nature

https://doi.org/10.1038/s41586-021-03676-z

Accelerated Article Preview

Nanobodies from camelid mice and llamas neutralize SARS-CoV-2 variants

Received: 4 March 2021

Accepted: 27 May 2021

Accelerated Article Preview Published online 7 June 2021

Cite this article as: Xu, J. et al. Nanobodies from camelid mice and llamas neutralize SARS-CoV-2 variants. *Nature* https://doi.org/10.1038/s41586-021-03676-z (2021).

Jianliang Xu, Kai Xu, Seolkyoung Jung, Andrea Conte, Jenna Lieberman, Frauke Muecksch, Julio Cesar Cetrulo Lorenzi, Solji Park, Fabian Schmidt, Zijun Wang, Yaoxing Huang, Yang Luo, Manoj Nair, Pengfei Wang, Jonathan E. Schulz, Lino Tessarollo, Tatsiana Bylund, Gwo-Yu Chuang, Adam S. Olia, Tyler Stephens, I-Ting Teng, Yaroslav Tsybovsky, Tongqing Zhou, Vincent Munster, David D. Ho, Theodora Hatziioannou, Paul D. Bieniasz, Michel C. Nussenzweig, Peter D. Kwong & Rafael Casellas

Nanobodies from camelid mice and llamas neutralize SARS-CoV-2 variants

https://doi.org/10.1038/s41586-021-03676-z

Received: 4 March 2021

Accepted: 27 May 2021

Published online: 7 June 2021

Jianliang Xu[1,13]✉, Kai Xu[2,12,13], Seolkyoung Jung[1], Andrea Conte[1], Jenna Lieberman[1], Frauke Muecksch[3], Julio Cesar Cetrulo Lorenzi[4], Solji Park[1], Fabian Schmidt[3], Zijun Wang[4], Yaoxing Huang[5], Yang Luo[5], Manoj Nair[5], Pengfei Wang[5], Jonathan E. Schulz[6], Lino Tessarollo[7], Tatsiana Bylund[2], Gwo-Yu Chuang[2], Adam S. Olia[2], Tyler Stephens[8], I-Ting Teng[2], Yaroslav Tsybovsky[8], Tongqing Zhou[2], Vincent Munster[6], David D. Ho[5], Theodora Hatziioannou[3], Paul D. Bieniasz[3,9], Michel C. Nussenzweig[4,9]✉, Peter D. Kwong[2,13]✉ & Rafael Casellas[1,10,11,13]✉

Since the start of the COVID-19 pandemic, SARS-CoV-2 has caused millions of deaths worldwide. While many vaccines have been deployed to date, the continual evolution of the viral receptor-binding domain (RBD) has challenged their efficacy. In particular, emerging variants B.1.1.7 (U.K.), B.1.351 (South Africa) and P.1 (Brazil) have compromised convalescent sera and immunotherapies that received emergency use authorization[1–3]. One potential alternative to avert viral escape is the use of camelid VHHs or nanobodies, which can recognize epitopes often inaccessible to conventional antibodies[4]. Here, we isolate anti-RBD nanobodies from llamas and "nanomice" we engineered to produce VHHs cloned from alpacas, dromedaries and camels. We identified two sets of highly neutralizing nanobodies. Group 1 circumvents antigenic drift by recognizing an RBD region that is highly conserved in coronaviruses but rarely targeted by human antibodies. Group 2 is almost exclusively focused to the RBD-ACE2 interface and fails to neutralize variants carrying E484K or N501Y substitutions. Notably however, group 2 nanobodies retain full neutralization activity against variants when expressed as homotrimers, rivaling the most potent antibodies produced to date against SARS-CoV-2. These findings suggest that multivalent nanobodies overcome SARS-CoV-2 mutations through two separate mechanisms: enhanced avidity for the ACE2 binding domain, and recognition of conserved epitopes largely inaccessible to human antibodies. Therefore, while new SARS-CoV-2 mutants will continue to emerge, nanobodies represent promising tools to prevent COVID-19 mortality when vaccines are compromised.

Generation and characterization of nanomice

In contrast to mouse and human antibody binding domains, which are ~50 kDa in size, camelid VHHs (Variable Heavy chain domains of a Heavy chain, also known as nanobodies (Nbs)) retain full antigen specificity at ~15 kDa. This feature and extended CDRs allow Nbs to bind epitopes not normally accessible to conventional antibodies[4], such as conserved viral domains which are often masked by glycan shields. Nbs can be readily humanized[5] and in recent clinical trials they appeared safe and of low immunogenicity[6]. Despite these advantages, Nbs are still not widely used. One reason is that camelids are large animals not suitable for academic facilities. There are also few reagents available to isolate antigen-specific memory B cells from immunized camelids[7]. To bypass these hurdles, we sought to produce Nbs in mice by combining 18 alpaca, 7 dromedary and 5 camel VHH genes in a 25 Kb insertion cassette (Fig.1a). Each gene was fused to a VH promoter, leader exons, and recombination signal sequences (RSSs) to ensure physiological expression and recombination (Extended Data Fig. 1). By means of CRISPR-Cas9, the VHH cassette was inserted in lieu of the VH locus in mouse embryonic stem (ES) cells (Fig.1a).

Camelid Nbs are only expressed in conjunction with dedicated IgG2 and IgG3, which splice out the CH1 exon during transcription[4].

[1]Lymphocyte Nuclear Biology, NIAMS, NIH, Bethesda, MD, 20892, USA. [2]Vaccine Research Center, NIAID, NIH, Bethesda, MD, 20892, USA. [3]Laboratory of Retrovirology, The Rockefeller University, New York, NY, 10065, USA. [4]Laboratory of Molecular Immunology, The Rockefeller University, New York, NY, 10065, USA. [5]Aaron Diamond AIDS Research Center, Columbia University Vagelos College of Physicians and Surgeons, New York, NY, 10032, USA. [6]Laboratory of Virology, Division of Intramural Research, NIAID, NIH, Rocky Mountain Laboratories, Hamilton, MT, 59840, USA. [7]Mouse Cancer Genetics Program, CCR, NCI, NIH, Frederick, MD, 21702, USA. [8]Electron Microscopy Laboratory, Cancer Research Technology Program, Frederick National Laboratory for Cancer Research sponsored by the National Cancer Institute, Frederick, MD, 21701, USA. [9]Howard Hughes Medical Institute, The Rockefeller University, New York, NY10065, USA. [10]The NIH Regulome Project, NIH, Bethesda, MD, 20892, USA. [11]Center for Cancer Research, NCI, NIH, Bethesda, MD, 20892, USA. [12]Department of Veterinary Biosciences, College of Veterinary Medicine, The Ohio State University, Columbus, OH, 43210, USA. [13]These authors contributed equally: Jianliang Xu, Kai Xu, Peter D. Kwong, Rafael Casellas. ✉e-mail: jianliang.xu@nih.gov; nussen@mail.rockefeller.edu; pdkwong@nih.gov; rafael.casellas@nih.gov

In conventional antibodies, the hydrophobic surface of CH1 helps to pair heavy and light chain constant domains. To recapitulate this configuration in the mouse genome, we deleted CH1 from both *Igμ* and *Igγ1* in the ES cells (Fig. 1a). The targeted allele was germline transmitted from mouse chimeras to F1 offspring (hereafter referred to as nanomice).

As expected, in wild type (WT) mice ~85% of splenic B220$^+$ B cells were IgM$^+$Igκ$^+$ (Fig.1b, left panel). In stark contrast, in heterozygous nanomice 72% displayed an IgM$^+$Igκ$^-$ phenotype (Fig.1b, right panel). Of these, less than 2% were IgM$^+$Igλ$^+$ (Extended Data Fig. 2a), implying that a large fraction of nanomouse B cells develop expressing single-chain antibodies. We confirmed this observation by amplifying VHH-DJ joining events using gene-specific primers. We found that all 30 VHHs were recombined to downstream JHs in bone marrow and spleen samples (Extended Data Fig. 2b). A deep-sequencing analysis confirmed that all VHH genes undergo V(D)J recombination and are thus potentially available for expansion during the immune response (Extended Data Fig. 2c).

In VHH homozygous mice the B cell compartment was largely normal, displaying all developmental stages including B1, B2 and marginal zone B cells (Extended Data Fig. 3a). One difference was an increased number of IgM$^+$ transitional and immature B cells in the bone marrow and spleen respectively, indicative of enhanced selection as cells transition from the short- to the long-lived CD23^{high}CD21^{low} compartment, which in nanomice was reduced 1.7-fold (Extended Data Fig. 3a). Another distinct feature was the absence of IgD (Extended Data Fig. 3b). This phenotype likely results from differential mRNA splicing due to CH1 deletion at *Igμ*, because IgD was also absent in *Igμ*-CH1$^{-/-}$ only mice, where VHs and *Igγ1* CH1 are intact (Extended Data Fig. 3b). Taken together, the data show that mouse B cells can mature expressing single-chain antibodies.

Activation and hypermutation in nanomice

To probe activation, splenic B cells were isolated and cultured in the presence of lipopolysaccharide (LPS) and interleukin-4 (IL-4). Under these conditions, VHH-expressing cells underwent proliferation and switch recombination to IgG1 (Extended Data Fig. 3b-c). To examine activation *in vivo*, we performed intraperitoneal immunizations with keyhole limpet hemocyanin (KLH). Twelve days post-immunization, nanomice showed equivalent numbers of B220$^+$CD95^{high}IgG1$^+$ germinal center B cells relative to controls (Fig. 1c).

To study affinity maturation against a specific antigen, we immunized nanomice with human immunodeficiency virus-1 (HIV-1) envelop trimer (BG505 DS-SOSIP)[8] (Extended Data Fig. 3d). Hypermutation of VHH genes was increased relative to unimmunized controls (1.1e-2 vs. 7.5e-4, Fig. 1d). The mutation spectra revealed an enrichment in G to A and C to T transitions (Extended Data Fig. 3e), consistent with activation-induced cytidine deaminase (AID) catalysis[9].

To measure the antibody response against BG505 DS-SOSIP, we characterized 16 nanobodies that were enriched for HIV-1 trimer recognition. Sequence analysis showed CDR3s to be highly diverse in this group in terms of JH usage, mutations, and size (9-16 amino acids, Extended Data Fig. 4a). To measure binding kinetics, we applied bio-layer interferometry (BLI). The analysis identified four VHH9 variants, which displayed dissociation constants (K_Ds) ranging from 2 to 13 nM, demonstrating that they represent high-affinity binders (Extended Data Fig. 4b and Extended Data Table 1). We conclude that mouse B cells expressing single-chain antibodies can undergo affinity maturation and produce highly specific Nbs upon immunization.

SARS-CoV-2 neutralizing nanobodies

We next sought to produce neutralizing Nbs against severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2). To this end, 3 nanomice and 1 llama were immunized with RBD and the stabilized prefusion spike (Fig. 2a). Peripheral blood mononuclear cells (PBMCs) were isolated post-immunization and VHHs were amplified and cloned into a phagemid vector. Following phage display, RBD-specific Nbs were enriched using an ELISA-based binding screen. Deep-sequencing analysis identified on average 26,000 Nb variants per library, representing a total of 192 unique CDR3s for llama and 199 for nanomice (Extended data Fig. 5a). Nbs were then clustered by CDR3s (see Methods) and representative clones from each subgroup were isolated and tested for blocking RBD binding to the angiotensin-converting enzyme 2 (ACE2) receptor *in vitro*[10]. Six llama and six nanomouse Nbs were selected with this method (Extended data Fig. 5b-c).

To refine the list of candidates we measured RBD binding affinity by BLI. The analysis identified 4 llama (15, 17, 19, and 56) and 2 nanomouse (12 and 30) Nbs with dissociation constants tighter than 30 nM (Fig. 2b, Extended Data Fig. 5d and Extended Data Table 1). The off-rate varied from $7.1x10^{-3}$ to $1.1x10^{-3}$ s^{-1}, demonstrating slow dissociation for all Nbs (Extended Data Fig. 5e). We next explored neutralization *in vitro* using lentiviral particles pseudotyped with the SARS-CoV-2 spike[11]. The Nb monomers displayed nM and sub-nM half-maximal inhibitory concentration (IC_{50}), ranging from 11.7 nM (168.5 ng/ml) for Nb12 to 0.335 nM (4.6 ng/ml) for Nb19 (Fig. 2c).

A crucial advantage of Nbs over conventional antibodies is that they can be easily assembled into multimers, often resulting in remarkable avidity[12,13]. To explore this property, Nbs were fused as trimers using flexible GGGGS(x3) linkers and connected to human IgG1 Fc via the human hinge domain or its much longer and flexible llama counterpart (Fig. 2d). By fusing two VHHs to IgG1 Fc we also created bivalent antibodies (Fig. 2d). We found that neutralization increased with the number of linked monomers, from 3-fold for Nb15 to 180-fold for Nb12 (Fig. 2c and 2e). Notably, the four most potent multimeric Nbs (12, 17, 19, and 56) reached IC50 values in the picomolar range, from 65 pM to 9 pM (Fig. 2c and 2e). These values rank among the best reported to date for anti-SARS-CoV-2 Nbs[14].

Nanobodies overcome SARS-CoV-2 mutants

With the worldwide spread of SARS-CoV-2, several variants carrying RBD mutations have emerged that increase transmissibility or allow escape from antibody neutralization. Of particular interest is the B.1.1.7 variant, containing an N501Y substitution, that caused an upsurge in COVID-19 cases in the United Kingdom[15]. A second variant of concern is the South African B.1.351, which combines N501Y with two additional RBD substitutions: K417N and E484K. A third variant that spread rapidly in Brazil, P.1, share similar changes: N501Y, K417T and E484K[16]. All of these mutations were shown to reduce the efficacy of serum antibodies elicited by the Moderna and Pfizer vaccines[1,2].

We first explored whether our leading Nbs could neutralize SARS-CoV-2 S pseudotyped viruses carrying the RBD mutations. The R683G substitution, which increases infectivity *in vitro*[17] was also included as a control. In contrast to their efficacy against the WT virus, Nb17, Nb19, and Nb56 were unable to neutralize viruses carrying the E484K mutation alone or in combination with K417N and N501Y (KEN construct, Fig. 3a). Similarly, Nb15 was ineffective against N501Y. Surprisingly, however, with the exception of Nb17, they all remained highly potent binders and neutralizers as bivalent or trivalent Nbs (Extended Data Fig. 5e, 6a and Fig. 3a). In the case of Nb15 and Nb56 trimers, IC50 values reached 30 pM and 14 pM respectively. Thus, the E484K and N501Y mutations allow viral escape from monomeric but not multimeric Nbs.

