21T$_{EK}$	9.18	9.57
D20-20	18.37	N/A
D21-21	16.11	N/A
D21-34	11.99	N/A
D21-36	46.99	25.45
D21-93	51.40	N/A
D21-95	28.31	N/A
D21-105	15.35	N/A

Figure 5. Development of multivalent Nb cocktails for highly efficient SARS-CoV-2 neutralization

5A: Pseudotyped SARS-CoV-2 neutralization assay of multivalent Nbs. The average neutralization percentage of each data point was shown (n = 2). ANTE-CoV2-Nab20T$_{GS/EK}$: homo-trimeric Nb20 with the GS/EK linker; ANTE-CoV2-Nab21T$_{GS/EK}$: homo-trimeric Nb21 with the GS/EK linker.

5B: SARS-CoV-2 PRNT of monomeric and trimeric forms of Nbs 20 and 21. The average neutralization percentage of each data point was shown (n = 2 for the trimers, n=4 for the monomers).

5C: A summary table of the neutralization potency measurements of the multivalent Nbs. N/A: not tested.

5D: Mapping mutations to localization of Nb epitopes on the RBD. The x-axis corresponds to the RBD residue numbers (333 to 533). Rows in different colors represent different epitope residues. Epitope I: 351, 449-450, 452-453, 455-456, 470, 472, 483-486, 488-496; Epitope II: 403, 405-406, 408,409, 413-417, 419-421, 424, 427, 455-461, 473-478, 487, 489, 505; Epitope III: 53, 355, 379-383, 392-393, 396, 412-413, 424-431, 460-466, 514-520; Epitope IV: 333-349, 351-359, 361, 394, 396-399, 464-466, 468, 510-511, 516; Epitope V: 353, 355-383, 387, 392-394, 396, 420, 426-431, 457,459-468, 514, 520.

Supplementary Materials

Materials and Methods

Figures S1 to S13

Tables S1 to S4

References (41-60)

Materials and methods

Camelid immunization and proteomic identification of high-affinity RBD-Nbs

A male Llama “Wally” was immunized with an RBD-Fc fusion protein (Acro Biosystems, Cat#SPD-c5255) at a primary dose of 0.2 mg (with complete Freund’s adjuvant), followed by three consecutive boosts of 0.1 mg every 2 weeks. ~480 ml blood from the animal was collected 10 days after the final boost. All the above procedures were performed by the Capralogics, Inc. following the IACUC protocol. ~ 1 $x10^9$ peripheral mononuclear cells were isolated using Ficoll gradient (Sigma). The mRNA was purified from the mononuclear cells using an RNeasy kit (Qiagen) and was reverse-transcribed into cDNA by the Maxima™ H Minus cDNA Synthesis kit (Thermo). The VHH genes were PCR amplified, and the P5 and P7 adapters were added with the index before sequencing (*25*). Next-generation sequencing (NGS) of the VHH repertoire was performed by Illumina MiSeq with the 300 bp paired-end model in the UPMC Genome Center.

For proteomic analysis of RBD-specific Nbs, plasma was first purified from the immunized blood by the Ficoll gradient (Sigma). V_HH antibodies were then isolated from the plasma by a two-step purification protocol using protein G and protein A sepharose beads (Marvelgent) (*41*). RBD-specific V_HH antibodies were affinity isolated and subsequentially eluted by either increasing stringency of high pH buffer or salt. All the eluted V_HHs were neutralized and dialyzed into 1x DPBS before the quantitative proteomics analysis. RBD-specific V_HH antibodies were reduced, alkylated and in-solution digested using either trypsin or chymotrypsin (*25*). After proteolysis, the peptide mixtures were desalted by self-packed stage-tips or Sep-Pak C18 columns (Waters) and analyzed with a nano-LC 1200 that is coupled online with a Q Exactive™ HF-X Hybrid Quadrupole Orbitrap™ mass spectrometer (Thermo Fisher). Proteomic analysis was performed as previously described and by using the Augur Llama- a dedicated software that we developed to facilitate reliable identification, label-free quantification, and classification of high-affinity Nbs (*25*). This analysis led to thousands of RBD-specific, high-affinity Nb candidates that belong to ~ 350 unique CDR3 families. From these, we selected 109 Nb sequences with unique CDR3s for DNA synthesis and characterizations.

Nb DNA synthesis and cloning

The monomeric Nb genes and the homotrimeric Nbs 20 and 21 with the $(GGGGS)_5$ linkers were codon-optimized and synthesized (Synbio). All the Nb DNA sequences were cloned into a pET-21b(+) vector using EcoRI and HindIII restriction sites. The monomeric Nbs 20, 21, and 89, as well as the homotrimeric Nbs 20 and 21, were also cloned into a pET-22b(+) vector at the BamHI and XhoI sites for periplasmic purification. To produce a heterodimeric Nb such as Nb21- Nb34, the DNA fragment of the monomeric Nb34 was first PCR amplified from the pET-21b(+) vector to introduce a linker sequence and two restriction sites of XhoI and HindIII that facilitate cloning. The PCR fragment was then inserted into the Nb21 pET-21b(+)vector to produce the heterodimer Nb21-$(GGGGS)_2$-Nb34. Homodimers or homotrimers with the EK linker were produced in-house. For the in-house generated constructs, a flexible linker sequence of EGKSSGSGSESKSTGGGGSEGKSSGSGSESKST introduced using the following two oligos:

CCGCTCGAGTGCTGCGGCCGCGGTGCTTTTGCTTTCGCCGCTACCGCTGCTTTTACCTTCGCTGCCACC, and
CCCAAGCTTGAAGGTAAAAGCAGCGGTAGCGGCGAAAGCAAAAGCACCGGTGGCGGTGGCAGCGAAGG T (Integrated DNA Technologies).

Briefly, after digestion using XhoI/HindIII restriction sites, the linker was inserted into pET21b(+)_Nb21 or pET21b(+)_Nb20 vector. For simplicity, here we used the Nb21 trimer for illustration of the cloning strategy. To shuffle the second copy of an Nb, we amplified Nb 21 from the pET21b(+) vector and introduced XhoI/NotI restriction sites. After digestion, the XhoI/NotI Nb fragment was inserted into the Nb_linker vector to produce an Nb homodimer. The following DNA oligos were used to facilitate the cloning of the second copy of the linker in which EcoRI and BamHI sites were introduced:

CCGGAATTCGAGGTGCTTTTGCTTTCGCCGCTACCGCTGCTTTTACCTTCGCTGCCACCGCCACCGGTGC ,
and
CGCGGATCCGGGTTCGAGCTCGGAAGGTAAAAGCAGCGGTAGCGGCGAAAGCAAAAGCACCGGTGGCG G (Integrated DNA Technologies).

After digestion, the second linker was then inserted into the homodimeric pET21b(+)_Nb21 vector. The third copy of the Nb21 gene was PCR amplified from the vector and inserted into the homodimeric pET21b(+)_Nb21 construct using BamHI/SacI sites. All the DNA constructs were verified by sanger sequencing.

Expression and purification of proteins
Nb DNA constructs were transformed into BL21(DE3) cells and plated on Agar with 50 µg/ml ampicillin at 37 °C overnight. Cells were cultured in an LB broth to reach an O.D. of ~ 0.5- 0.6 before IPTG (0.5 mM) induction at 16°C overnight. Cells were then harvested, sonicated, and lysed on ice with a lysis buffer (1xPBS, 150 mM NaCl, 0.2% TX-100 with protease inhibitor). After cell lysis, protein extracts were collected by centrifugation at 15,000 X g for 10 mins and the his-tagged Nbs were purified by Cobalt resin and natively eluted by imidazole buffer (Thermo). Eluted Nbs were subsequently dialyzed in a dialysis buffer (e.g., 1x DPBS, pH 7.4). For the periplasmic preparation of Nbs (Nbs 20, 21, and 89 and the homotrimeric constructs), cell pellets were resuspended in the TES buffer (0.1 M Tris-HCl, pH 8.0; 0.25 mM EDTA, pH 8.0; 0.25 M Sucrose) and incubated on ice for 30 min. The supernatants were collected by centrifugation and subsequently dialyzed to DPBS. The resulting Nbs were then purified by Cobalt resin as described above.

The RBD (residues 319-541) of the SARS-Cov-2 S protein was expressed as a secreted protein in *Spodoptera frugiperda* Sf9 cells (Expression Systems) using the Bac-to-bac baculovirus method (Invitrogen). To facilitate protein purification, a FLAG-tag and an 8 × His-tag were fused to its N terminus, and a tobacco etch virus (TEV) protease cleavage site was introduced between the His-tag and RBD. Cells were infected with baculovirus and incubated at 27 °C for 60 h before harvesting. The conditioned media was added with 20 mM Tris pH 7.5 and incubated at RT for 1 h in the presence of 1 mM $NiSO_4$ and 5mM $CaCl_2$. The supernatant was collected by centrifugation at 25,000 g for 30 min and then incubated with Nickel-NTA agarose resin (Clontech) overnight at 4 °C. After washing with buffer containing 20 mM Hepes pH 7.5, 200 mM NaCl, and 50mM imidazole, the RBD protein was eluted with the same buffer containing 400 mM Imidazole. Eluted protein was treated by TEV protease overnight to remove extra tags and further purified by size exclusion chromatography using the Superdex 75 column (Fisher) with a buffer containing 20 mM HEPES pH 7.5 and 150 mM NaCl. To obtain RBD and Nb20 complex, purified RBD was mixed with purified Nb20 in a molar ratio of 1: 1.5 and then incubated on ice for 2 hours. The complex was further purified using the Superdex 75 column with a buffer

containing 20 mM Hepes pH 7.5 and 150 mM NaCl. Purified RBD-Nb20 complex was concentrated to 10-15 mg/ml for crystallization.