In contrast to llama Nbs, nanomouse Nb12 and Nb30 were intriguing in that their neutralization potencies were largely unaltered by RBD mutations (Fig. 3a), suggesting that they recognize a region different from the receptor binding motif. To explore whether multimeric Nbs function against authentic virus, we repeated the neutralization assay with trivalent Nb15, Nb56, Nb12, and bivalent Nb30 using SARS-CoV-2 WA1 and variants B1.1.7, B1.351, and P.1. The results closely recapitulated the pseudovirus findings, showing neutralization of WT and the

3 variants by all 4 Nbs (Fig. 3b and Extended Data Fig. 6b). Of note, the trimers were most effective against the B1.1.7 variant, with IC50 values ranging between 4 pM for Nb15 to 538 pM for Nb30, and relatively less effective against the B.1.351 variant, showing a range of 18 pM for Nb56 to 2,755 pM for Nb30 (Extended Data Fig. 6c). Neutralization of the P.1 variant was intermediate (Extended Data Fig. 6b-c).

The fact that llama and nanomouse Nbs are differentially affected by the variants suggests that they recognize different RBD epitopes. To test this idea, we applied BLI in which a preformed Nb-RBD immunocomplex was incubated with a second Nb (Fig. 3c). We found that all four llama Nbs, but not Nb30, could bind the Nb12-RBD immunocomplex (Fig. 3d). Likewise, Nb30-RBD interfered with Nb12 binding, while llama Nbs bound freely to it (Fig. 3e). At the same time, Nb12 and Nb30 recognized all combinations of llama Nb-RBD complexes, whereas llama Nbs could not (Fig. 3f and Extended Data Fig. 6d). Thus, nanomouse and llama Nbs recognize two distinct neutralizing RBD regions.

As is often the case with single-chain antibodies, both llama and nanomouse Nbs were largely thermostable and could be aerosolized with commercially available mesh nebulizers without losing neutralization activity (Extended Data Fig. 6e-g).

Nanobody structures

To define the region bound by nanomouse Nbs, we collected single particle cryo-electron microscopy (EM) data on a Titian Krios for Nb12 and Nb30 in complex with HexaPro[10], a prefusion construct of the SARS-CoV-2 spike (Extended Data Fig. 7 and 8; Extended Data Table 2). In both cases, we used particle subtraction, classification and location refinement to enhance the resolution of the Nb-spike interface.

The structure of the Nb12-spike complex revealed Nb12 to induce a 2-RBD-up, 1 RBD-down spike conformation, with Nb12 recognizing a region towards the middle of the RBD, outside of the ACE2-binding region and distal from the 417, 484 and 501 mutations in emerging variants of concern (Fig. 4a and Extended Data Fig. 9a). Meanwhile, the structure of the Nb30-spike complex revealed Nb30 to induce a 3-RBD-up conformation, with Nb30 recognizing a region at the opposite end of RBD from the ACE2 binding motif and escape mutations (Fig. 4b and Extended Data Fig. 9b).

To understand how the two nanomouse Nbs neutralized despite recognizing surfaces outside the ACE2 binding domain, we superimposed the structure of the ACE2-RBD complex[18–20] with those of Nb12 and Nb30 (Extended Data Fig. 9c). We observed a substantial portion of Nb12 domain clashing with ACE2, indicating Nb12 and ACE2 binding to be sterically incompatible. With Nb30, a more subtle clash with glycan N322 on ACE2 was observed, which nonetheless also indicated Nb30 and ACE2 binding to be sterically incompatible.

To obtain a structural understanding of the neutralizing regions recognized by nanomouse and llama Nbs, we also determined 3d-negative stain EM reconstructions of each of the Nbs in complex with HexaPro. These reconstructions revealed that llama Nbs uniformly target the ACE2-binding interface, with Nb17, Nb19, and Nb56 inducing 1-RBD-up conformation, while Nb15 associates with all-RBD-down spikes (Extended Data Fig. 9d-h). By contrast, both nanomouse Nb12 and Nb30 recognize RBD at a surface outside the ACE2-binding site (Fig. 4c).

RBD regions recognized by mouse nanobodies

To provide insight into the prevalence of regions on RBD recognized by nanomouse versus human antibodies, we superimposed 51 RBD-directed human neutralizing antibodies in the PDB and quantified the recognition prevalence at the residue level (Fig. 4d; Extended Data Table 3). While recognition extended over much of the RBD, the prevalence of human antibody recognition was much higher in the ACE2-binding region where emerging mutations reside. By contrast, the regions recognized by Nb12 and Nb30 were conserved in Sarbecoviruses and displayed substantially lower prevalence of human antibody recognition. Interestingly, the epitope of Nb12 overlaps considerably with those recognized by previous Nbs specific for SARS-CoV-1[21] and -CoV-2[13,22], raising the possibility that those Nbs might also block the new SARS-CoV-2 variants. Nb30 binding footprint however is distanced further away from the ACE2 RBD motif and cover a surface area that is 79% conserved among Sarbecoviruses including SARS-CoV-1, Bat-CoV and SARS-CoV-2, compared to 54% for Nb12 and on average 23% for human antibodies (Fig. 4c and d; Extended Data Table 3). Consistent with these findings, we found that Nb12 and Nb30 bind to SARS-CoV-1 and the bat coronavirus WIV16 RBD and neutralize HIV-1-based pseudoviruses carrying their spike, while Nb56 does not (Extended Data Fig. 10a and b). Similar neutralization patterns were observed with VSV-based pseudoviruses carrying spike proteins from pangolin and 6 additional bat coronaviruses (Extended Data Fig. 10c). We conclude that nanomouse VHHs circumvent RBD antigenic drift by recognizing a Sarbecoviral conserved region outside the ACE2-binding motif.

Discussion

A key contribution of our study is the creation of Nb-producing mice. Previous work explored transgenic expression of 1-2 llama VHHs[23]. In our model, the 30 VHHs replace the entire VH domain, leading to physiological recombination and selection during ontogeny. While capable of producing high-affinity Nbs, our nanomouse 1.0 can be improved further by increasing the number of available VHHs. This could be done by engineering a second allele carrying VHHs from llamas, vicuñas and guanacos, the three camelids not represented in our insertion cassette. We anticipate that this and similar improvements in animal models will help popularize the development of Nbs against infectious diseases or for basic applications.

As a proof of principle, we used the nanomice to produce highly specific Nbs against SARS-CoV-2 RBD. To date, numerous monoclonal antibodies isolated from COVID-19 patients or humanized mice have been shown to block the RBD-ACE2 interface. Not surprisingly, immunotherapies involving such antibodies are vulnerable to escape variants carrying mutations at or around the ACE2 binding motif[1–3]. The anti-RBD Nbs we isolated overcome this limitation in two important ways. First, similar to human antibodies, llama Nbs (Nb15, Nb56) hinder ACE2 binding to the spike of the original virus, but they are ineffective against viruses that carry E484K or N501Y substitutions. However, in multimeric form, these Nbs overcome the block and display remarkable neutralization potency. This reversal is likely the result of increased avidity for the trimeric spike, or possibly the simultaneous cross-linking of multiple spikes on the viral membrane. Another possibility is that trimers occlude ACE2 access to the RBD. Second, Nbs isolated from nanomice (Nb12, Nb30) associate with an RBD region that is highly conserved among Sarbecoviruses[21], but remains inaccessible to most human antibodies. As this region lies outside the ACE2 binding motif, Nb-RBD contacts are unaffected by E484K or N501Y. Importantly, even though the conserved domain does not overlap with the ACE2 binding motif, our structural studies suggest that Nbs of this class sterically interfere with ACE2-RBD associations. Based on these features we propose that our leading Nbs may provide valuable tools for passive immunotherapy or pulmonary delivery against current or future SARS-CoV-2 variants of concern.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-021-03676-z.

1. Wang, Z. *et al*. mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating variants. *Nature* **592**, 616-622, https://doi.org/10.1038/s41586-021-03324-6 (2021).
2. Wang, P. *et al*. Antibody resistance of SARS-CoV-2 variants B.1.351 and B.1.1.7. *Nature* **593**, 130-135, https://doi.org/10.1038/s41586-021-03398-2 (2021).
3. Wu, K. *et al*. mRNA-1273 vaccine induces neutralizing antibodies against spike mutants from global SARS-CoV-2 variants. *bioRxiv*, https://doi.org/10.1101/2021.01.25.427948 (2021).
4. Muyldermans, S. Nanobodies: natural single-domain antibodies. *Annu Rev Biochem* **82**, 775-797, https://doi.org/10.1146/annurev-biochem-063011-092449 (2013).
5. Muyldermans, S. Applications of Nanobodies. *Annu Rev Anim Biosci* **9**, 401-421, https://doi.org/10.1146/annurev-animal-021419-083831 (2021).
6. Scully, M. *et al*. Caplacizumab Treatment for Acquired Thrombotic Thrombocytopenic Purpura. *N Engl J Med* **380**, 335-346, https://doi.org/10.1056/NEJMoa1806311 (2019).
7. Hussen, J. & Schuberth, H. J. Recent Advances in Camel Immunology. *Front Immunol* **11**, 614150, https://doi.org/10.3389/fimmu.2020.614150 (2020).
8. Kong, R. *et al*. Antibody Lineages with Vaccine-Induced Antigen-Binding Hotspots Develop Broad HIV Neutralization. *Cell* **178**, 567-584 e519, https://doi.org/10.1016/j.cell.2019.06.030 (2019).
9. Pham, P., Bransteitter, R., Petruska, J. & Goodman, M. F. Processive AID-catalysed cytosine deamination on single-stranded DNA simulates somatic hypermutation. *Nature* **424**, 103-107, https://doi.org/10.1038/nature01760 (2003).
10. Hsieh, C. L. *et al*. Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. *Science* **369**, 1501-1505, https://doi.org/10.1126/science.abd0826 (2020).
11. Robbiani, D. F. *et al*. Convergent antibody responses to SARS-CoV-2 in convalescent individuals. *Nature* **584**, 437-442, https://doi.org/10.1038/s41586-020-2456-9 (2020).
12. Schoof, M. *et al*. An ultrapotent synthetic nanobody neutralizes SARS-CoV-2 by stabilizing inactive Spike. *Science* **370**, 1473-1479, https://doi.org/10.1126/science.abe3255 (2020).
13. Xiang, Y. *et al*. Versatile and multivalent nanobodies efficiently neutralize SARS-CoV-2. *Science* **370**, 1479-1484, https://doi.org/10.1126/science.abe4747 (2020).
14. Saelens, X. & Schepens, B. Single-domain antibodies make a difference. *Science* **371**, 681-682, https://doi.org/10.1126/science.abg2294 (2021).
15. Kupferschmidt, K. Fast-spreading U.K. virus variant raises alarms. *Science* **371**, 9-10, https://doi.org/10.1126/science.371.6524.9 (2021).
16. Faria, N. R. *et al*. Genomics and epidemiology of a novel SARS-CoV-2 lineage in Manaus, Brazil. *medRxiv*, https://doi.org/10.1101/2021.02.26.21252554 (2021).
17. Schmidt, F. *et al*. Measuring SARS-CoV-2 neutralizing antibody activity using pseudotyped and chimeric viruses. *J Exp Med* **217**, https://doi.org/10.1084/jem.20201181 (2020).
18. Benton, D. J. *et al*. Receptor binding and priming of the spike protein of SARS-CoV-2 for membrane fusion. *Nature* **588**, 327-330, https://doi.org/10.1038/s41586-020-2772-0 (2020).
19. Lan, J. *et al*. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **581**, 215-220, https://doi.org/10.1038/s41586-020-2180-5 (2020).
20. Zhou, T. *et al*. Cryo-EM Structures of SARS-CoV-2 Spike without and with ACE2 Reveal a pH-Dependent Switch to Mediate Endosomal Positioning of Receptor-Binding Domains. *Cell Host Microbe* **28**, 867-879 e865, https://doi.org/10.1016/j.chom.2020.11.004 (2020).
21. Wrapp, D. *et al*. Structural Basis for Potent Neutralization of Betacoronaviruses by Single-Domain Camelid Antibodies. *Cell* **181**, 1004-1015 e1015, https://doi.org/10.1016/j.cell.2020.04.031 (2020).
22. Koenig, P. A. *et al*. Structure-guided multivalent nanobodies block SARS-CoV-2 infection and suppress mutational escape. *Science* **371**, https://doi.org/10.1126/science.abe6230 (2021).
23. Janssens, R. *et al*. Generation of heavy-chain-only antibodies in mice. *Proc Natl Acad Sci U S A* **103**, 15130-15135, https://doi.org/10.1073/pnas.0601108103 (2006).