ELISA (Enzyme-linked immunosorbent assay)

Indirect ELISA was carried out to measure the relative affinities of Nbs. RBD was coated onto a 96-well ELISA plate (R&D system) at two ng/well in coating buffer (15 mM sodium carbonate, 35 mM sodium bicarbonate, pH 9.6) overnight at 4°C and was blocked with a blocking buffer (DPBS, 0.05% Tween 20, 5% milk) at room temperature for 2 hrs. Nbs were serially 10X diluted in the blocking buffer, starting from 1 µM to 0.1 pM, and 100 µl of each concentration was incubated with RBD-coated plates for 2 hrs. HRP-conjugated secondary antibodies against T7-tag (Thermo, Cat# PA1-31449) were diluted 1:7500 and incubated with the well for 1 hr at room temperature. After PBST (DPBS, 0.05% Tween 20) washes, the samples were further incubated under dark with freshly prepared w3,3′,5,5′-Tetramethylbenzidine (TMB) substrate for 10 mins to develop the signals. After the STOP solution (R&D system), the plates were read at multiple wavelengths (the optical density O.D. at 550 nm wavelength subtracted from the density at 450 nm) on a plate reader (Multiskan GO, Thermo Fisher). A non-binder was defined if any of the following two criteria were met: i) The ELISA signal was under detected at 1 µM concentration. ii) The ELISA signal could only be detected at a concentration of 1 µM and was under detected at 0.1 µM concentration. The raw data was processed by Prism 7 (GraphPad) to fit into a 4PL curve and to calculate logIC50.

Competitive ELISA with hACE2

The COVID-19 spike-ACE2 binding assay kit was purchased from RayBiotech (Cat# CoV-SACE2-1). A 96-well plate was pre-coated with recombinant RBD. Nbs were 10-fold diluted (from 1 µM to 1 pM) in the assay buffer containing a saturating amount of hACE2 and then incubated with the plate at room temperature for 2.5 hrs. The plate was washed by the washing buffer to remove the unbound hACE2. Goat anti-hACE2 antibodies were incubated with the plate for 1 hr at room temperature. HRP-conjugated anti-goat IgG was added to the plate and incubated for an hour. TMB solution was added to react with the HRP conjugates for 0.5 hr. The reaction was then stopped by the Stop Solution. The signal corresponding to the amount of the bound hACE2 was measured by a plate reader at 450 nm. The resulting data were analyzed by Prism 7 (GraphPad) and plotted.

Pseudotyped SARS-CoV-2 neutralization assay

The 293T-hsACE2 stable cell line (Cat# C-HA101, Lot# TA060720C) and the pseudotyped SARS-CoV-2 (Wuhan-Hu-1 strain) particles with GFP (Cat# RVP-701G, Lot#CG-113A) or luciferase (Cat# RVP-701L, Lot# CL109A, and CL-114A) reporters were purchased from the Integral Molecular. The neutralization assay was carried out according to the manufacturers' protocols. In brief, 10-fold serially diluted Nbs were incubated with the pseudotyped SARS-CoV-2-GFP for 1 hr at 37 °C for screening, while 3- or 5-fold serially diluted Nbs / immunized serum / immunized V_HH mixture was incubated with the pseudotyped SARS-CoV-2-luciferase for accurate measurements. At least eight concentrations were tested for each Nb. Pseudovirus in culture media without Nbs was used as a negative control. 100 µl of the mixtures were then incubated with 100 µl 293T-hsACE2 cells at 2.5x10e5 cells/ml in the 96-well plates. The infection took ~72 hrs at 37 °C with 5% CO_2. The GFP signals (ex488/em530) were read using the Tecan Spark 20M with auto-optimal settings, while the luciferase signal was measured using the *Renilla*-Glo luciferase assay system (Promega, Cat# E2720) with the luminometer at 1 ms integration time. The obtained relative fluorescent/luminescence signals (RFU/RLU) from the negative control wells were normalized and used to calculate the neutralization percentage at each concentration. For SARS-CoV-2-GFP screening, the 49 tested Nbs were divided into 6 groups based on their lowest tested concentration of 100% neutralization. For SARS-CoV-2-luciferase, data was processed by Prism7 (GraphPad) to fit into a 4PL curve and to calculate the logIC50 (half-maximal inhibitory concentration).

SARS-CoV-2 Munich plaque reduction neutralization test (PRNT)

Nbs were diluted in a 2- or 3-fold series in Opti-MEM (Thermo). Each Nb dilution (110 µl) was mixed with 110 µl of SARS-CoV-2 (Munich strain) containing 100 plaque-forming units (p.f.u.) of the virus in Opti-MEM. The serum–virus mixes (220 µl total) were incubated at 37 °C for 1 h, after which they were added dropwise onto confluent Vero E6 cell (ATCC® CRL-1586™) monolayers in the six-well plates. After incubation at 37 °C, 5% (v/v) CO_2 for 1 h, 2 ml of 0.1% (w/v) immunodiffusion agarose (MP Biomedicals) in Dulbecco's modified eagle medium (DMEM) (Thermo) with 10% (v/v) FBS and 1× pen-strep was added to each well. The cells were incubated at 37 °C, 5% CO_2 for 72 hrs. The agarose overlay was removed and the cell monolayer was fixed with 1 ml/well formaldehyde (Fisher) for 20 min at room temperature. Fixative was discarded and 1 ml/well of 1% (w/v) crystal violet in 10% (v/v) methanol was added. Plates were incubated at room temperature for 20 min and rinsed thoroughly with water. Plaques were then enumerated and the 50% plaque reduction neutralization titer (PRNT50) was calculated. A validated SARS-CoV-2 antibody-negative human serum control, a validated NIBSC SARS-CoV-2 plasma control, was obtained from the National Institute for Biological Standards and Control, UK) and an uninfected cells control were also performed to ensure that virus neutralization by antibodies was specific.

Surface plasmon resonance (SPR)

Surface plasmon resonance (Biacore 3000, GE-Healthcare) was used to measure Nb affinities. Briefly, recombinant Nb was immobilized to the flow channels of an activated CM5 sensor-chip. Nb was diluted to 10 µg/ml in 10 mM sodium acetate, pH 5.5, and injected into the SPR system at 5 µl/min for 420 s. The surface was then blocked by 1 M ethanolamine-HCl (pH 8.5). For RBD analyte, a series of dilution (spanning ~ 1,000-fold concentration range) was injected in duplicate, with HBS-EP+ running buffer at a flow rate of 30 µl/min for 180 s, followed by 20-min dissociation. Between each injection, the sensor chip surface was regenerated twice with 6M Guanidine-HCl at a flow rate of 40-50 µl/min for 30 s - 1 min. Binding sensorgrams for each Nb were processed and analyzed using BIA evaluation by fitting with the 1:1 Langmuir model.

Phylogenetic tree analysis and sequence logo

Sequences were first aligned and numbered according to Martin's numbering scheme by ANARCI (*42*). The phylogenetic tree was constructed from aligned sequences by Molecular Evolutionary Genetics Analysis (MEGA)(*43*) using the Maximum Likelihood method. The sequence logo was plotted from aligned sequences by logomaker(*44*).

Epitope screening by size exclusion chromatography (SEC)

Recombinant RBD and Nb proteins were mixed at a ratio of 1:1 (w:w) and incubated at 4°C for 1 hr. The complexes were analyzed by the SEC (Superdex75, GE Healthcare) at a low rate of 0.4 ml/min for 1 hr using a running buffer of 20 mM HEPES, 150 mM NaCl, pH 7.5. Protein signals were detected by ultraviolet light absorbance at 280 nm.

Chemical cross-linking and mass spectrometry (CXMS)

Recombinant Nbs were first pre-incubated with the trypsin resin for approximately 2-5 mins to remove the N terminal T7 tag, which is highly reactive to the crosslinker. Nb was incubated with RBD in the pH 7.4 buffer at 4°C for 1 hr to allow the formation of the complex. The reconstituted complexes were then cross-linked with 2 mM disuccinimidyl suberate (DSS, ThermoFisher Scientific) for 25 min at 25°C with gentle agitation. The reaction was then quenched with 50 mM ammonium bicarbonate (ABC) for 10 min at room temperature. After protein reduction and alkylation, the cross-linked samples were separated by a 4–12% SDS-PAGE gel (NuPAGE, Thermo Fisher). The regions corresponding to the monomeric, cross-linked species (~45-50 kDa) were sliced and digested in-gel with trypsin and Lys-C (*45-47*). After efficient proteolysis, the cross-link peptide mixtures were desalted and analyzed with a nano-LC 1200 (Thermo Fisher) coupled to a Q Exactive™ HF-X

Hybrid Quadrupole-Orbitrap™ mass spectrometer (Thermo Fisher). The cross-linked peptides were loaded onto a Picochip column (C18, 3 µm particle size, 300 Å pore size, 50 µm × 10.5 cm; New Objective) and eluted using a 60 min LC gradient: 5% B–10% B, 0 – 2 min; 10% B – 40% B, 2 – 50 min; 40% B–100% B, 50 – 60 min; mobile phase A consisted of 0.1% formic acid (FA), and mobile phase B consisted of 0.1% FA in 80% acetonitrile. The QE HF-X instrument was operated in the data-dependent mode. The top 4 most abundant ions (with the mass range of 350 to 2,000 and the charge state of +4 to +7) were fragmented by high-energy collisional dissociation (normalized HCD energy 30). The target resolution was 120,000 for MS and 15,000 for MS/MS analyses. The quadrupole isolation window was 1.6 Th and the maximum injection time for MS/MS was set at 300 ms. After the MS analysis, the data was searched by pLink for the identification of cross-linked peptides. The mass accuracy was specified as 10 and 20 p.p.m. for MS and MS/MS, respectively. Other search parameters included cysteine carbamidomethylation as a fixed modification and methionine oxidation as a variable modification. A maximum of three trypsin missed-cleavage sites was allowed. Initial search results were obtained using the default 5% false discovery rate, estimated using a target-decoy search strategy. The crosslink spectra were manually checked as previously described (*45, 46, 48, 49*).