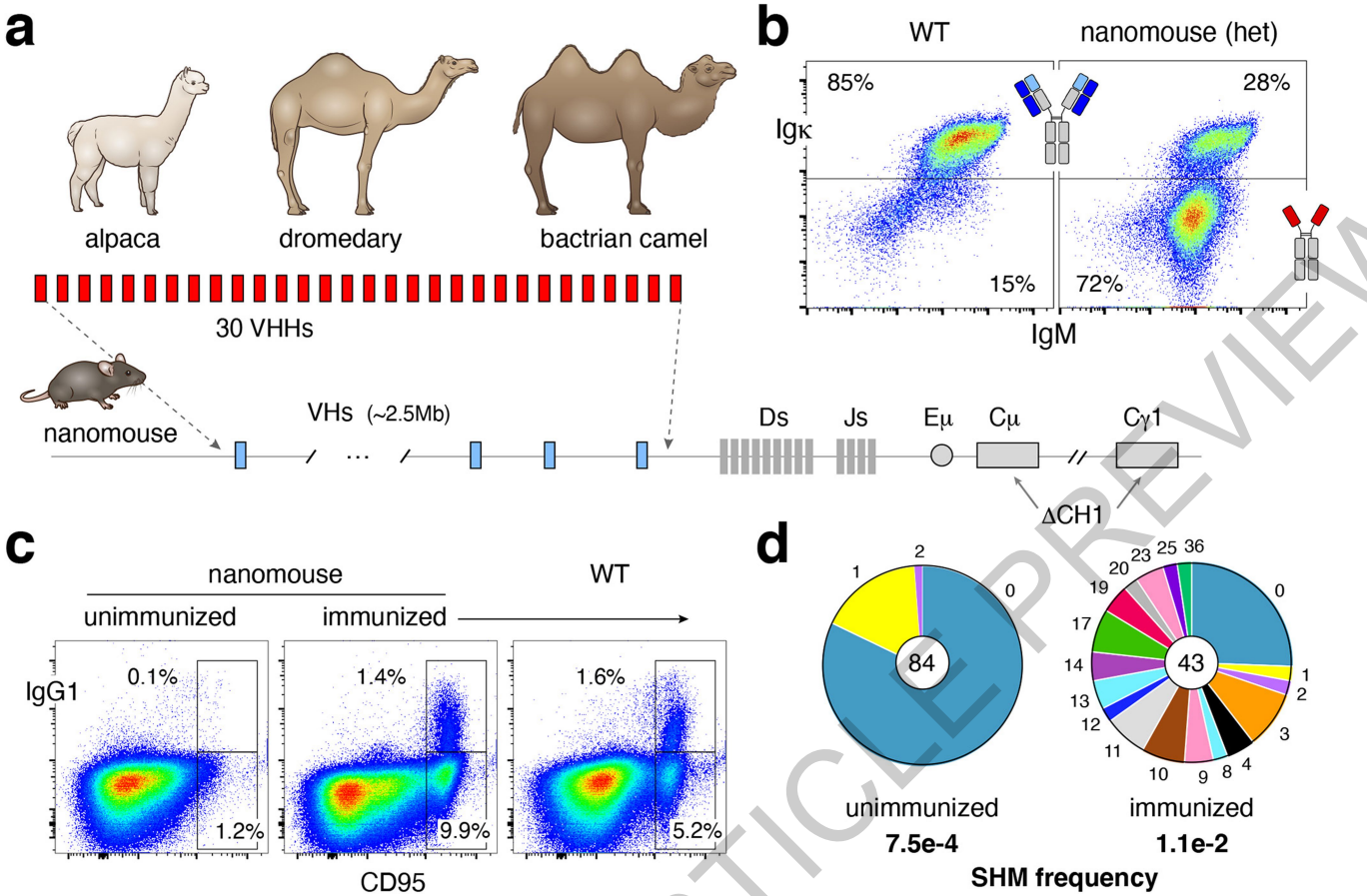
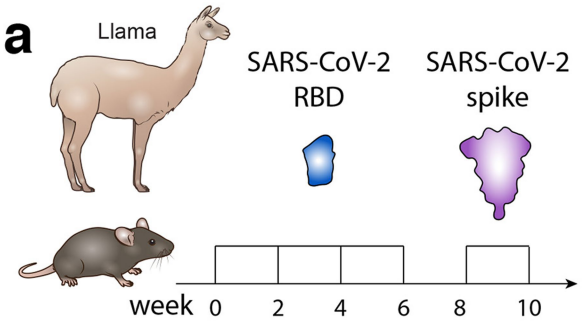
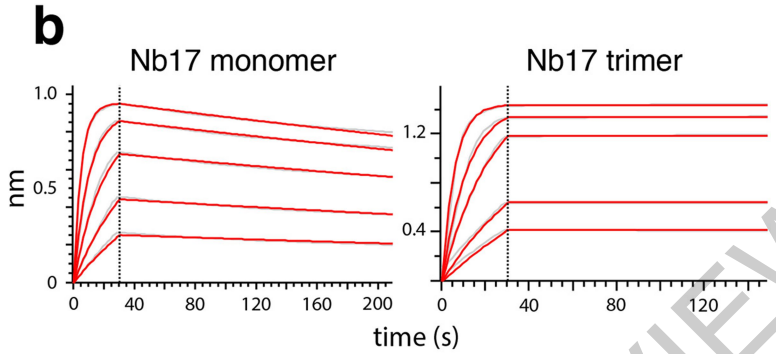


Fig. 1 | Production of nanomice. a, 30 VHHs selected from alpaca, dromedary and camel were inserted via CRISPR-Cas9 in lieu of the 2.5MB mouse VH locus. CH1 exons from Cμ and Cγ1 were also deleted to avoid misfolding of the antibody heavy chain. **b**, Flow cytometry analysis of splenic B220+ B cells from WT or heterozygous nanomice. IgM+Igκ+ represent cells expressing conventional heavy-light chain antibodies, whereas IgM+Igκ- are mostly Igλ+ in WT (not shown) or single-chain antibody B cells in nanomice. **c**, Flow cytometry analysis of splenic cells from unimmunized and immunized nanomice and controls stained with CD95 and IgG1. **d**, Pie charts showing VHH somatic hypermutation in unimmunized and immunized nanomice. Pie segments are proportional to the VHH sequences carrying the mutations indicated on the periphery of the chart. The middle circle shows the total number of sequences, while mutation frequency is shown in bold below.





c

		IC50	(ng/ml)
Nb12 mouse	monomer	11.7 nM	168.5
	bivalent-llama	0.7 nM	30.0
	bivalent-human	0.5 nM	21.6
	trimer-human	0.1 nM	7.7
	trimer-llama	65 pM	4.7
Nb15 llama	monomer	0.4 nM	6.2
	bivalent-human	0.3 nM	14.3
	trimer-human	0.1 nM	10.0
Nb17 llama	monomer	0.6 nM	8.0
	bivalent-human	79 pM	3.1
	trimer-human	18 pM	1.2
Nb19 llama	monomer	0.3 nM	4.6
	bivalent-human	66 pM	2.5
	trimer-human	9 pM	0.7
Nb30 mouse	monomer	6.9 nM	99.6
	bivalent-llama	1.9 nM	81.7
	bivalent-human	1.8 nM	70.7
Nb56 llama	monomer	0.9 nM	13.7
	bivalent-human	0.1 nM	4.6
	trimer-human	55 pM	3.9

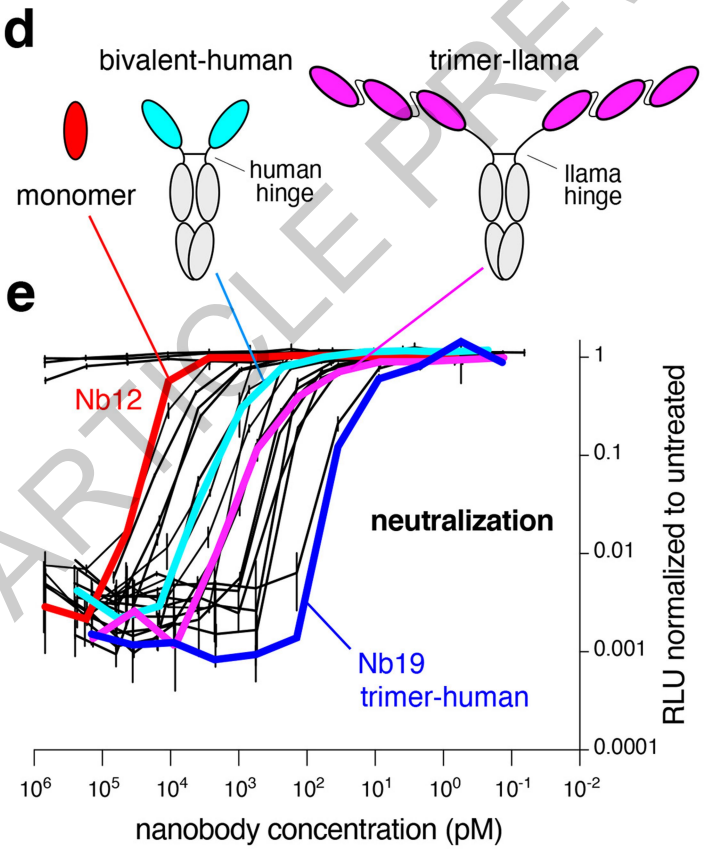


Fig. 2 | Isolation of Nbs against SARS-CoV-2. a, Immunization of llama and nanomice to obtain high affinity Nbs against SARS-CoV-2 RBD. **b**, BLI analysis of difference concentrations of Nb17 monomer and trimer binding to immobilized RBD. Red trace represents the raw data, and the kinetic fit is shown in grey underneath. Equilibrium (*KD*) constants are provided. **c**, Table summarizing pseudovirus neutralization potency (IC50) of selected Nbs. Values are provided in molarity or as ng/ml. **d**, Diagrams showing Nbs used in neutralization assays as monomers, bivalent or trimers (the last two fused to human IgG Fc via the human or llama hinge domain). **e**, Neutralization of SARS-CoV-2 pseudovirus by the 20 Nbs shown in panel **c**. Nb12 monomer (red), bivalent (cyan) and trimer (magenta), as well as Nb19 trimer (blue) are highlighted. Data are representative of two independent experiments and the error bars represent the mean ± s.d. of triplicates.

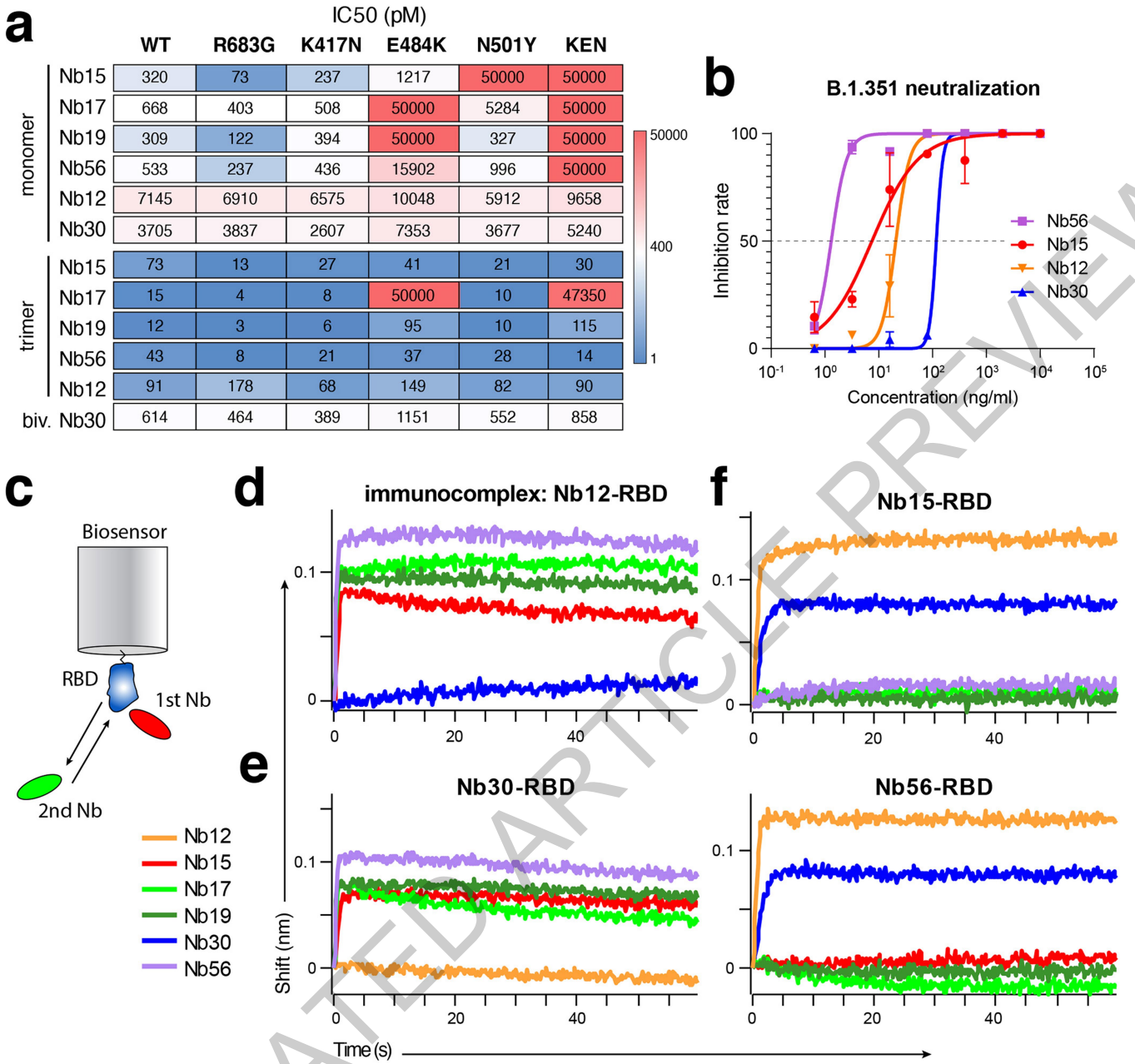


Fig. 3 | Neutralization of SARS-CoV-2 RBD WT and mutants. a, Neutralization assays (IC50 values) of pseudoviruses carrying WT or mutant SARS-CoV-2 spikes. Color gradient indicates values ranging from 0 (blue) to 50,000 pM (red). Pseudotyped viruses containing E484K or KEN also contain the R683G mutation. **b**, Neutralization assay showing the sensitivity of SARS-CoV-2 B.1.351 to different concentrations of trivalent Nb15, Nb56, Nb12, and bivalent Nb30. Data are representative of two independent experiments and the error bars represent the mean ± s.d. of triplicates. **c**, Schematics summarizing the BLI competition assay, where Nb-RBD immunocomplexes attached to a biosensor are incubated with different Nbs to measure binding. **d,e,f**, Binding of Nbs to Nb12-RBD, Nb30-RBD, Nb15-RBD and Nb56-RBD immunocomplexes respectively.

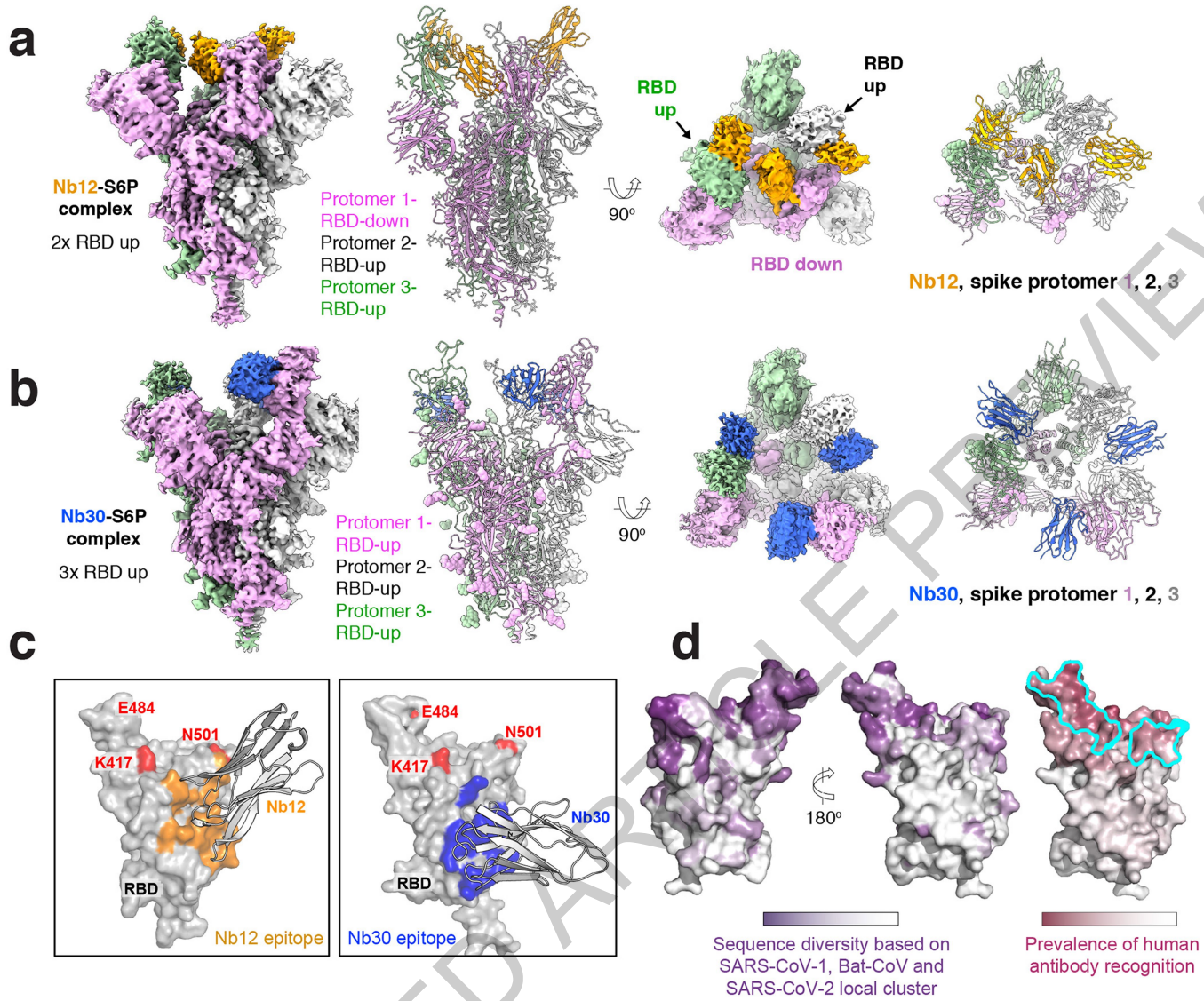


Fig. 4 | Structures of leading Nbs in complex with SARS-CoV-2 spike. **a**, Cryo-EM structure of Nb12 in complex with SARS-CoV-2 spike. **b**, Same as panel **a** for Nb30. **c**, Interface between Nbs and spike. **d**, Surface properties of RBD including sequence diversity (dark purple indicates diversity among Sarbecoviruses), and prevalence of RBD-recognized regions by human antibodies (dark raspberry indicates high prevalence) and binding site for ACE2 (cyan).

Methods

Data reporting

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Construction of exchange-cassette and VHH minigene

To replace the entire mouse VH locus (mm10, chr12:113,567,224-116,010,427) we assembled a targeting vector (pLH28-exchange-cassette) and a VHH minigene. The targeting vector was built by inserting a selection cassette composed of pEF1a-Puro-TK-2A-EGFP in between wtLoxP and LoxP257 sites within the pLH28 exchange vector[24]. As homology arms, 1Kb and 0.8Kb fragments flanking the VH deletion domain were cloned 5′ and 3′ of the LoxP sites. To build the VHH minigene, VHH genes (18 from alpaca, 7 from dromedary and 5 from camel) were selected based on published sequences[25–27]. 30 mouse VH promoters (250 bps) were next chosen based on their expression as measured by GRO-Seq in resting and activated mouse B cells. VHHs were codon optimized and complemented with mouse leading exons, introns, and recombination signal sequences (RSSs). The 30 units were pieced together by Gibson assembly (NEB) into the pBeloBAC11 vector.

ES cell targeting

E14 cells were cultured in Glasgow's MEM (Thermo Fisher Scientific, 11710035) supplemented with 10% FBS (ATCC, SCRR-03-2020), Glutamax, sodium pyruvate, NEAA, Pen(50 units/ml)-Strep(50 μg/ml) and β-mercaptoethanol (Thermo Fisher Scientific, 35050061, 11360070, 11140050, 15140122, 21985023, respectively) at 37 °C and 5% CO_2. mLIF (GeminiBio, 400-495, 10,000x), MEKi (Stemgent, 0400602, 10,000x) and GSKi (Stemgent, 0400402, 3,333x) were added to the medium before use. Dishes and plates were coated with 0.1% glycine for 15 minutes at room temperature before use. To delete the CH1 exon of Cμ, sgRNAs targeting the flanking introns were cloned into the CRISPR-Cas9 plasmid pX458 (Addgene, 48138). A 100nt long single-stranded oligonucleotide (ODN) donor (100μM, 3 μl) was co-transfected with the two Cas9-sgRNA plasmids (2 μg each) into ES cells (2 million cells) with the Amaxa nucleofection kit (Lonza, VPH-1001, program A030). After 24 hours of culture, GFP-high ES cells were FACS sorted and cultured in 10-cm dishes at a concentration of 2,000 cells per dish. 7 days later, colonies were transferred into 96-well plates and cultured for an additional 3 days. Genomic DNA was then extracted and genotyped for Cμ exon deletion. Clones with homozygous deletions were selected to next delete the Cγ1 exon with the same strategy. To delete the entire VH locus, sgRNAs targeting sequences upstream of Ighv1-86 (first Ighv) and downstream of Ighv5-1 (last Ighv) respectively were cloned into pX458. Selected ES cells (2 million cells) were co-transfected with the two Cas9-sgRNA plasmids (1.5 μg each) and the pLH28-exchange-cassette plasmid (1.5 μg) and then cultured in 10-cm dishes. 24 hours later, cells were selected with puromycin (0.8 μg/ml) for 10 days and individual colonies were picked for expansion and genotyping by long-range PCR. Positive clones (2 million cells) were co-transfected with VHH minigene vector (3 μg) and a Cre-expressing plasmid (1 μg) and cultured in 10-cm dish for 3 days. Cells were then selected with ganciclovir (2 μg/ml) for 7 days before individual colonies were picked for expansion and genotyping. sgRNAs and ODN primers are listed in Extended Data Table 4.

Generation of nanomice

Two modified ES cell clones with normal karyotype were injected into C57BL/6 blastocysts, which were then transferred to the uteri of pseudo-pregnant C56BL/6 recipients. High percentage chimeras were mated to C57BL/6 mice and offspring were genotyped for VHH minigene knock-in and Cμ and Cγ1 exon deletion. One out of three chimeras produced F1 offspring that showed germline transmission. Three F1 male mice were backcrossed with C57BL/6 mice. F2 heterozygous mice were inbred to produce mice homozygous for all three modifications. Two F1 offspring from the same chimera were Igμ$Ch1^{-/-}$ but WT for VH and CH1 exon of Cγ1. These mice were used as controls for Extended Data Fig. 3b.

Fluorescence activated cell sorting (FACS) analysis

B cells were activated by culturing them in RPMI 1640 supplemented with 10% FBS, HEPES, sodium pyruvate, NEAA, Pen/Str and β-mercaptoethanol at 37 °C and 5% CO_2 in the presence of lipopolysaccharide (LPS), interleukin-4 (IL-4) and αCD180 (1:2000, BD Pharmingen, 552128) antibody for 72 hours. For proliferation assays, cells were stained with CellTracer Violet (Thermo Fisher Scientific, C34557) at room temperature for 20 minutes before culturing for 96 hours. For all FACS staining, cells were incubated in FACS buffer (PBS, 2% FBS) at 4 °C for 20 minutes. Antibodies used for staining were: anti-B220-PerCP-Cy5.5 (1:500, eBioscience, 45-045-82), anti-B220-APC (1:500, Invitrogen, 17-0452-83), anti-IgM-APC (1:500, eBioscience, 17-5790-82), anti-Igκ-PE (1:500, BD Pharmingen, 559940), anti-Igκ-FITC (1:500, BD Pharmingen, 550003), anti-Igλ-FITC (1:200, BD Pharmingen, 553434), anti-IgG1-PE (1:200, BD Pharmingen, 550083), anti-IgG1-APC (1:200, BD Pharmingen, 550874), anti-IgD-FITC (1:200, BD Pharmingen, 553439), anti-CD95-PE (1:200, BD Pharmingen, 554258), anti-CD43-PE (1:200, BD Pharmingen, 553271), anti-CD23-PE (1:200, BD Pharmingen, 553139), anti-CD21-FITC (1:200, Biolegend, 123408), Viability Dye eFluor506 (1:1000, Invitrogen, 1923275). Data were acquired using BD FACSCanto and FACSDiva software and analyzed with FlowJo software. Gating strategy is shown in Extended Data Fig. 11d.

Analysis of VHH(D)J recombination

Genomic DNA from bone marrow or splenic samples was extracted with the DNeasy Blood & Tissue kit (Qiagen, 69506). VHH(D)J joints were PCR amplified from 100ng of DNA with a framework primer unique for each of the 30 VHHs, and a common downstream JH4 primer. PCR products were loaded onto 1% agarose gel to resolved them by size. Primers are listed in Extended Data Table 4.

VHH(D)J recombinants phagemid library construction

VHH(D)J phagemid libraries from unimmunized mice were constructed by first extracting RNA from nanomouse splenic samples with Trizol (Thermo Fisher Scientific, 15596026) and reverse transcribed to cDNA with SuperScript™ III (Thermo Fisher Scientific, 18080400) according to the manufacturer's instructions with some modifications.10 μg of total RNA was denatured and annealed with gene specific primers corresponding to Ighμ CH2 of gene. After elongation at 50 °C for 50 minutes, template switching oligonucleotide (TSO, 3′-propyl modified) linker was added to the 3′ end of the first strand cDNA with 90 minutes incubation at 42 °C. The reaction was inactivated at 85 °C for 5 minutes and 2 μl of cDNA was used as template for VHH(D)J amplification by two-step PCR with HiFi PCR Premix (Takara, 639298). For the first-step PCR, unmodified TSO and Ighμ CH2-specific oligonucleotides were used. 30 ng of the first-step PCR product was then amplified with a primer mix of 30 forward primers corresponding to framework (FR1) of 30 VHH genes and 4 reverse primers corresponding to JH1-JH4. pMES4 phagemid (Addgene, 98223) was amplified with primers to introduce SfiI sites on both ends. VHH(D)J and pMES4 fragments were then digested with SfiI (NEB, R0123L) and ligated (100 and 200 ng respectively) with T4 ligase (NEB, M0202L) at 16 °C overnight. Ligation product was purified with DNA Clean & Concentrator (Zymo Research, D4014) and eluted into 12 μl of water. 3 μl of DNA were electroporated into 60 μl of TG1 cells (Lucigen, 60502-2) in 1.0 mm cuvette (HARVARD Apparatus, 450134) with BTX electroporation system ECM 630 at the setting of 25 μF, 200 Ohms, 1600 Volts. After 1 hour recovery in 37 °C in shaker incubator, TG1 cells were plated on 5 of 10-cm LB Agar plates supplemented with 100 μg/ml carbenicillin (KD Medical, BPL-2400). Plates were placed in 37 °C bacteria incubator overnight and then bacteria scaped off plates and

phagemid library DNA extracted with Zymo Plasmid Miniprep kit (Zymo Research, D4054). Primers are listed in Extended Data Table 4.

Sanger sequencing for somatic hypermutation analysis

VHH(D)J recombinants from splenic cells of two nanomice were PCR amplified as described in the phagemid library construction section, and then cloned directly into pCR-Blunt II-TOPO vector (Thermo Fisher Scientific, 450245) and transformed into Stabl3 competent E. coli (Thermo Fisher Scientific, C737303). 96 colonies were randomly picked for Sanger sequencing. TG1 cells from BG505 DS-SOSIP immunized nanomouse phagemid library were plated onto carbenicillin-containing plates and 50 colonies picked for Sanger sequencing. Sequence alignment was performed using Snapgene software.

Immunizations

All animal related procedures were performed by following our NIAMS ACUC protocol. To monitor the germinal center reaction, three nanomice and two C57BL/6 mice were immunized intraperitoneally with 50 μg of keyhole limpet hemocyanin (KLH) in the presence of complete Freund's adjuvant (CFA). A boost injection was performed in the footpads with 25 μg of KLH in the presence of incomplete Freund's adjuvant (IFA) on day 6. Spleen samples were harvested on day 12 for analysis.

To isolate nanobodies recognizing HIV-1 envelop trimer, one nanomouse was immunized intraperitoneally with 50 μg of BG505 DS-SOSIP in the presence of CFA on day 0, and boost immunized with 25 μg of BG505 DS-SOSIP in the presence of IFA or phosphate-buffered saline (PBS) on day 22 and 44, respectively. Bone marrow, spleen and blood were harvested on day 48.

To isolate neutralizing nanobodies against SARS-CoV-2, a llama (Capralogics, Inc.) was immunized subcutaneously with 1 mg of recombinant RBD protein in the presence of CFA at day 0, and boost immunized with 0.5 mg of RBD protein in the presence of IFA on day 14, 28, 42. Two more boost immunizations with 0.5 mg of recombinant Spike protein in the presence of IFA were performed on day 56 and 70. On day 80, 500 ml of whole blood were collected for library preparation.

To isolate SARS-CoV-2 neutralizing nanobodies from nanomice, two groups of mice (five for group 1 and six for group 2) were immunized with RBD and/or Spike protein and bleeds were collected after a 62-day immunization protocol. Mice were immunized intraperitoneally with 50 μg of RBD protein (group 1) or Spike protein (group 2) in the presence of complete Freund's adjuvant (CFA) on day 0, and boost immunized intraperitoneally with 25 μg of RBD protein (group 1) or Spike protein (group 2) in the presence of IFA on day 14, 28 and 42. Mice were further immunized with 25 μg of Spike protein in PBS on day 56 and 59, intraperitoneally and intravenously, respectively. Bone marrow, spleen and blood samples were harvested on day 62. Best responders nanomouse 1 (group 1) and nanomice 2 and 3 (group 2) were selected for phage library construction.

Llama and nanomouse phage library construction

The llama phage library was constructed as previously described[28] with some modifications. Briefly, 300 ml of whole blood was collected from llama and peripheral blood mononuclear cells (PBMC) were enriched using Ficoll-Paque plus (GE Healthcare, 17-1440-03). 50 μg of extracted RNA were reverse transcribed to cDNA with random hexamers and 2.5 μl of cDNA was used for first round RT-PCR with gene-specific primers CALL001 and CALL002. The PCR reaction was repeated in 12 individual tubes with cDNA added into reactions separately. PCR fragments of about 700 bp long were gel purified and used as template (30 ng for each reaction, repeated in 12 individual tubes) for second round PCR with nested primers VHH-Back and VHH-For. PCR product from individual reactions were pooled and gel purified. Nanobody fragments and pMES4 phagemid were digested with PstI-HF and BstEII-HF restriction enzymes (NEB: R3140L, R3162L) and ligated (1 μg and 2 μg respectively) with T4 ligase at 16 °C overnight. Ligation product was column purified (into 12 μl of H2O) and electroporated into 360 μl of TG1 cells. After 1 hour of recovery in 37 °C in shaker incubator, cells were plated on 6 of 245mm x 245mm dish (Thermo Fisher Scientific, 431301) containing 2-YT Agar supplemented with 100 μg/ml carbenicillin and 2% (wt/vol) glucose. Plates were placed in 37 °C bacteria incubator overnight and then bacteria were scraped off of plates and archived as glycerol stocks. Cells were infected with VCSM13 helper phage (Agilent Technologies, 200251) followed by precipitation of culture supernatant with 20% polyethylene glycol 8000 (Sigma, 89510) /2.5M sodium chloride on ice to purify the nanobody phage particles. Phage particles were resuspended in 1 ml PBS, 300 μl were used for screening immediately and the remaining phages were stored in -80 °C in the presence of 10% glycerol.