Integrative structural modeling

Structural models for Nbs were obtained using a multi-template comparative modeling protocol of MODELLER (*50*). Next, we refined the CDR3 loop (*51*) and selected the top 5 scoring loop conformations for the downstream docking in addition to 5 models from comparative modeling. Each Nb model was then docked to the RBD structure (PDB 6LZG) by an antibody-antigen docking protocol of PatchDock software that focuses the search to the CDRs and optimizes CXMS-based distance restraints satisfaction (*52, 53*). A restraint was considered satisfied if the Ca-Ca distance between two DSS cross-linked residues is within 28Å. The models were then re-scored by a potential statistical SOAP (*54*). The antigen interface residues (distance <6Å from Nb atoms) among the top 10 scoring models, according to the SOAP score, were used to determine the epitopes. The precision was estimated based on the convergence among the ten top-scoring models which was measured as the average RMSD between all the pairs.

Crystallization, data collection, and structure determination of RBD-Nb20 complex

Crystallization trials were performed with the Crystal Gryphon robot (Art Robbins). The RBD-Nb20 complex was crystallized using the sitting-drop vapor diffusion method at 17 °C. The crystals were obtained in conditions containing 100 mM sodium cacodylate pH 6.5 and 1 M sodium citrate. For data collection, the crystals were transferred to the reservoir solution supplemented with 20% glycerol before freezing in liquid nitrogen. X-ray diffraction data were collected at the Advanced Photon Source (APS) beamline 23IDB of GM/CA with a 10µm-diameter microbeam. The data were processed using HKL2000 (*55*). Diffraction data from six crystals were merged to obtain a complete dataset with a resolution of 3.3 Å.

The structure was determined by the molecular replacement method in Phaser (*56*) using the crystal structures of RBD (PDB 6LZG) and an Nb (V_HH-72, PDB 6WAQ) as search models. The initial model was refined in Phenix (*57*)and adjusted in COOT (*58*). The model quality was checked by MolProbity (*59*). The final refinement statistics were listed in Table S3.

Nb21 comparative modeling was done using the Nb20 structure as a template in MODELLER. All structure visualization figures were prepared using UCSF ChimeraX (*60*).

Thermostability analysis of Nbs

Nb thermostability was measured by differential scanning fluorimetry (DSF). To prepare DSF samples, Nbs were mixed with SYPRO orange dye (Invitrogen) in PBS to reach a final concentration of 12 µM or 6 µM. The

samples were analyzed in triplicate using a 7900HT Fast Real-Time PCR System (Applied Biosystems) as previously described (*40*). The melting point was calculated by the first derivatives method (*61*).

Nb stability test

For the stability test, Nb was eluted and collected in the SEC running buffer (20 mM HEPES, 150 mM NaCl, pH 7.5) and then concentrated to 1 ml (1 mg/ml). 0.5 ml of the concentrated Nb was lyophilized by snap freezing in liquid nitrogen before dried in a speed-vac. ddH2O was then used to reconstitute the Nb. The other 0.5 ml was aerosolized by using a portable mesh atomizer nebulizer (MayLuck). No obvious dead volume was observed. The aerosols were collected in a microcentrifuge tube. SEC analysis and pseudovirus neutralization assays were performed as described above.

Supplementary Figure 1

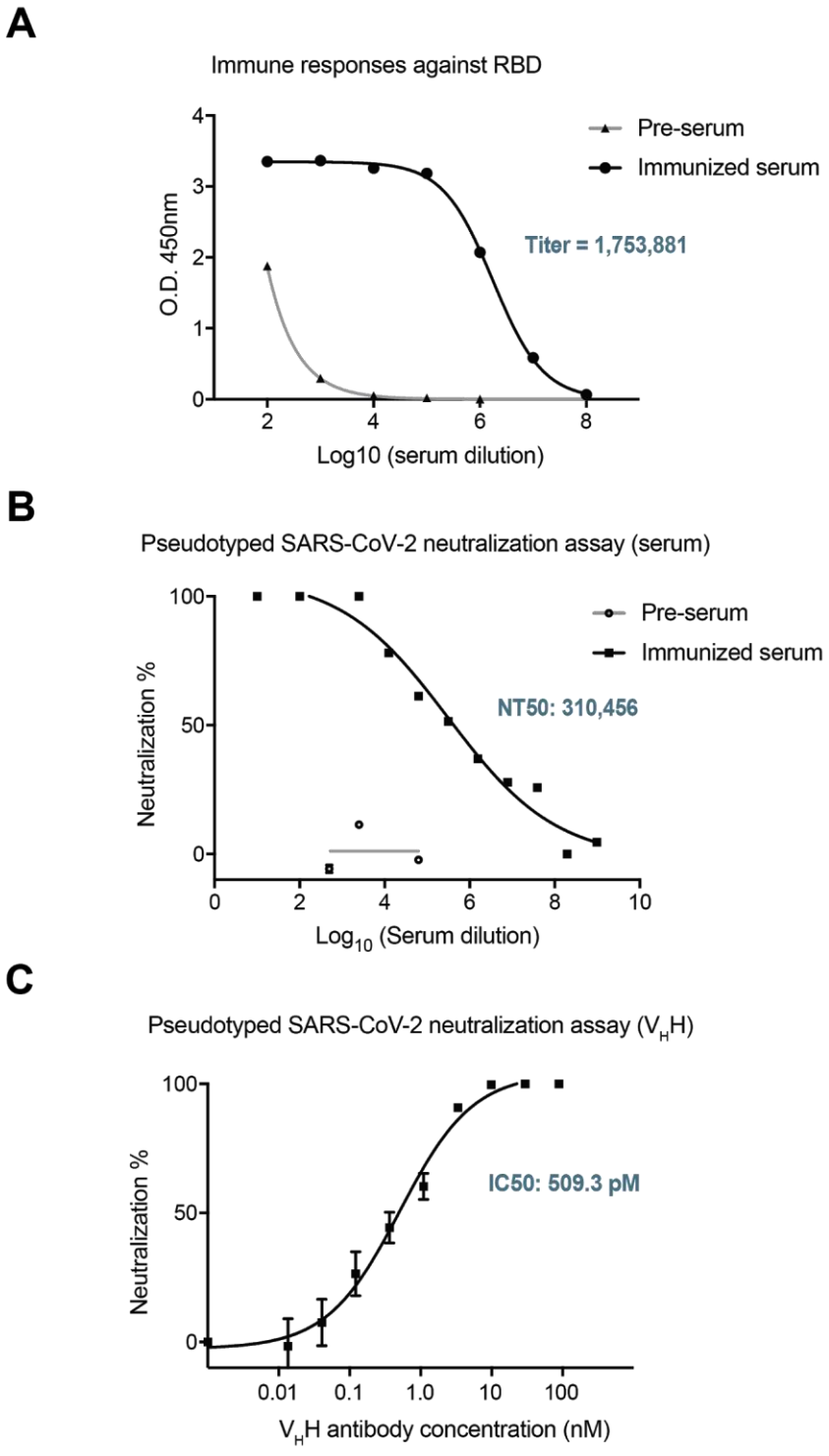


Figure S1. Development of RBD-specific Nbs for potent SARS-CoV-2 neutralization.

S1A: Detection of strong and specific serologic activities after immunization of SARS-CoV-2 $_{RBD.}$

S1B: Neutralization potency of the immunized camelid's serum against pseudotyped SARS-CoV-2 (luciferase). NT50: half-maximal serum neutralization titer.

S1C: Neutralization potency of Nbs against pseudotyped SARS-CoV-2 (luciferase).