Nanomouse nanobody phage libraries were constructed the same way as nanomouse VHH(D)J region phagemid library construction with some modifications. Briefly, total RNA was extracted from splenic cells, bone marrow cells and PBMC of immunized mice and processed separately until TG1 cell electroporation step. RNA from splenic cells, bone marrow and PBMCs (50 μg, 50 μg and all respectively) were reversed transcribed to cDNA with Ighγ1 CH2-specific primer in separate tubes. 2 μl of cDNA were used as template for PCR variable domain amplification (12 reactions each), using unmodified TSO and Ighγ1 CH2-specific oligonucleotide as primers. Second PCR was repeated in 12 reactions using 30 ng of the first-step PCR product as template and 30 FR1 and 4 JH oligonucleotide mix as primers. PCR products were gel purified, digested with SfiI and ligated with pMES4 (200 ng and 400 ng respectively). Ligation products from splenic cells, bone marrow and PBMC samples were pooled and column purified (into 12 μl of water) and electroporated into 360 μl of TG1 cells and phage libraries prepared as described above. Primers are listed in Extended Data Table 4.

Library construction for Illumina MiSeq deep sequencing

Phagemid DNA extracted from TG1 cell libraries was used as starting material for constructing MiSeq libraries to measure VHH usage and nanobody diversity. Briefly, 1.2 μg of phagemid DNA was used as template and VHH(D)J inserts were amplified with primers recognizing the pMES4 backbone using CloneAmp HiFi PCR Premix (Takara, 639298) in a 50 μl reaction (9 cycles). To avoid MiSeq failure due to low complexity at initial cycles and to enable multiplex sequencing, 1-9 nt long staggers were introduced into forward primers. Without purification, 5 μl of the first PCR product were used as template for a second PCR (9 cycles) to add Illumina P5 and P7 primers on both ends. PCR product was then loaded onto a 2% agarose gel and the ~580 bp size band was purified with Zymoclean Gel DNA Recovery Kit (Zymo Research, D4002). DNA concentration was determined by Qubit 4 Fluorometer (Thermo Fisher Scientific, Q33238) and average DNA size was determined by TapeStation 4150 (Agilent). DNA was then adjusted to 2 nM in elution buffer containing 0.1% Tween-20. For unimmunized nanomice VHH(D)J library, DNA (2nM) from 3 mice were mix at 1:1:1 ratio and loaded for MiSeq run. For immunized llama and nanomice nanobody diversity analysis, DNA (2nM) of pre-selection and post-selection libraries were mixed at 10:1 ratio first and then samples from individual animals were pooled at 1:1 ratio before loading for MiSeq sequencing. Primers are listed in Extended Data Table 4.

Deep sequencing analysis

For unimmunized nanomouse VHH usage analysis, pooled library from 3 mice were sequenced by MiSeq (pair end, 270 cycles x 2). Pair end reads were merged with NGmerge[29] with default settings. Nucleotides corresponding to pMES4 were trimmed using pTrimmer program[30], leaving clean VHH(D)J sequences in the merged reads. Reads with undetermined N nucleotides, low quality sequence or less than 300 nt in length were removed with the fastp program[31]. Fastq format sequences were converted to fasta format for further analysis. To calculate VHH usage, a BLAST database was built from a fasta format file vhh.exon.

fa containing the exon sequence of all 30 VHH genes, using BLAST+. VHH(D)Js were then aligned to VHH genes using igblast program[32]. Alignment output file was simplified to retain only sequence ID and VHH(D)J recombination information.

For immunized llama and nanomouse nanobody diversity analysis, in total 8 libraries (pre- and post-selection) were sequenced by MiSeq (pair end, 300 cycles x 2). The 3' end low-quality sequences were trimmed using the Sickle program (v1.33, available at https://github.com/najoshi/sickle.). For different libraries, the minimum length of trimmed sequence was adjusted based on the length of staggers in the primers used for library construction. Paired sequences were merged by flash program (v1.2.11)[33] and translated. To extract nanobody sequences and to locate CDR3 region, we used ANARCI program[34] to annotate VHH genes with IMGT numbering. Protein sequences with greater than or equal to 100 amino acids in total and greater than or equal to 1 amino acid in CDR3 region were extracted for further analysis. Enrichment of individual sequences were calculated by comparing their frequencies in pre- and post-selection libraries. Sequences that were enriched more than 10 times and had greater than or equal to 5e-05 frequency were selected for CDR3 clustering using cd-hit program (v4.6.8)[35].

Expression and purification of BG505 DS-SOSIP and SARS-CoV-2 proteins

BG505 DS-SOSIP protein was expressed and purified as described[36]. The spike protein of SARS-CoV-2 and its receptor binding domain (RBD) were expressed and purified as described[37,38] with some modifications. Briefly, 1 mg of pCAGGS-Spike or pCAGGS-RBD plasmid was transfected into 1 liter of Expi293 cells (Thermo Fisher Scientific, A14528) with Turbo293 transfection reagent (Speed Biosystem, PXX1002). Supernatants from transfected cells were harvested on day 4 post-transfection by centrifugation of the culture at 12,000 g for 15 min. Supernatant was then filtered through 0.2 m aPES filter (Thermo Fisher Scientific, 5670020) and incubated with 10 ml of cOmplete His-tag purification resin (Roche, 50434600) for 1 hour at room temperature. Next, His-tag resin was collected through gravity flow columns (BioRad, 9704652), washed with 100 ml of washing buffer (15 mM imidazole, 50 mM TrisHCl, 300 mM NaCl) and eluted with 25 ml of elution buffer (300 mM imidazole, 50 mM TrisHCl, 300 mM NaCl). Eluate was concentrated in 10 kDa Amicon Centrifugal Units (EMD Millipore, UFC901024) and then dialyzed in PBS using Slide-A-Lyzer dialysis cassette (Thermo Fisher Scientific, 66381). Proteins were analyzed by NuPAGE gel (Thermo Fisher Scientific, NP0336BOX) and visualized by InstantBlue staining (Abcam, ab119211). Soluble spike trimers or monomeric RBD proteins were aliquoted, snap frozen by liquid nitrogen and stored at -80 °C before being used for immunization. RBD and Spike (HexaPro) proteins used for phage screening, BLI, negative-stain and cryo-EM were done as previously described[10,39].

Phage screening for BG505 DS-SOSIP, RBD and Spike binding Nbs

RBD, Spike and BG505 DS-SOSIP were coated by different methods onto MaxiSorp 96-well plate (Thermo Fisher Scientific, 439454) for phage screening. For RBD screening, two wells were coated with 50 µl of RBD protein (100 µg/ml in PBS) in 4 °C overnight. Another well with 50 µl of PBS was used as an un-coated control. Wells were washed with PBS with 0.1% Tween-20 three times and blocked with 5% non-fat milk in PBS at room temperature for 1 hour. For Spike or BG505 DS-SOSIP screening, three wells were coated with 50 µl of lectin (EMD Millipore, L8275, 100 µg/ml in PBS) in 4 °C overnight. Wells were washed and blocked with 10% non-fat milk in room temperature for 1 hour. After three washes, 50 µl of 100 µg/ml of BG505 DS-SOSIP or Spike were added to 2 wells, incubated at room temperature for 2 hours and washed. A third well contained PBS and served as an un-coated control. 300 µl of phage particles was mixed with 300 µl of 10% non-fat milk and rotated gently at room temperature for 1 hour. 150 µl of blocked phage particles were then added into each well and incubated in room temperature for 2 hours with gentle shaking. After 15 washes, phages were eluted with TrypLE Express Enzyme (Thermo Fisher Scientific, 12605010) by shaking plates at 700 rpm at room temperature for 30 minutes and used immediately for selection efficiency estimation (10 µl of phage eluate) and recovery infection (the remaining eluate) as described[28]. Anti-RBD libraries were selected with RBD protein once, and libraries constructed from BG505 DS-SOSIP or Spike immunized animals were selected with BG505 DS-SOSIP or Spike (HexaPro) proteins twice.

ELISA selection of anti-BG505 DS-SOSIP and anti-RBD Nbs

After one or two rounds of selection, recovered TG-1 cells were plated and colonies were picked to prepare periplasmic extracts containing crude nanobodies for ELISA. Briefly, individual colonies were picked and grown in 96 deep-well plates (Thermo Fisher Scientific, 278743) in 2YT medium supplemented with 100 µg/ml of Carbenicillin and 0.1% Glucose. IPTG (final 1mM) was added when OD600 reached 1 and protein expression was induced in 30 °C for 16 hours. Periplasmic extracts were prepared by resuspending bacteria pellet in 200 µl of PBS and rapidly frozen in liquid nitrogen. Frozen cells were thawed slowly at room temperature and centrifuged at 4100 g for 15 minutes. Maxisorp plates were coated with lectin (2 µg/ml) followed by BG505 DS-SOSIP (2 µg/ml) or with RBD (2 µg/ml). After blocking, 100 µl of Nb-containing supernatant were transferred to the plates and incubated for 2 hours at room temperature. Plates were washed and then incubated with horse radish peroxidase (HRP) conjugated goat anti-alpaca VHH domain specific antibody (Jackson ImmunoResearch, 128-035-232) for 1 hour at room temperature. Plates were washed and then developed by addition of 50 µl of tetramethylbenzidine (TMB) (Thermo Fisher Scientific, 34028) for 10 minutes, then the reaction was stopped by adding 50 µl of 1M H_2SO_4. Absorbance at 450 nm was measured with Synergy microplate reader (BioTek Gen5).

Expression and purification of nanobodies

Phagemids from lead candidates identified by ELISA were extracted from TG-1 cells and transformed into WK6 cells (ATCC, 47078). Cultures were grown in 30ml of 2YT medium (100 µg/ml of carbenicillin and 0.1% glucose) at 37 °C and 220rpm until OD600 reached 1. Protein expression was induced by 1mM IPTG at 30 °C for 16 hours and then pelleted at 4100 g for 15 minutes. The resulting pellets were resuspended in 1 ml of PBS plus 30 µl of 0.5 MU/ml polymyxin B (Sigma, P1004) and incubated at 37 °C with shaking for 1 hour. Cell debris were pelleted at 12,000 g for 5 minutes and nanobodies in the supernatant were purified using Capturem His-tagged purification kit (Takara, 635710). For larger scale of nanobody production (0.2 to 1 liter of culture), nanobodies in the supernatant were purified by cOmplete His-tag purification resin and dialyzed in PBS as described above. Proteins were filtered sterile by 0.22 µm PVDF membrane (EMD Millipore, UFC30GVNB) before used for downstream assay.

Expression and purification of Fc conjugated nanobodies and RBD in Expi293 cells

Monomeric or trimeric nanobody sequences were fused to the Fc region of human IgG1 with 6xHis tag at the C terminal end and cloned into the pVRC8400 vector. In trimeric form, nanobody units were connected through (GGGGS)x3 flexible linkers. In some cases, llama IgG2a hinge region was used in lieu of human IgG1 hinge. The Fc-fusion constructs were expressed in Expi293 cells as described above at 33 °C from day 2 to day 4. Antibodies in the supernatant were purified using either His-tag (Roche, 05893801001) or protein A (Thermo Fisher Scientific, A26457). When protein A resin was used, antibodies were eluted by IgG elution buffer (Thermo Fisher Scientific, 21009) and brought to neutral pH by adding 1/10 volume of Tris-HCl (1M, pH 8). Antibodies were concentrated, dialyzed and filtered. Nb-Fc yields were up to 100 mg/L. RBD region of SARS-CoV-1, SARS-CoV-2 and BatCoV-WIV16 spike protein were fused to the Fc region of human IgG1 and cloned

into pVRC8400 vector. RBD-Fc proteins were expressed in Expi293 cells and purified with protein A.

SARS-CoV-2 surrogate virus neutralization test (sVNT)

RBD-ACE2 interaction blocking potential of nanobodies was tested using the SARS-CoV-2 surrogate virus neutralization test kit (Genscript, L00847) according to the manufacturer's instructions. Briefly, HRP-RBD was diluted and incubated with specified concentrations of nanobodies for 30 minutes at 37 °C. Samples were then transferred onto ACE2-coated plates and incubated for 15 minutes at 37 °C. Plates were washed, and the assay was developed using TMB reagent and quenched with stop solution. Absorbance at 450 nm was measured with a Synergy microplate reader (BioTek Gen5). Inhibition rate was calculated and plotted using Microsoft Excel according to manufacturer's instruction of the sVNT kit.

Pseudotyped virus neutralization assay

A panel of plasmids expressing RBD-mutant SARS-CoV-2 spike proteins in the context of pSARS-CoV-2-SD19 have been described previously[1,40]. The mutants E484K and KEN (K417N+E484K+N501Y) were constructed in the context of a pSARS-CoV-2-$S_{\Delta 19}$ variant with a mutation in the furin cleavage site (R683G). The IC50 of these pseudotypes were compared to a wildtype SARS-CoV-2 spike sequence carrying R683G in the subsequent analyses, as appropriate. Generation of SARS-CoV-2 pseudotyped HIV-1 particles and pseudovirus neutralization assay was performed as previously described[11]. Briefly, 293T cells were transfected with pNL4-3DEnv-nanoluc and pSARS-CoV-2-SD19 and pseudotyped virus stocks were harvested 48 hours after transfection, filtered and stored at -80 °C. Serially diluted nanobodies were incubated with the pseudotyped virus for 1 h at 37 °C. The mixture was added to $293T_{ACE2}$[11] (for analysis of WT neutralization activity, Figure 2) or HT1080Ace2 cl.14[17] cells (for analysis of spike mutant panel, Figure 3) and after 48 hours cells were washed with PBS and lysed with Luciferase Cell Culture Lysis 5x reagent (Promega). Nanoluc Luciferase activity in lysates was measured using the Nano-Glo Luciferase Assay System (Promega) with Modulus II Microplate Reader User interface (TURNER BioSystems). The relative luminescence units were normalized to those derived from cells infected with SARS-CoV-2 pseudotyped virus in the absence of antibodies. Neutralization of HIV-1-based SARS-CoV-1 and bat coronavirus BatCov-WIV16 pseudotypes were performed in HT1080/ACE2cl.14 cells as previously described[41]. The half-maximal inhibitory concentration for nanobodies (IC50) was determined using four-parameter nonlinear regression (GraphPad Prism).