Supplementary Figure 2

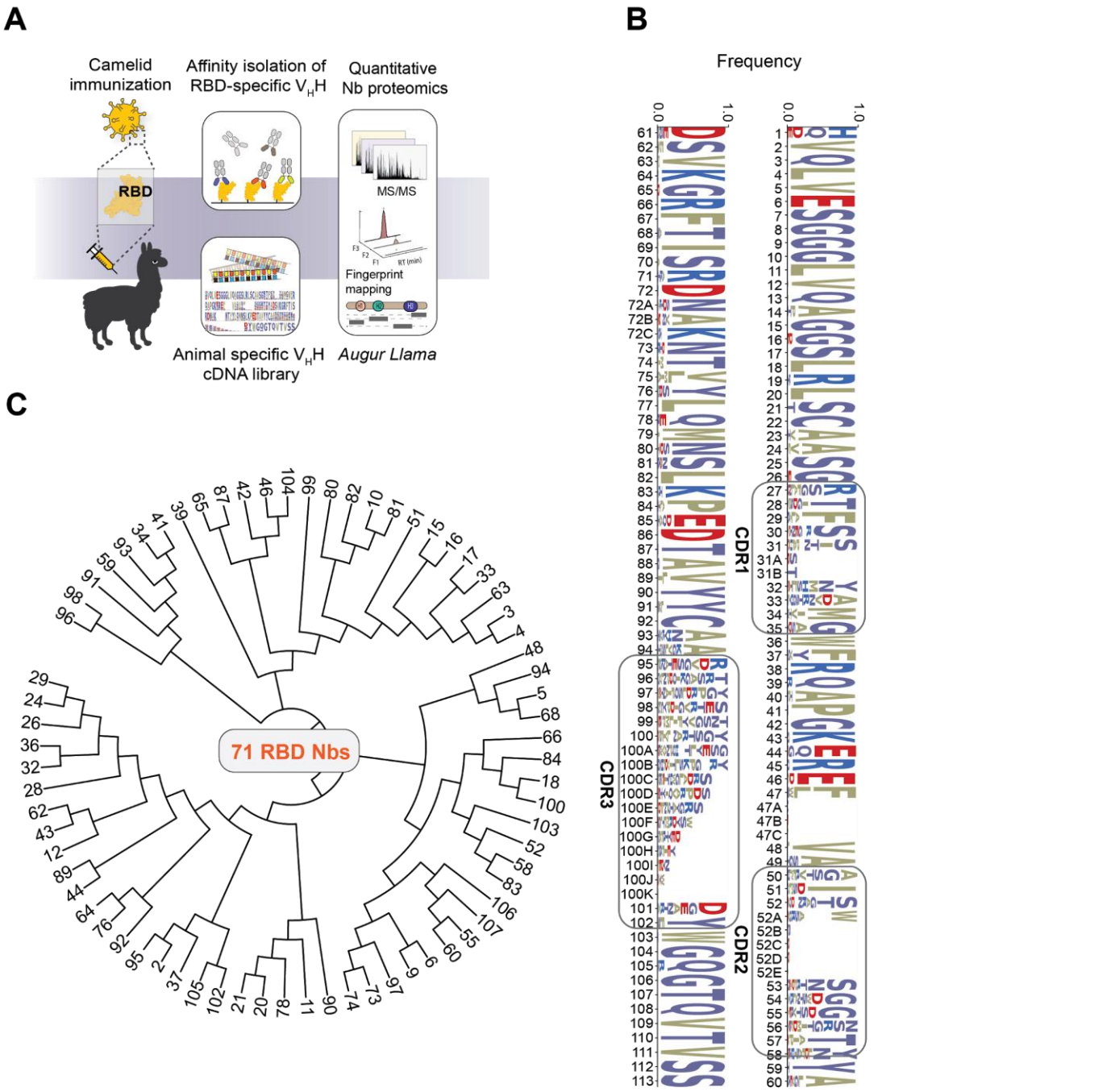


Figure S2. Identification of high-affinity RBD- Nbs.

S2A: High-affinity RBD-specific Nb identification by camelid immunization and quantitative Nb proteomics. Briefly, a camelid was immunized with the RBD protein. After immunization, RBD-specific V_HH antibodies were isolated from the serum and analyzed by quantitative proteomics to identify the high-affinity, RBD-specific Nbs (see Methods). A V_HH (Nb) cDNA library from the plasma B cells of the immunized camelid was generated to facilitate proteomic analysis.

S2B: Sequence logo of 71 RBD Nbs. Each Nb has a unique CDR3 sequence. The amino acid occurrence at each position is shown. CDR: complementarity determining region. FR: framework.

S2C: The phylogenetic tree of the Nbs constructed by the maximum likelihood model.

Supplementary Figure 3

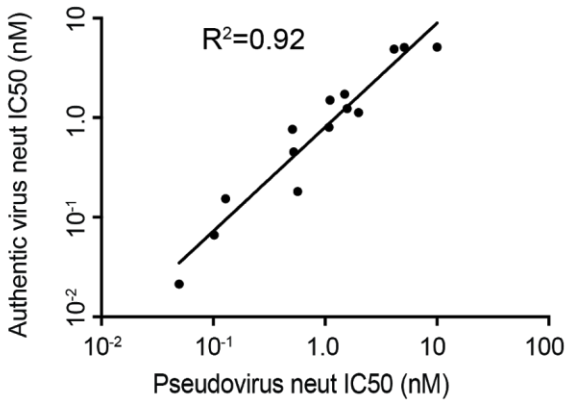


Figure S3. Correlation analysis of 18 highly potent SARS-CoV-2 neutralizing Nbs.
A plot showing a linear correlation of Nb neutralization IC50s between the pseudotyped virus neutralization assay and the SARS-CoV-2 PRNT.

Supplementary Figure 4

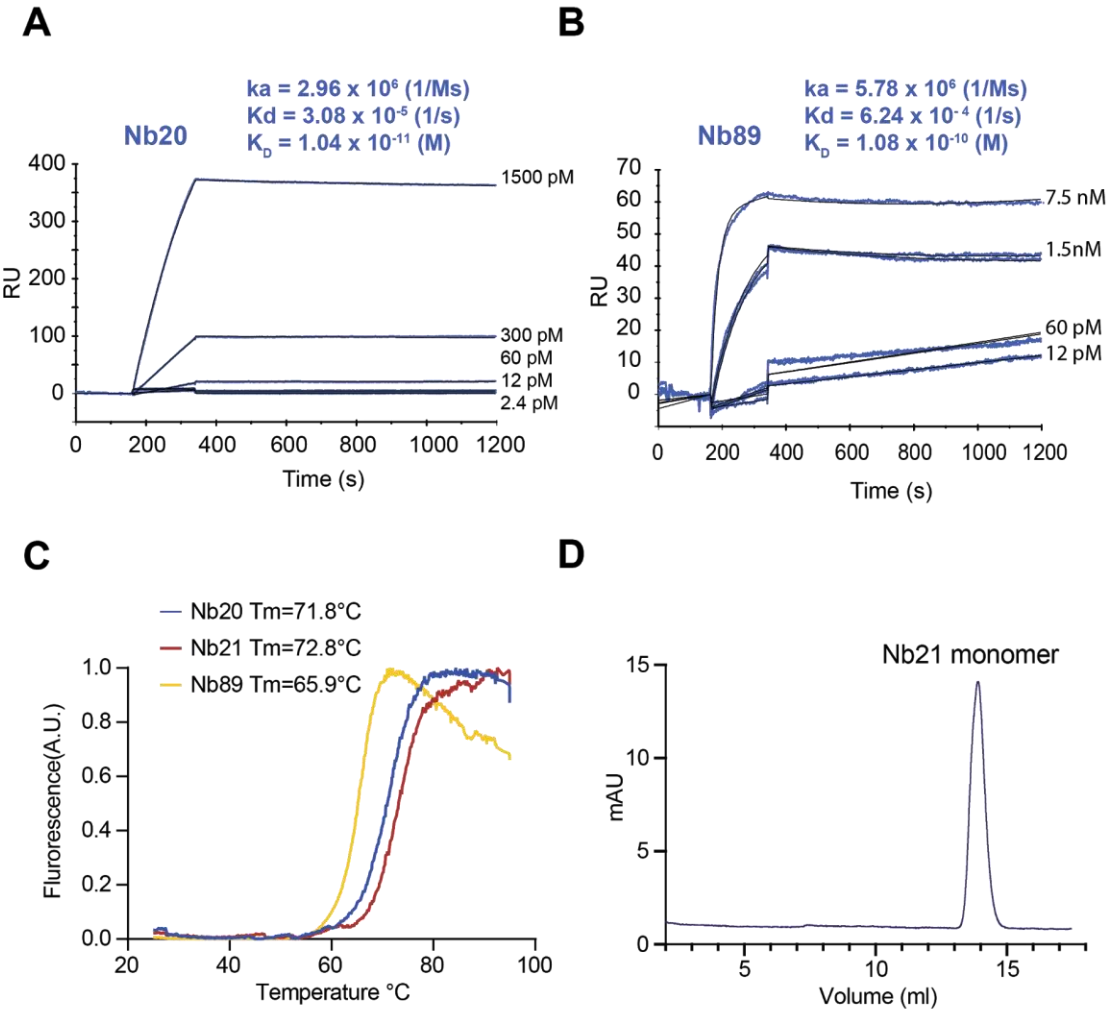


Figure S4. Biophysical analysis of the outstanding neutralizing Nbs.

S4A-B: Binding kinetics of Nbs 20 and 89 by SPR.

S4C: Thermostability analysis of Nbs 20, 21, and 89. The values represent the average thermostability (Tm, °C) based on three replicates. The standard deviations (SD) of the measurements are 0.17, 0.93, and 0.8 °C for Nbs 20, 21, and 89.

S4D: Stability analysis of Nb21 by SEC. Purified recombinant Nb21 was stored at room temperature for ~ 6 weeks before subject to SEC analysis. The dominant peak represents the Nb21 monomer.

Supplementary Figure 5

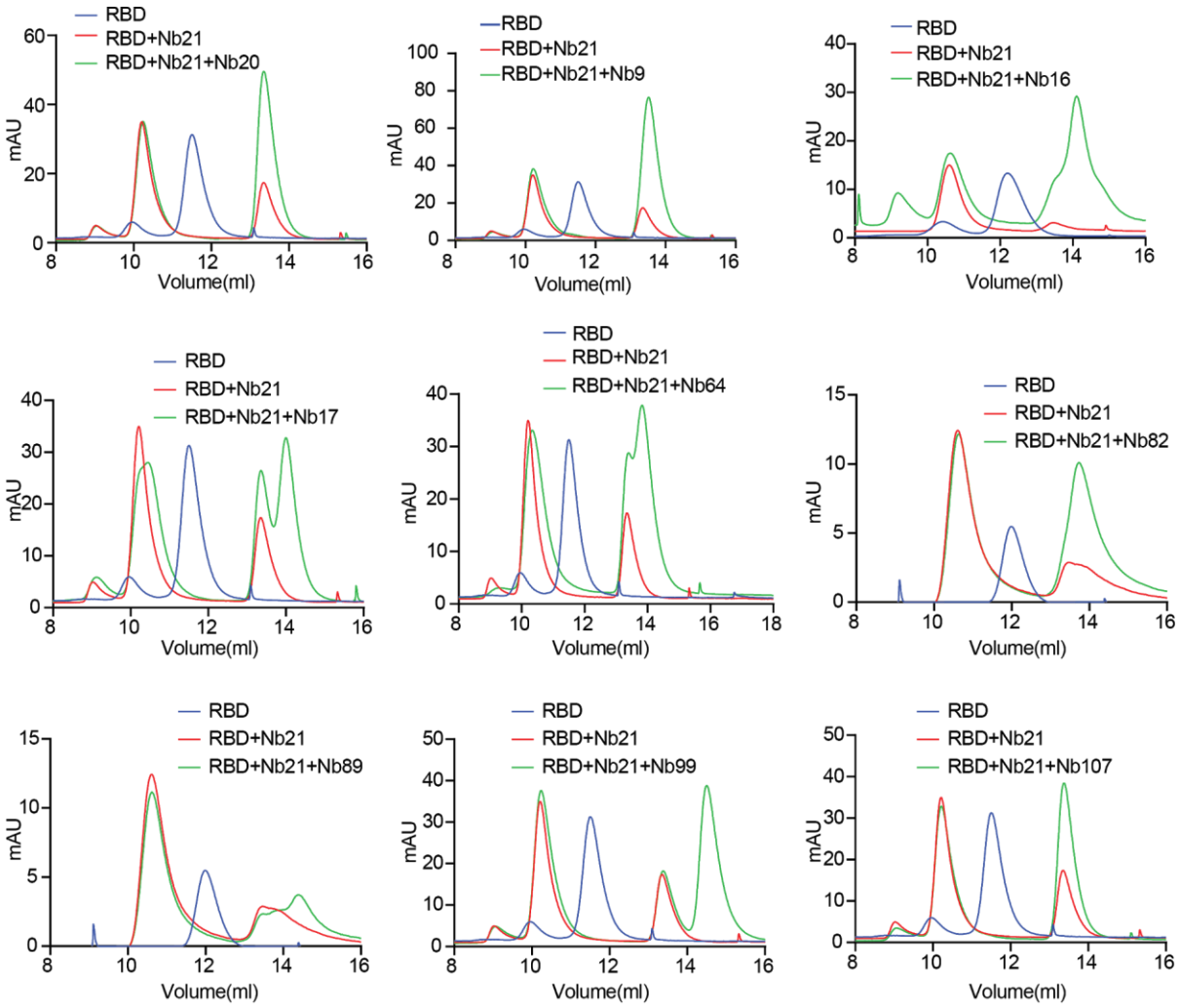


Figure S5. SEC analysis of RBD-Nb complexes.

The SEC profiles of RBD-Nb complexes showing 9 Nbs that have overlapping epitopes with Nb21.

Supplementary Figure 6

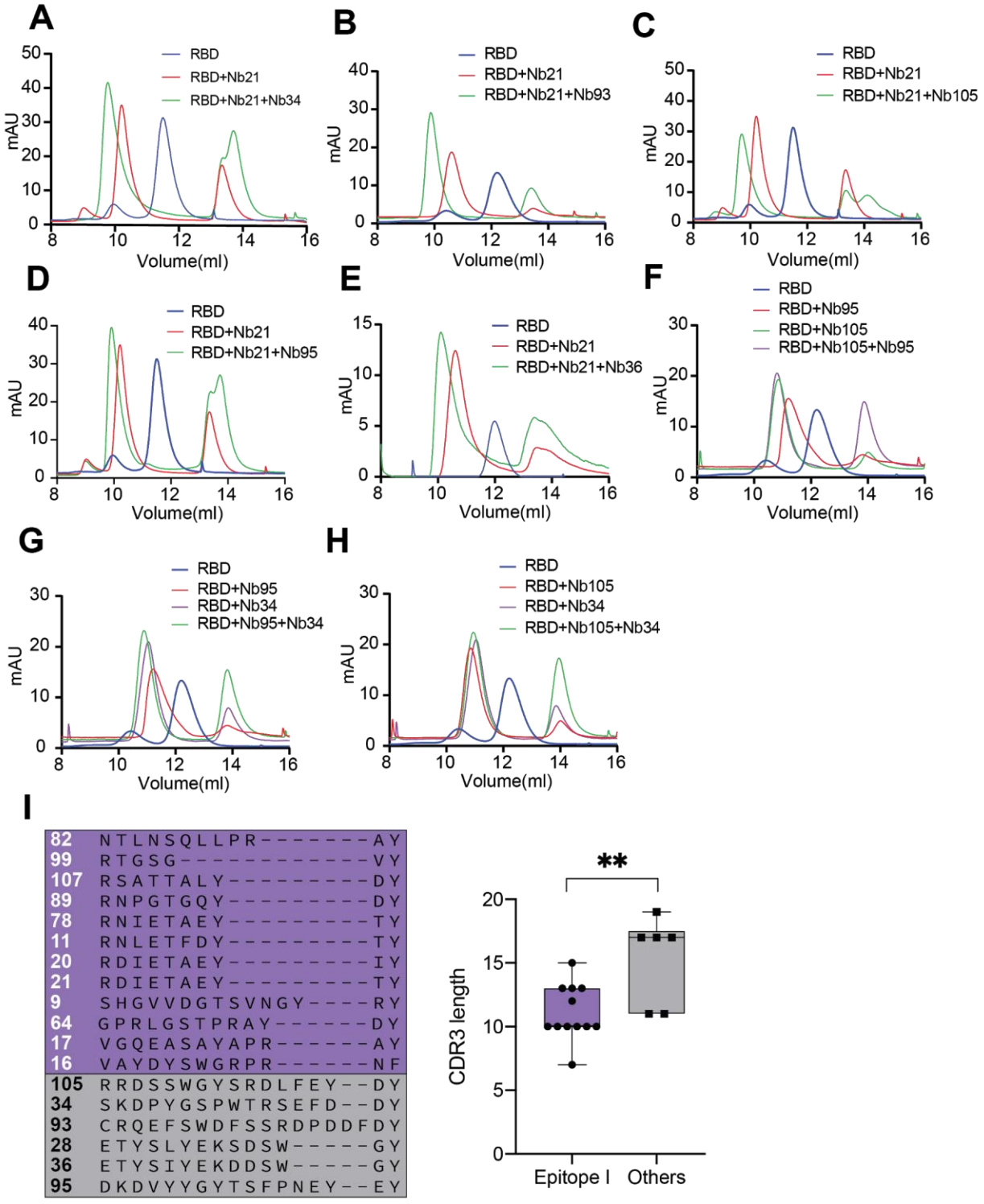


Figure S6. SEC analysis of RBD-Nb complexes.

S6A-H: The SEC profiles of RBD-Nb complexes showing five Nbs that have unique and non-overlapping epitopes with Nb21.

S6I: Sequence alignment of the CDR3s of 18 highly potent neutralizing Nbs and CDR3 lengths comparing Nbs from epitope I and others.