Recombinant Indiana VSV (rVSV) expressing different coronavirus spikes (SARS-CoV-2, RaTG13, GDPangolin, GXPangolin, SARS-CoV-1, WIV1, SHC014, LYRa11, Rs7327, Rs4084, Rs4231) were generated as previously described[2,42,43]. Briefly, HEK293T cells were grown to 80% confluency before transfection with the spike gene using Lipofectamine 3000 (Invitrogen). Cells were cultured overnight at 37 °C with 5% CO_2, and VSV-G pseudo-typed ΔG-luciferase (G*ΔG-luciferase, Kerafast) was used to infect the cells in DMEM at an MOI of 3 for 2 hours before washing the cells with 1X DPBS three times. The next day, the transfection supernatant was harvested and clarified by centrifugation at 300 g for 10 min. Each viral stock was then incubated with 20% I1 hybridoma (anti-VSV-G, ATCC: CRL-2700) supernatant for 1 hour at 37 °C to neutralize contaminating VSV-G pseudo-typed ΔG-luciferase virus before measuring titers and making aliquots to be stored at −80 °C. Neutralization assays were performed by incubating pseudoviruses with serial dilutions of antibodies and scored by the reduction in luciferase gene expression as previously described[2,42,43]. Briefly, 293T-ACE2 cells were seeded in 96-well plates (2×10^4 cells per well). Pseudoviruses were incubated with serial dilutions of the antibodies in triplicate for 30 min at 37 °C. The mixture was added to cultured cells and incubated for an additional 16 hrs. Luminescence was measured using Luciferase Assay System (Promega), and IC_{50} was defined as the dilution at which the relative light units were reduced by 50% compared with the virus control wells (virus + cells) after subtraction of the background in the control groups with cells only. The IC50 values were calculated using a five-parameter dose-response curve in GraphPad Prism.

Authentic SARS-CoV-2 Microplate Neutralization

The SARS-CoV-2 viruses USA-WA1/2020 (WA1), USA/CA_CDC_5574/2020 (B.1.1.7), hCoV-19/South Africa/KRISP-EC-K005321/2020 (B.1.351), and hCoV-19/Japan/TY7-503/2021 (P.1) were obtained from BEI Resources (NIAID, NIH) and propagated for one passage using Vero-E6 cells. Virus infectious titer was determined by an end-point dilution and cytopathic effect (CPE) assay on Vero-E6 cells as described previously[42]. An end-point dilution microplate neutralization assay was performed to measure the neutralization activity of nanobodies. In brief, nanobodies were subjected to successive 5-fold dilutions starting from 10μg/ml. Triplicates of each dilution were incubated with SARS-CoV-2 at an MOI of 0.1 in EMEM with 7.5% inactivated fetal calf serum (FCS) for 1 hour at 37 °C. Post incubation, the virus-nanobody mixture was transferred onto a monolayer of Vero-E6 cells grown overnight. The cells were incubated with the mixture for ~70 hours. CPE of viral infection was visually scored for each well in a blinded fashion by two independent observers. The results were then reported as percentage of neutralization at a given nanobody dilution. The half-maximal inhibitory concentration for nanobodies (IC50) was determined using nonlinear regression (normalized response, variable slope) in GraphPad Prism.

Nanobody stability studies

Nanobody was nebulized with a portable mesh nebulizer (Philips, InnoSpire Go) producing 2-5μm particles at a final concentration of 0.4 mg/ml. The resulting aerosol was collected by condensation into a 50 ml tube cooled on ice. Pre- and post-nebulization samples were analyzed by NuPAGE gel and visualized by InstantBlue staining. SARS-CoV-2 surrogate virus neutralization test was also performed to compare the neutralization potency of pre- and post-nebulization samples. For thermostability tests, nanobodies supplemented with loading buffer (Thermo Fisher Scientifc, NP0007) and β-mercaptoethanol were heated at 98 °C for 10 minutes and then analyzed on a NuPAGE gel and visualized by InstantBlue staining.

BLI assay to measure nanobody affinity

The BLI assay was performed using a fortéBio Octet Red384 instrument to determine the affinity of nanobodies to RBD. Briefly, biotinylated-RBD was immobilized onto streptavidin coated biosensors and then dipped into a solution containing the Nb for 30 seconds followed by dissociation for 2-3 minutes. To assay the binding of nanobodies to RBD-Fc (SARS-CoV-1, SARS-CoV-2, BatCoV-WIV16), 6xHis-tagged nanobody was immobilized onto Ni-NTA coated biosensors and then dip into RBD-Fc solution for association for 1 minute followed by dissociation for 1 minute. Sensorgrams of the concentration series were corrected with corresponding blank curves and fitted globally with Octet evaluation software using a 1:1 Langmuir model of binding.

Nanobody-RBD binding competition assay

Nanobody-RBD binding competition assay was performed using a fortéBio Octet Red384 instrument. Biotinylated-RBD was first immobilized onto streptavidin coated biosensors and allowed to associate with one of the six nanobodies, then dipped into a solution contained a second nanobody.

Negative-staining EM analysis for the structure of nanobody-spike complex

Nanobody-spike complexes were prepared by mixing the two proteins at a 1:1 weight ratio, then diluted with a buffer containing 10 mM HEPES, pH 7.4, 150 mM NaCl, adsorbed to a freshly glow-discharged carbon-coated copper grid, washed with the above buffer, and stained

with 0.75% uranyl formate. Images were collected at a magnification of 57,000 using EPU on a Thermo Fisher Talos F200C microscope equipped with a 4k x 4k CETA 16M camera and operated at 200 kV. The pixel size was 2.5 Å for the CETA camera. Particles picking, reference-free 2D classification, 3D reconstruction and refinement were performed using cryoSPARC.

Cryo-EM data collection and processing

Nanobody-spike complexes (Nb12-S6P and Nb30-S6P) were prepared by manual mixture of the two proteins in a 1:1 weight ratio, then diluted to a final concentration of 0.5 mg/ml. Samplea (2.7 µl) were applied to a glow-discharged Quantifoil R 2/2 gold grids and vitrified using a Vitrobot Mark IV with a blot time of 3 s before the grid was plunged into liquid ethane. Data were acquired using the Leginon system installed on Titan Krios electron microscopes operating at 300kV and equipped with a K3-BioQuantum direct detection device. The dose was fractionated over 40 raw frames and collected over a 2 s exposure time. Motion correction, CTF estimation, particle picking, 2D classifications, ab initio model generation, heterogeneous refinements, 3D variability analysis and homogeneous 3D refinements were carried out with cryoSPARC. Local refinement was performed to resolve the RBD-Nb interface by using a mask encompassing one copy of the RBD-Nb complex for refinement, after removing the rest density by particle subtraction.

Cryo-EM model fitting

For initial fits to the cryo-EM reconstructed maps, we used the coordinates of the SARS-CoV-2 spike from PDB ID 7JZL, and a nanobody model predicted by the ABodyBuilder server[44]. These initial models were docked into the cryo-EM maps using Chimera. The coordinates were then fit to the electron density more precisely through an iterative process of manual fitting using Coot and real space refinement within Phenix, Molprobity and EMRinger were used to check geometry and evaluate structures at each iteration step. Figures were generated in UCSF ChimeraX and PyMOL (https://pymol.org). Map-fitting cross correlations were calculated using Fit-in-Map feature in UCSF Chimera. Overall and local resolution of cryo-EM maps was determined using cryoSPARC.

Informatics analysis

Sequence entropy are based on 9 strains with the following uniprot ID: SARS-Cov-2: P0DTD1, B.1.1.7, B.1.351; SARS-CoV: A7J8L4, Q202E5, P59594; and Bat SARS-like coronavirus: MG772933, WIV16 (A0A0U2IWM2), RsSHC014 (U5WLK5). The entropy was calculated for each residue based on aligned sequences with the formula:

$\text{Entropy} = -\sum_{i=1}^{21} p(xi)\log(p(xi))$, where x_i are standard amino acids, plus gap.

The buried surface area on the RBD were calculated for 51 human antibody/SARS-CoV-2 RBD complexes based using the Naccess program.

Data presentation

Figures arranged in Adobe Illustrator 2020.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

Data are provided in Extended Data Tables 1-4. The accession numbers for the deep-sequencing data reported in this paper are found at GSE167310. Coordinates and maps for reported cryo-EM structures have been deposited in the Electron Microscopy Data Bank and Protein Data Bank. Public accession codes are EMD-24078, EMD-24077, PDB-7MY3 and PDB-7MY2.

24. Han, L., Masani, S. & Yu, K. Overlapping activation-induced cytidine deaminase hotspot motifs in Ig class-switch recombination. *Proc Natl Acad Sci U S A* **108**, 11584-11589, https://doi.org/10.1073/pnas.1018726108 (2011).
25. Achour, I. *et al*. Tetrameric and homodimeric camelid IgGs originate from the same IgH locus. *J Immunol* **181**, 2001-2009, https://doi.org/10.4049/jimmunol.181.3.2001 (2008).
26. Bactrian Camels Genome, S. *et al*. Genome sequences of wild and domestic bactrian camels. *Nat Commun* **3**, 1202, https://doi.org/10.1038/ncomms2192 (2012).
27. Wu, H. *et al*. Camelid genomes reveal evolution and adaptation to desert environments. *Nat Commun* **5**, 5188, https://doi.org/10.1038/ncomms6188 (2014).
28. Pardon, E. *et al*. A general protocol for the generation of Nanobodies for structural biology. *Nat Protoc* **9**, 674-693, https://doi.org/10.1038/nprot.2014.039 (2014).
29. Gaspar, J. M. NGmerge: merging paired-end reads via novel empirically-derived models of sequencing errors. *BMC Bioinformatics* **19**, 536, https://doi.org/10.1186/s12859-018-2579-2 (2018).
30. Zhang, X. *et al*. pTrimmer: An efficient tool to trim primers of multiplex deep sequencing data. *BMC Bioinformatics* **20**, 236, https://doi.org/10.1186/s12859-019-2854-x (2019).
31. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884-i890, https://doi.org/10.1093/bioinformatics/bty560 (2018).
32. Ye, J., Ma, N., Madden, T. L. & Ostell, J. M. IgBLAST: an immunoglobulin variable domain sequence analysis tool. *Nucleic Acids Res* **41**, W34-40, https://doi.org/10.1093/nar/gkt382 (2013).
33. Magoc, T. & Salzberg, S. L. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**, 2957-2963, https://doi.org/10.1093/bioinformatics/btr507 (2011).
34. Dunbar, J. & Deane, C. M. ANARCI: antigen receptor numbering and receptor classification. *Bioinformatics* **32**, 298-300, https://doi.org/10.1093/bioinformatics/btv552 (2016).
35. Li, W. & Godzik, A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658-1659, https://doi.org/10.1093/bioinformatics/btl158 (2006).
36. Chuang, G. Y. *et al*. Structure-Based Design of a Soluble Prefusion-Closed HIV-1 Env Trimer with Reduced CD4 Affinity and Improved Immunogenicity. *J Virol* **91**, https://doi.org/10.1128/JVI.02268-16 (2017).
37. Amanat, F. *et al*. A serological assay to detect SARS-CoV-2 seroconversion in humans. *Nat Med* **26**, 1033-1036, https://doi.org/10.1038/s41591-020-0913-5 (2020).
38. Stadlbauer, D. *et al*. SARS-CoV-2 Seroconversion in Humans: A Detailed Protocol for a Serological Assay, Antigen Production, and Test Setup. *Curr Protoc Microbiol* **57**, e100, https://doi.org/10.1002/cpmc.100 (2020).
39. Zhou, T. *et al*. Structure-Based Design with Tag-Based Purification and In-Process Biotinylation Enable Streamlined Development of SARS-CoV-2 Spike Molecular Probes. *SSRN*, 3639618, https://doi.org/10.2139/ssrn.3639618 (2020).
40. Weisblum, Y. *et al*. Escape from neutralizing antibodies by SARS-CoV-2 spike protein variants. *Elife* **9**, https://doi.org/10.7554/eLife.61312 (2020).
41. Muecksch, F. *et al*. Development of potency, breadth and resilience to viral escape mutations in SARS-CoV-2 neutralizing antibodies. *bioRxiv*, https://doi.org/10.1101/2021.03.07.434227 (2021).
42. Liu, L. *et al*. Potent neutralizing antibodies against multiple epitopes on SARS-CoV-2 spike. *Nature* **584**, 450-456, https://doi.org/10.1038/s41586-020-2571-7 (2020).
43. Wang, P. *et al*. SARS-CoV-2 neutralizing antibody responses are more robust in patients with severe disease. *Emerg Microbes Infect* **9**, 2091-2093, https://doi.org/10.1080/22221751.2020.1823890 (2020).
44. Leem, J., Dunbar, J., Georges, G., Shi, J. & Deane, C. M. ABodyBuilder: Automated antibody structure prediction with data-driven accuracy estimation. *MAbs* **8**, 1259-1268, https://doi.org/10.1080/19420862.2016.1205773 (2016).

Acknowledgements We thank Stefania Dell'Orso for assistance with deep-sequencing; Jim Simone, Kevin Tinsley, and Jeff Lay for sorting; Megan Laycock for proofreading the manuscript; Patrick and Nai-Ying Wilson for plasmids and reagents; Ethan Tyler for camelid and mouse drawings; Nancy Wong for ES cell karyotyping; Eileen Southon for ES cell injections; the Vaccine Production Program of the Vaccine Research Center, for BG505 DS-SOSIP; and members of the structural biology section, Vaccine Research Center for discussions and manuscript feedback. This work was supported by NIAMS and NCI funding and the *NIH Helix Systems* (http://helix.nih.gov). This work was supported by the intramural research program of the Vaccine Research Center, National Institute of Allergy and Infectious Diseases (NIAID), Federal funds from the Frederick National Laboratory for Cancer Research under Contract HHSN261200800001E (T.S., Y.T.). Cryo-EM datasets were collected at the National CryoEM Facility (NCEF) of the National Cancer Institute under contract HSSN261200800001E. We are especially grateful to U. Baxa, A. Wier, M. Hutchison, and T. Edwards of NCEF for collecting cryo-EM data and for technical assistance with cryo-EM data processing. MCN and PDB are HHMI Investigators.

Author contributions J.X., K.X., A.C., J.L., F.M., J.C.C.L., S.P., F.S., Z.W., Y.H., Y.L., M.N., P.W., and J.E.S. performed experiments. T.B., G-Y.C., J.X., and S.J. performed bioinformatic analyses. A.S.O., I-T.T., and T.Z. contributed critical reagents. T.S. and Y.T. assisted with cryo-EM samples and data collection. L.T., J.X., K.X., V.M., D.D.H., T.H., P.D.B., M.C.N., P.D.K. and R.C. designed experiments and discussed results. J.X., K.X., P.D.K. and R.C. wrote the manuscript.