Supplementary Figure 7

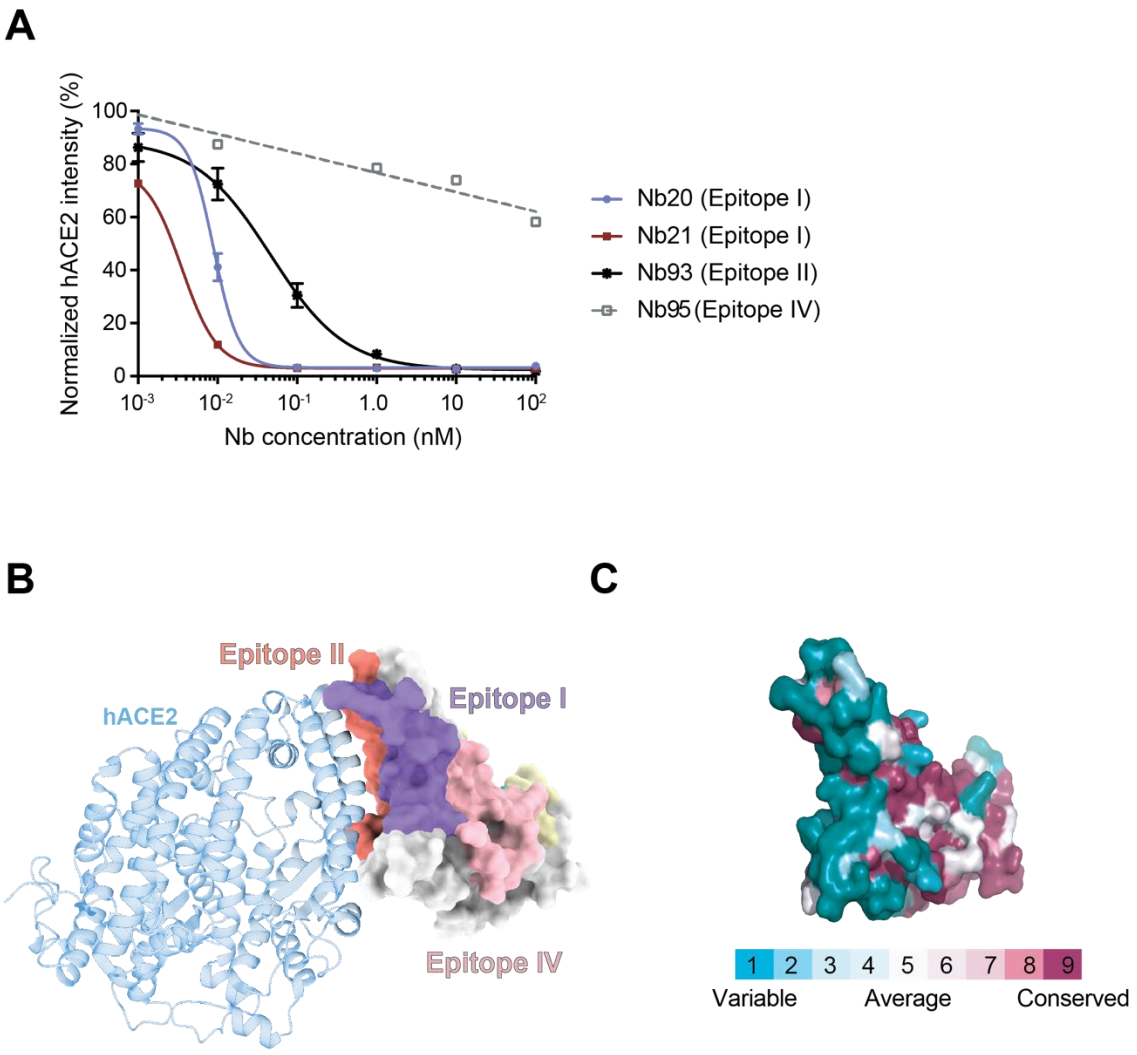


Figure S7. Competitive ELISA analysis of hACE2 and Nbs for RBD binding and conversation analysis of RBD across different coronaviruses.

S7A: Competitive ELISA of hACE2 and Nbs (20, 21, 93, and 95) for RBD binding. Y-axis: percentage of the normalized ACE2 signal. X-axis: Nb concentration (nM).

S7B: The surface display of different Nb neutralization epitopes on RBD in complex with hACE2 (cartoon model in blue).

S7C: The conservation analysis of the spike protein RBD using the ConSurf web server, oriented as in B. The conservation is based on 150 sequences automatically extracted by the ConSurf server.

Supplementary Figure 8

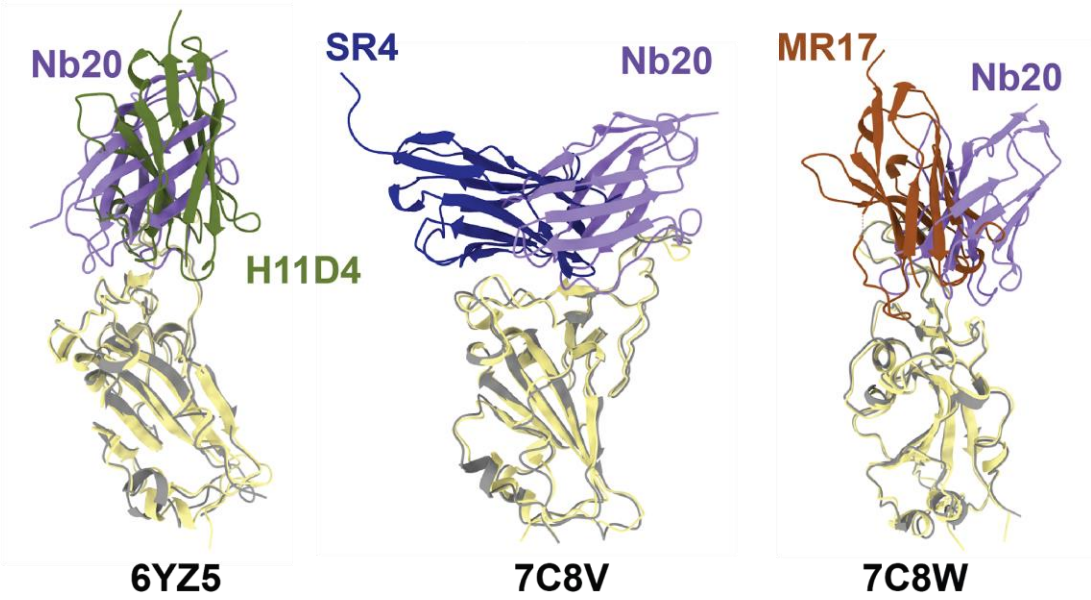


Figure S8. Structural comparisons of Nb20 with published RBD Nb structures.
Overlays of Nb20 (purple ribbon) and three other RBD-Nbs (PDBs 6YZ5, 7C8V, and 7C8W) in complex with RBD (yellow/grey ribbon).

Supplementary Figure 9

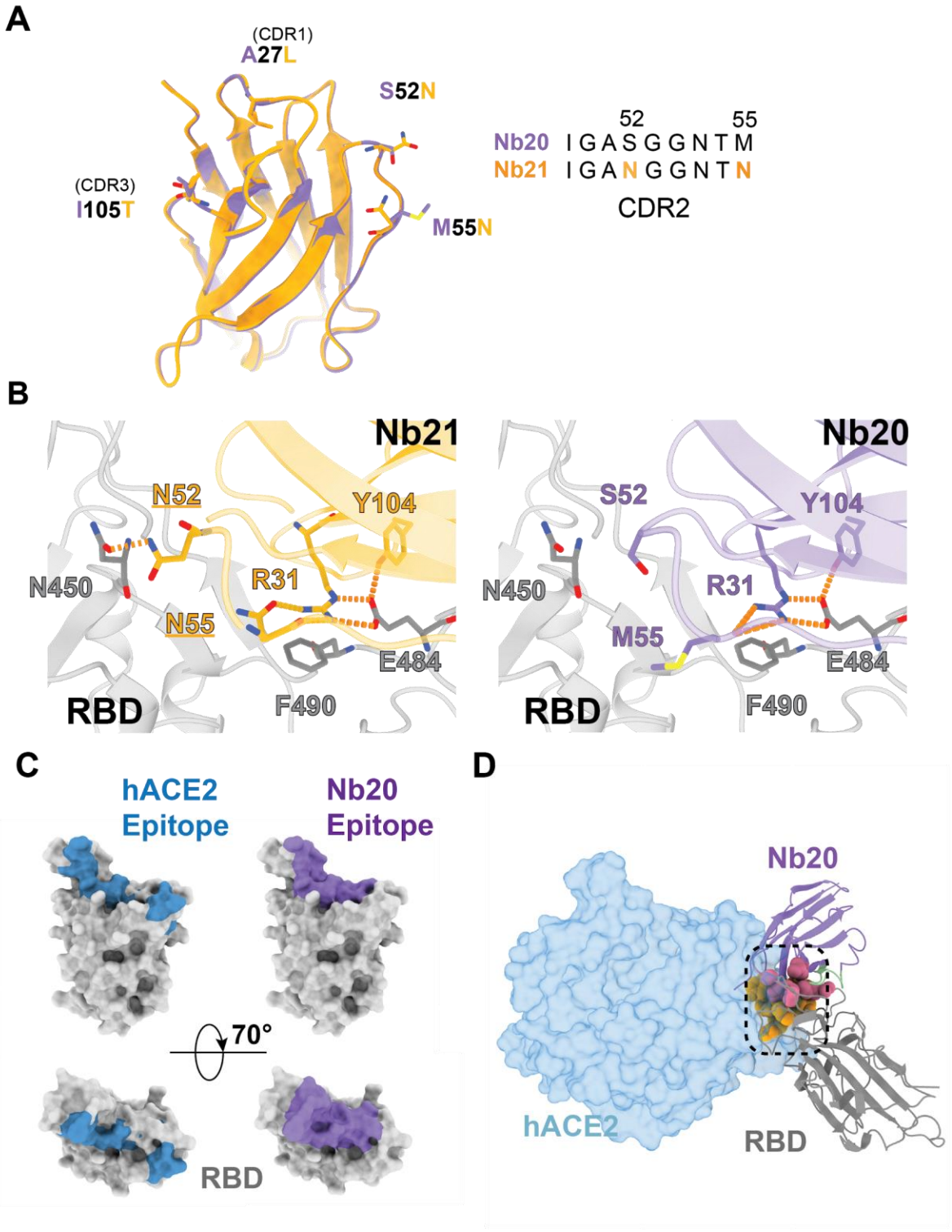


Figure S9. Structural modeling of Nb 21-RBD interaction based on the Nb20-RBD crystal structure.
S9A: Alignment of Nb21 with Nb20. The four residue differences between the two Nbs were shown.
S9B: Zoom-in views showing the addition of new polar interaction between N52 (Nb21) and N450 (RBD). The model of Nb21 is superimposed based on the crystal structure of Nb20.
S9C: Surface presentation of RBD. The hACE2 binding epitope is in steel blue and the Nb20 epitope is in medium purple.