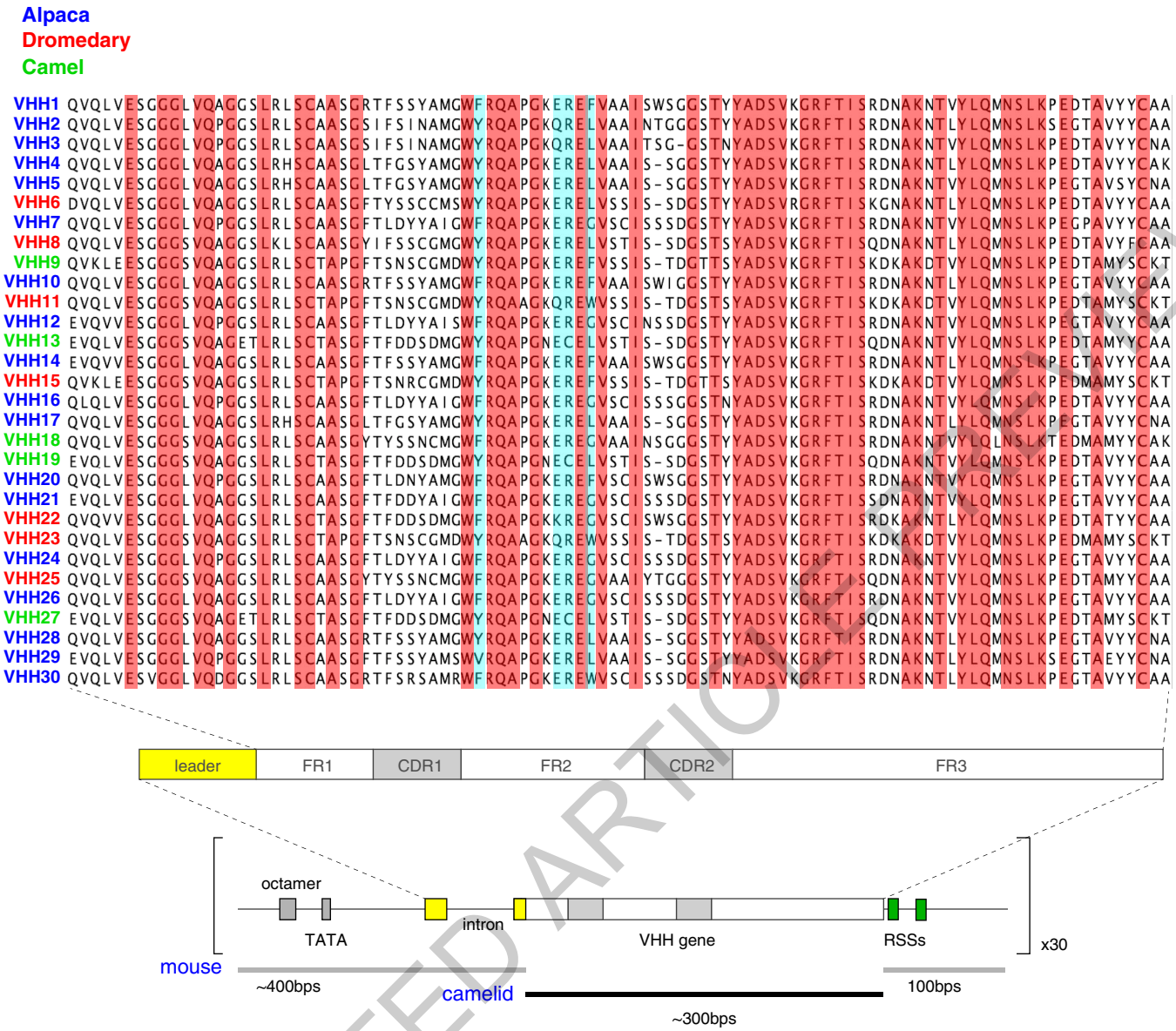
Competing interests The National Institutes of Health has filed a provisional patent application in connection with this work on which J.X. and R.C. are inventors (US patent 63-151,530).

Additional information

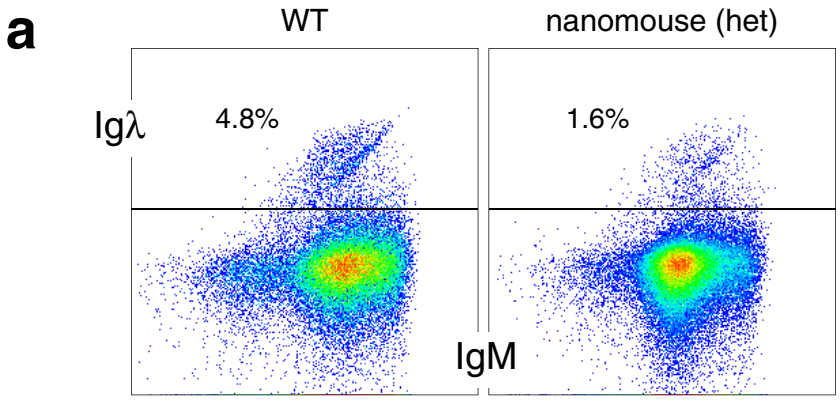
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-021-03676-z.

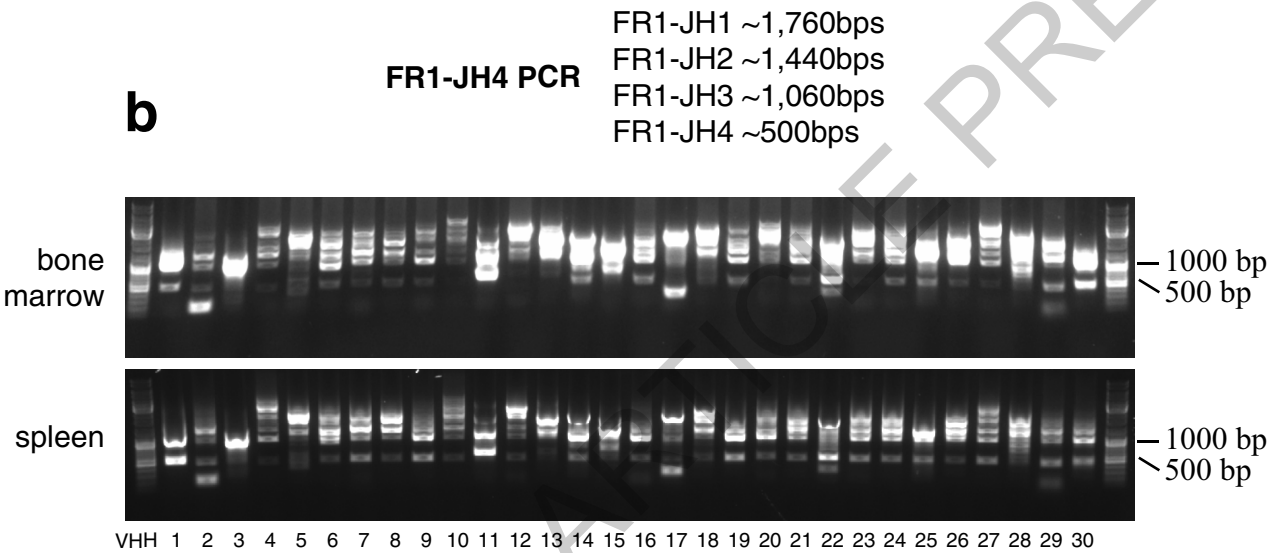
Correspondence and requests for materials should be addressed to J.X., M.C.N., P.D.K. or R.C.

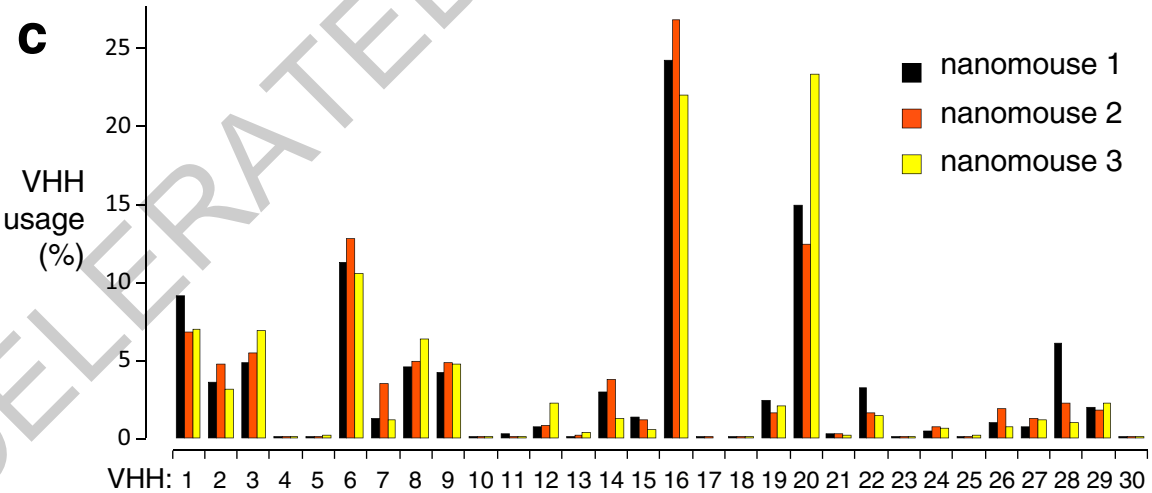
Peer review information *Nature* thanks Patrick Wilson and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.



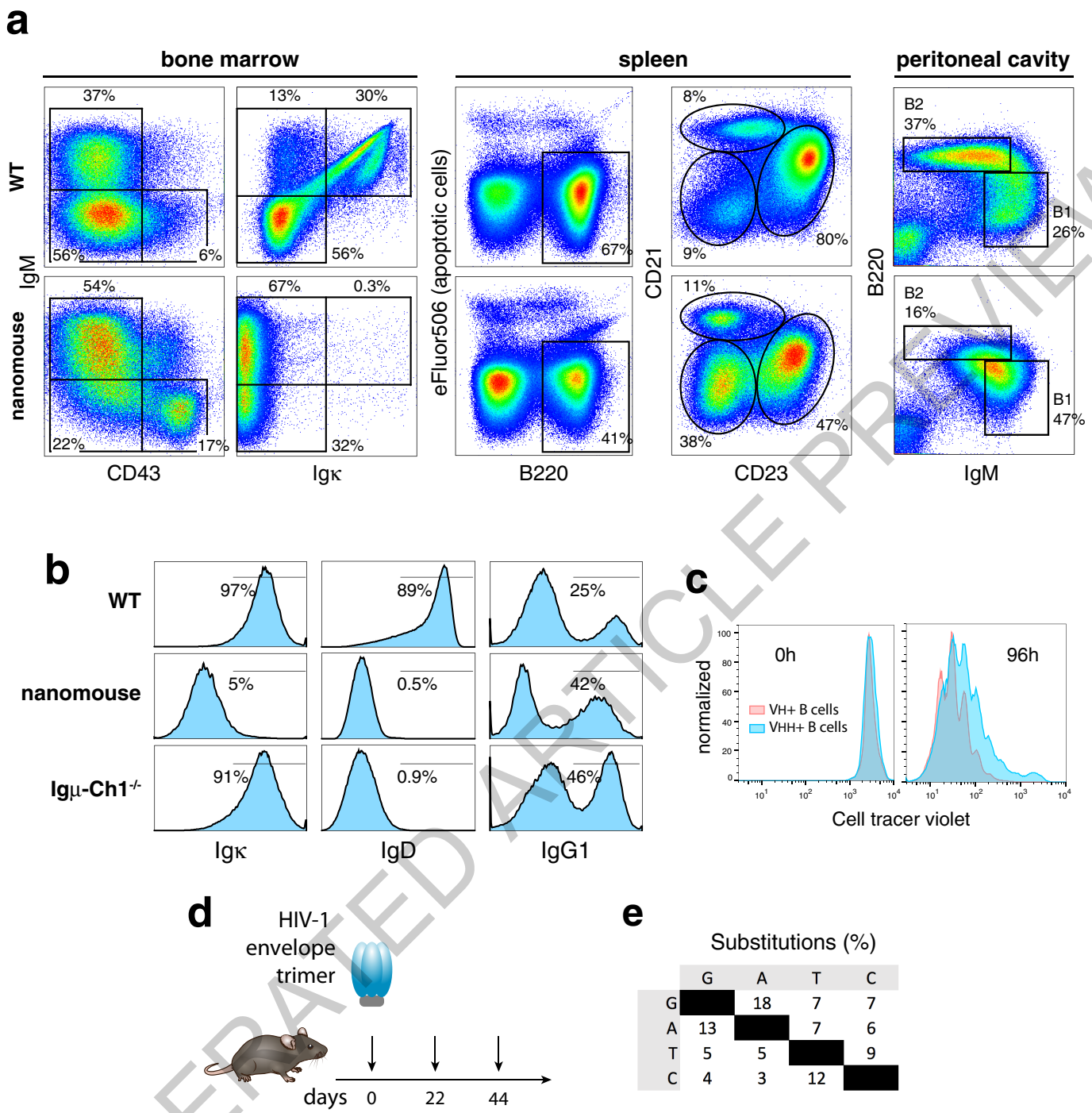
Extended Data Fig. 1 | VHH genes used in the array and gene unit assembly. Alignment of the 30 VHH genes highlighting in red 100% AA conservation and in blue the 4 hydrophilic AAs in FR2 [in VH proteins, those 4 AAs are hydrophobic and mediate the interaction with light chains]. Schematics below show the configuration of VHH gene units, composed of a mouse VH promoter (250 bps containing the octamer and TATA box); mouse leader exons-intron (~150 bp) codifying for the signal peptide cleaved off during heavy chain processing in the ER; the camelid VHH open reading frame (ORF, ~300 bps); mouse downstream sequences (100 bps) containing the recombination signal sequences (RSSs).