S9D: Structural alignment of Nb20-RBD complex with hACE2-RBD complex. The CDR1 and CDR3 residues (medium violet pink and goldenrod in spheres, respectively) of Nb20 overlap with the hACE2 binding site (steel blue) on RBD (grey ribbon).

Supplementary Figure 10

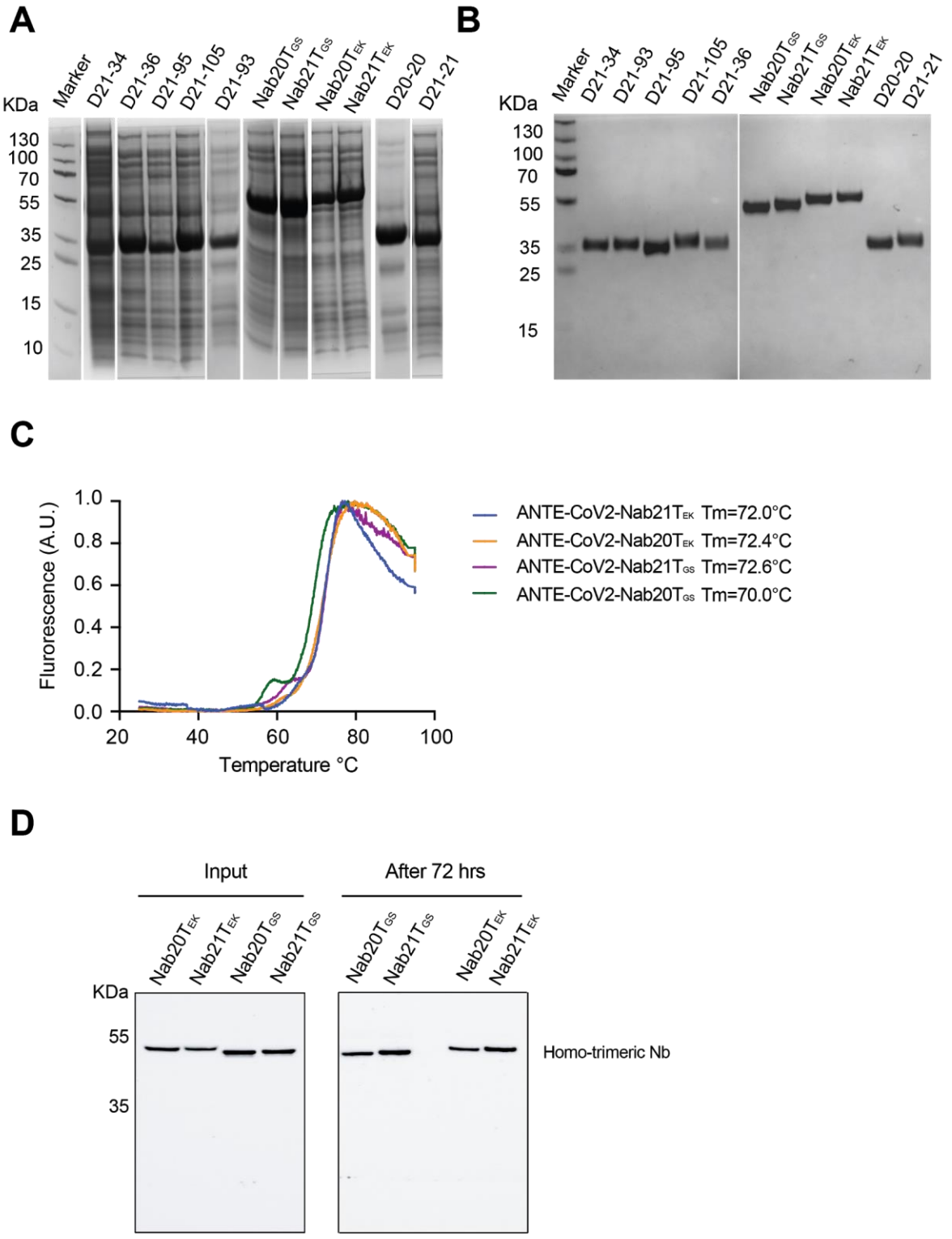
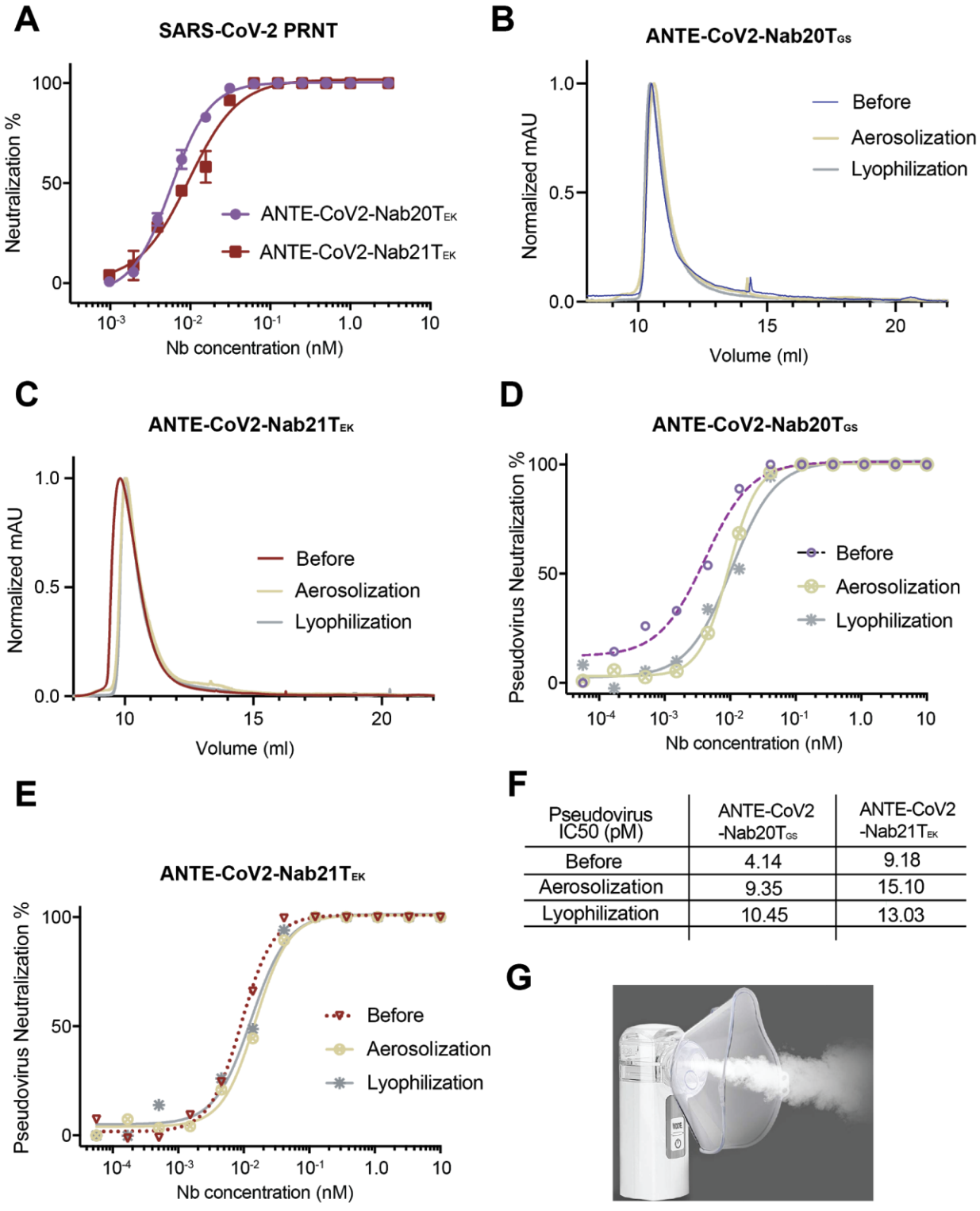


Figure S10. The biophysical properties of multivalent Nbs.
S10A: The expression levels of multivalent Nbs from *E.coli* whole cell lysates.
S10B: SDS-PAGE analysis of the purified multivalent Nbs.

S10C: Thermostability analysis of ANTE-CoV2-Nab21T$_{EK}$, ANTE-CoV2-Nab20T$_{EK}$, ANTE-CoV2-Nab21T$_{GS}$, and ANTE-CoV2-Nab20T$_{GS}$. The values represent the average thermostability (Tm, °C) based on three replicates. The standard deviations of the measurements are 0.6, 0.27, 0.169, and 0.72°C, respectively.

S10D: High stability of the multivalent Nbs under the pseudovirus neutralization condition. Different Nb constructs were incubated under the pseudovirus neutralization assay condition without the virus for 72 hours. An anti-His6 mouse monoclonal antibody (Genscript) was used to detect the Nb constructs (His6 tag at the C terminus) by western blot.

Supplementary Figure 11



F

Pseudovirus IC50 (pM)	ANTE-CoV2 -Nab20TGS	ANTE-CoV2 -Nab21TEK
Before	4.14	9.18
Aerosolization	9.35	15.10
Lyophilization	10.45	13.03

G

Figure S11. Stability test of the multivalent Nbs.
S11A: SARS-CoV-2 PRNT of the homo-trimeric forms of Nbs 20 and 21 with an EK linker. The average neutralization percentage and the standard deviation of each data point were shown (n = 2).

S11B-C: The SEC analysis of ANTE-CoV2-Nab20T$_{GS}$ and ANTE-CoV2-Nab21T$_{EK}$ before and after lyophilization or aerosolization.
S11D-E: Pseudotyped SARS-CoV-2 neutralization assay using ANTE-CoV2-Nab20T$_{GS}$ and ANTE-CoV2-Nab21T$_{EK}$ before and after lyophilization or aerosolization (n = 2).
S11F: A summary table of the neutralization potency measurements of the homo-trimeric Nbs.
S11G: A portable mesh nebulizer (producing ≤ 5μm aerosol particles) used in the study.

Supplementary Figure 12

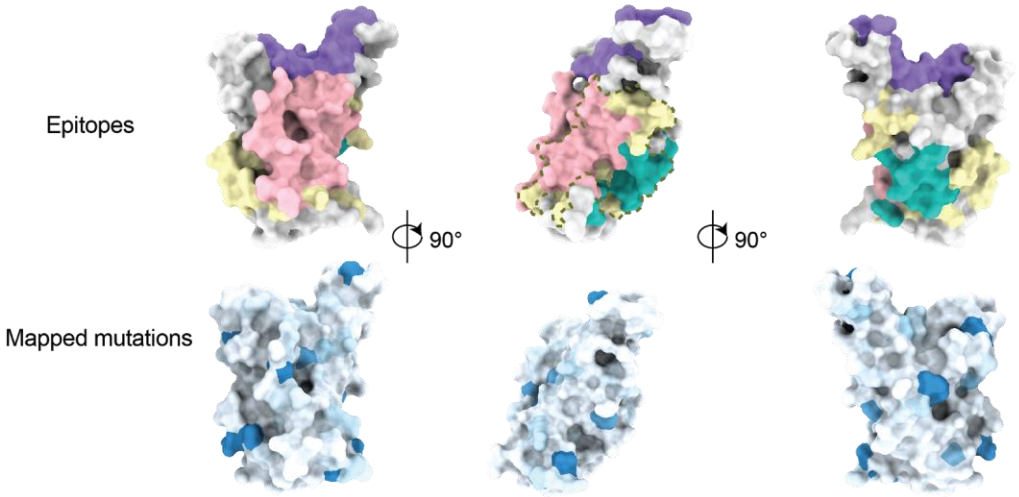


Figure S12. The neutralization epitopes and virus mutations mapped on the RBD crystal structure. The dashed line (upper panel) indicates epitope V that partially overlaps with epitopes III and IV. The mutations (lower panel) are colored in gradient blue (0-100 mutation count from the GISAID), where darker blue indicates more frequent mutations.

Supplementary Figure 13

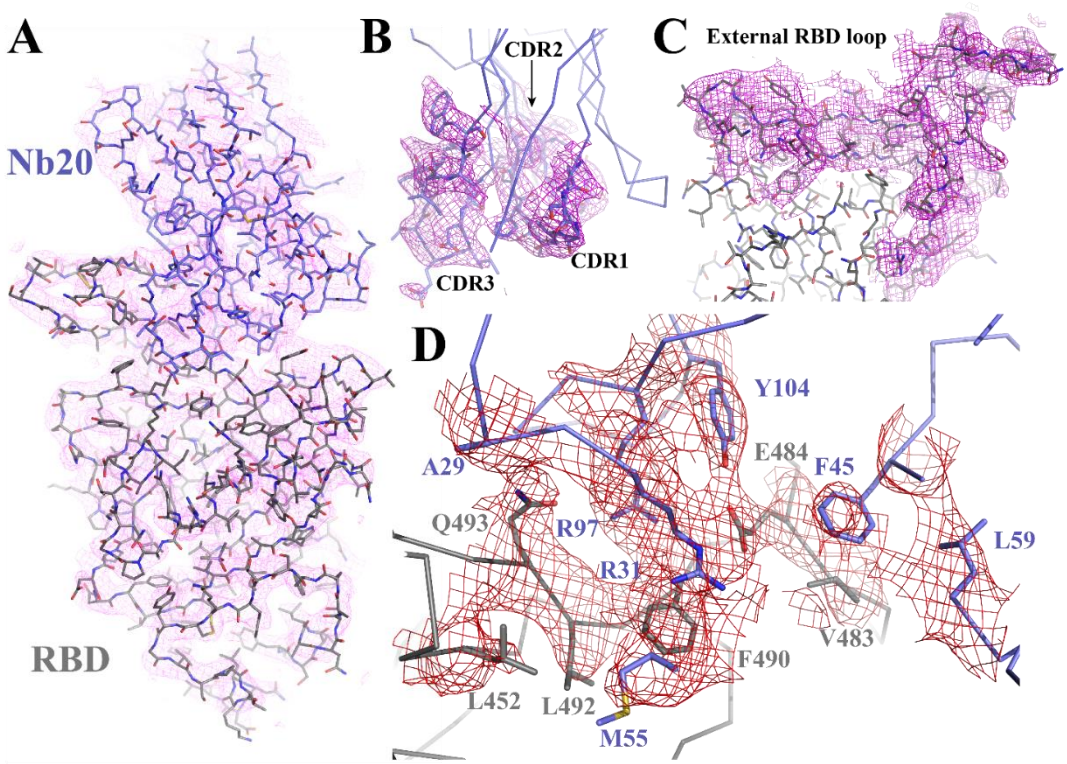


Figure S13. Composite 2Fc-Fo electron density maps of the representative areas of RBD-Nb20 complex contoured at 1.0σ.
S13A: Map of the whole complex shown as purple mesh.
S13B: Map of the three CDRs of Nb20 shown as purple mesh.
S13C: Map of the extended external loop region of RBD shown as purple mesh.
S13D: Map of residues involved in the interactions between RBD and Nb20 shown as red mesh. RBD is colored in gray and Nb20 is colored in blue.

Table S1. Summary of the biophysical, biochemical, and functional properties of the monomeric RBD-Nbs and the multivalent constructs.

Table S2. Summary of the cross-links of the Nbs from different epitopes.

Table S3. X-ray diffraction data collection and refinement statistics.

Table S4. Amino acid sequences of the monomeric and multivalent RBD-Nbs.

References

41. P. C. Fridy *et al.*, A robust pipeline for rapid production of versatile nanobody repertoires. *Nat Methods* **11**, 1253-1260 (2014).
42. J. Dunbar, C. M. Deane, ANARCI: antigen receptor numbering and receptor classification. *Bioinformatics* **32**, 298-300 (2016).
43. S. Kumar, G. Stecher, M. Li, C. Knyaz, K. Tamura, MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol* **35**, 1547-1549 (2018).
44. A. Tareen, J. B. Kinney, Logomaker: beautiful sequence logos in Python. *Bioinformatics* **36**, 2272-2274 (2020).
45. Y. Shi *et al.*, Structural characterization by cross-linking reveals the detailed architecture of a coatomer-related heptameric module from the nuclear pore complex. *Mol Cell Proteomics* **13**, 2927-2943 (2014).
46. Y. Shi *et al.*, A strategy for dissecting the architectures of native macromolecular assemblies. *Nat Methods* **12**, 1135-1138 (2015).
47. Y. Xiang, Z. Shen, Y. Shi, Chemical Cross-Linking and Mass Spectrometric Analysis of the Endogenous Yeast Exosome Complexes. *Methods Mol Biol* **2062**, 383-400 (2020).
48. S. J. Kim *et al.*, Integrative structure and functional anatomy of a nuclear pore complex. *Nature* **555**, 475-482 (2018).
49. A. Sali, T. L. Blundell, Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* **234**, 779-815 (1993).
50. A. Fiser, A. Sali, ModLoop: automated modeling of loops in protein structures. *Bioinformatics* **19**, 2500-2501 (2003).
51. D. Schneidman-Duhovny *et al.*, A method for integrative structure determination of protein-protein complexes. *Bioinformatics* **28**, 3282-3289 (2012).
52. D. Schneidman-Duhovny, H. J. Wolfson, Modeling of Multimolecular Complexes. *Methods Mol Biol* **2112**, 163-174 (2020).
53. G. Q. Dong, H. Fan, D. Schneidman-Duhovny, B. Webb, A. Sali, Optimized atomic statistical potentials: assessment of protein interfaces and loops. *Bioinformatics* **29**, 3158-3166 (2013).
54. Z. Otwinowski, W. Minor, Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**, 307-326 (1997).
55. A. J. McCoy *et al.*, Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674 (2007).
56. P. D. Adams *et al.*, PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221 (2010).
57. P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132 (2004).
58. C. J. Williams *et al.*, MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci* **27**, 293-315 (2018).
59. T. D. Goddard *et al.*, UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci* **27**, 14-25 (2018).
60. F. H. Niesen, H. Berglund, M. Vedadi, The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protoc* **2**, 2212-2221 (2007).