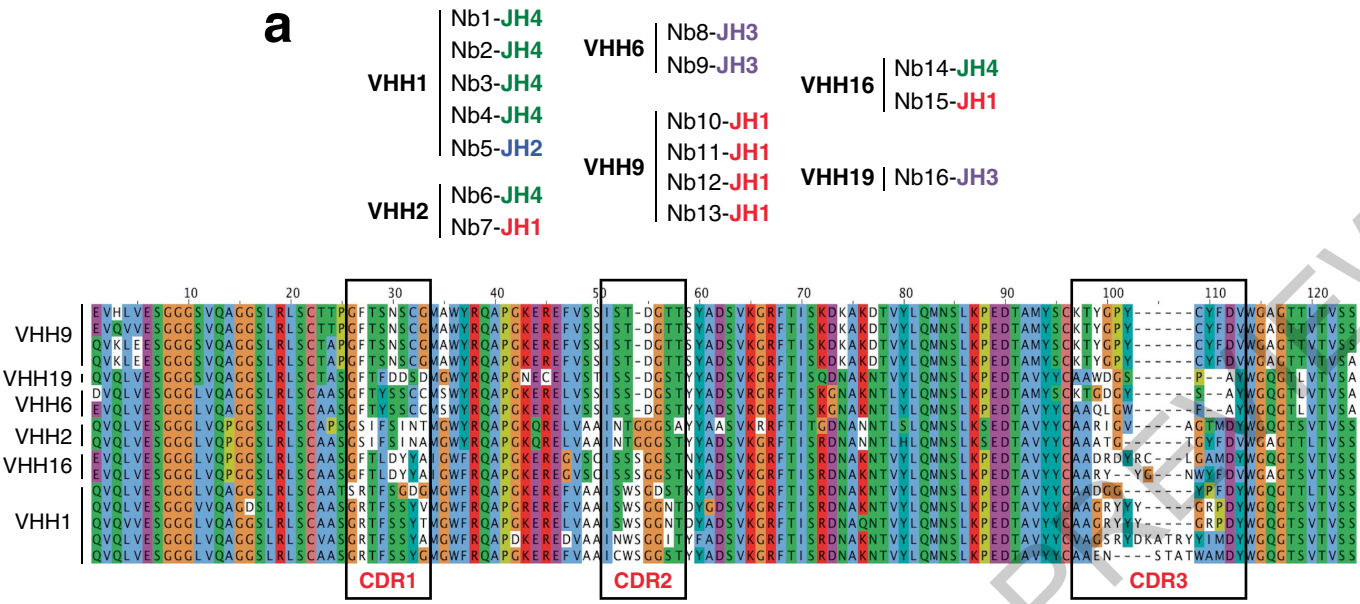


Extended Data Fig. 2 | Igλ expression and recombination frequency of VHH genes. a, Flow cytometry analysis of Igλ expression in B220^{+}IgM^{+} splenic B cells from wild type and heterozygous nanomice. **b**, VHH-DJ recombination was monitored by genomic PCR in bone marrow and spleen samples using a framework 1 (FR1) VHH-specific primer and a second primer downstream of JH4. The expected PCR products for each recombination event between a given VHH and JH1, JH2, JH3, or JH4 are provided. Data are representative of two independent experiments. **c**, Bar graph showing VHH percentage usage among splenic B cells in three nanomouse littermates.



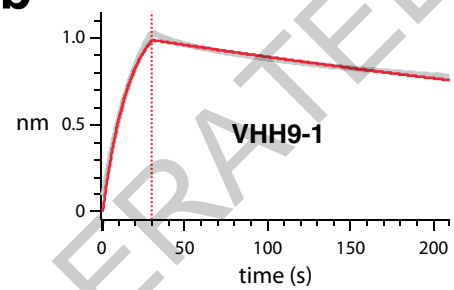
Extended Data Fig. 3 | B cell development in nanomice. a, Flow cytometry analysis of bone marrow (left two columns), spleen (third and fourth columns), and peritoneal cavity B cells (last column) in WT controls and nanomice. First column shows the percentage of B220-gated CD43^{+}IgM^{-} proB and CD43^{-}IgM^{+} immature B cells. Second column shows percentage of Igκ within the B220-gated IgM^{+} population. Third column denotes the total number of B220^{+} B cells in the spleen. Y axis shows viability staining with eFluor506 (eBiosciences). The fourth column shows the percentage of B220-gated CD23^{low}CD21^{low} immature, CD23^{high}CD21^{low} follicular, and CD23^{low}CD21^{high} marginal zone splenic cells in the two strains. The last column shows the percentage of B1 (IgM^{high}B220^{low}) and B2 (IgM^{low}B220^{high}) cells in the peritoneal cavity. Examples of gating for bone marrow and splenic B cells is provided in Extended Data Fig. 11d. **b**, Histograms depicting the percentage of Igκ (left), IgD (middle), and IgG1 (right row) in WT, nanomice, and Igμ-CH1$^{-/-}$ mice. The latter measured in LPS+IL-4+αCD180 *ex-vivo* cultures. Population gates are represented with a line and the percentage of total cells is provided. **c**, Proliferation assay of nanomouse and control B cells cultured for 96h with LPS+IL-4+αCD180. **d**, Immunization regimen. Nanomice were immunized with 50 μg HIV-1 envelop trimer at the indicated dates. **e**, Percent nucleotide substitutions (adjusted for base composition) observed in Nbs isolated from immunized animals. Phage library was selected for binding to HIV-1 envelop trimer.

a



BG505 DS-SOSIP derived nanobodies	VHH usage			D usage			JH usage		
	VHH	length	mutations	D	length	mutations	JH	length	mutations
Nb1	1	295	11	DSP2.10	10	0	JH4	51	4
Nb2	1	295	10	DFL16.1	12	0	JH4	39	0
Nb3	1	295	9	DFL16.1	12	0	JH4	39	0
Nb4	1	296	4	DFL16.2	9	0	JH4	44	0
Nb5	1	296	11	DSP2.3	7	0	JH2	41	0
Nb6	2	294	11	DFL16.1j	6	0	JH4	43	0
Nb7	2	295	3	DQ52-BALB/c	10	0	JH1	46	0
Nb8	6	280	5	DSP2.9	10	0	JH3	39	0
Nb9	6	290	0	DQ52-BALB/c	8	0	JH3	45	0
Nb10 (binder)	9	290	3	DFL16.2	7	0	JH1	51	1
Nb11 (binder)	9	290	2	DFL16.2	7	0	JH1	51	1
Nb12 (binder)	9	290	2	DFL16.2	7	0	JH1	48	2
Nb13 (binder)	9	290	3	DFL16.2	7	0	JH1	51	1
Nb14	16	290	0	DSP2.10	9	0	JH4	44	0
Nb15	16	287	0	DSP2.5	11	0	JH1	49	0
Nb16	19	290	0	DSP2.9	7	0	JH3	39	0

b



nanobody	***kon*** (1/Ms)	***koff*** (1/s)	***KD*** (M)
VHH9-1	2.50e+5	3.31e-3	1.32e-8
VHH9-2	2.88e+5	3.10e-3	1.08e-8
VHH9-3	2.84e+5	3.10e-3	1.09e-8
VHH9-4	2.37e+5	1.81e-3	2.37e-9

Extended Data Fig. 4 | Nanomouse immune response to HIV-1 envelop trimer. a, Upper table shows Nb and JH usage for the 16 VHH clones isolated from immunized nanomice. Middle graph shows protein alignment for VHHs isolated from HIV-1 trimer immunized nanomice. CDRs are boxed. Lower table shows hypermutation profiles for selected Nbs' VHH, D, and J domains. **b**, Left, BLI analysis of BG505 DS-SOSIP binding to immobilized VHH9-1. Red trace represents the raw data, and the kinetic fit is shown in grey underneath. Right, table showing the kinetic constants for association (*kon*), dissociation (*koff*) and equilibrium (*KD*) for all four VHH9 Nb variants.

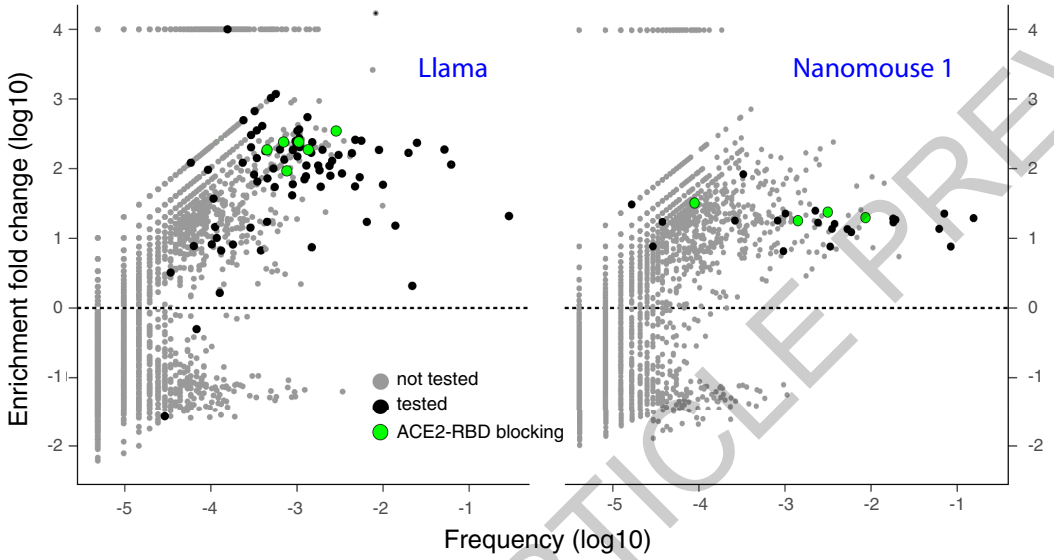
a

	Nbs sequenced	enriched (>= 10 fold)	unique CDR3s	CDR3 clusters
Llama	26,256	719	192	40
Nanomouse 1	23,536	576	35	2
Nanomouse 2	27,493	750	93	7
Nanomouse 3	27,112	629	71	9

b

	Inhibition Rate (%)			
	Exp. 1	Exp. 2	Exp. 3	Average
nanobody	0.1 ug/ml			0.1 ug/ml
Nb15	93.6	91.1	93.8	92.8
Nb17	87.0	79.6	89.3	85.3
Nb19	92.6	82.8	93.1	89.5
Nb56	84.9	85.7	91.8	87.5
Nb12	34.6	33.8	26.9	31.8
Nb30	40.6	36.0	43.7	40.1

c



d

	VHH usage			D usage			JH usage		
SARS-CoV-2 Nbs	VHH usage	length	mutated nt	D usage	length	mutated nt	JH usage	length	mutated nt
Nb30	1	296	11	DSP2.8	6	0	JH4	42	0
Nb12	9	292	11	DFL16.1	17	2	JH4	40	0

e

	kon (1/Ms)		***koff*** (1/s)		***KD*** (M)	
nanobody	monomer	trimer	monomer	trimer	monomer	trimer
Nb15	1.61e+5	3.20e+5	1.11e-3	5.26e-6	8.15e-9	1.64e-11
Nb17	2.13e+5	4.75e+5	1.06e-3	1.35e-7	5.59e-9	<1.0e-12
Nb19	2.60e+5	4.30e+5	1.13e-3	1.51e-5	4.72e-9	3.51e-11
Nb56	2.41e+5	3.29e+5	7.14e-3	1.26e-5	3.26e-9	3.82e-11
Nb12	1.05e+5	2.17e+5	2.40e-3	<1.0e-7	3.00e-8	<1.0e-12
Nb30	2.06e+5	1.17e+5	1.21e-3	9.59e-5	6.55e-9	8.20e-10

Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | Isolation of anti-SARS-CoV-2 RBD Nbs. a, Table indicating i) the total number of unique Nb genes identified from llama and three nanomice phage display libraries following selection for RBD binding, ii) the number of Nbs enriched at least 10-fold post-selection, iii) the number of Nbs with a unique CRD3, and iv) the different clusters of Nbs that share similar CDR3s (with no more than 2 AA differences). **b**, Table showing *in vitro* neutralization results for the 6 leading Nbs using GenScript's sVNT kit. **c**, Dot plot depicting the extent of enrichment (y-axis) and frequency (x-axis) of unique Nbs post RBD selection of llama (left) or nanomouse 1 (right) libraries. Green circles represent Nbs that block ACE2-RBD interactions *in vitro*, black circles are Nbs that do not efficiently block ACE2-RBD interactions, and grey dots represent untested Nbs. **d**, Upper graph shows protein alignment of the 6 Nbs isolated from llama and nanomice immunized with SARS-CoV-2 spike and RBD. Lower table shows detailed information of nanomouse Nbs VHH, D, and J domains. **e**, Table depicting equilibrium (*KD*), association (*Kon*) and dissociation (*Koff*) constants obtained for each Nb as a monomer (black) or trimer (red) form.

a

pseudovirus IC50 (pM)

		R683G	KEN
Nb12	monovalent	6910	9658
	bivalent	525	137
	trivalent	178	90
Nb15	monovalent	75	50000
	bivalent	45	180
	trivalent	13	30
Nb17	monovalent	403	50000
	bivalent	13	50000
	trivalent	4	47350
Nb19	monovalent	122	50000
	bivalent	5	2785
	trivalent	3	115
Nb30	monovalent	3837	5240
	bivalent	464	858
Nb56	monovalent	237	50000
	bivalent	19	384
	trivalent	8	14

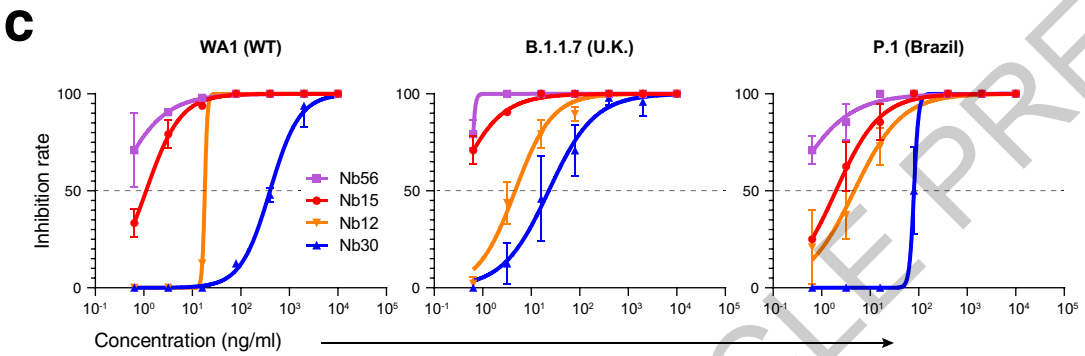
b

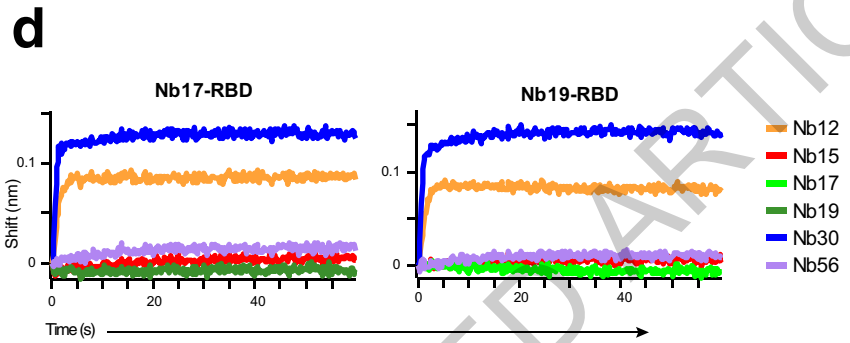
authentic virus IC50 (pM)

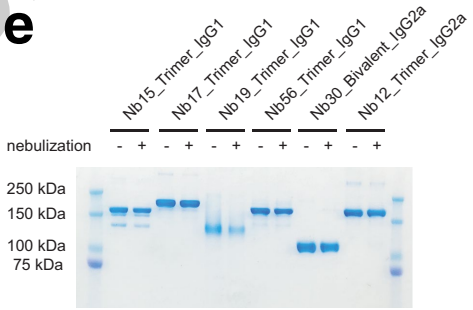
	WT	U.K.	S.A.	BR
	WA1	**B.1.1.7**	**B.1.351**	**P.1**
Nb15	16	4	106	28
Nb56	3	8	18	3
Nb12	249	64	286	65
Nb30	9374	538	2755	1874

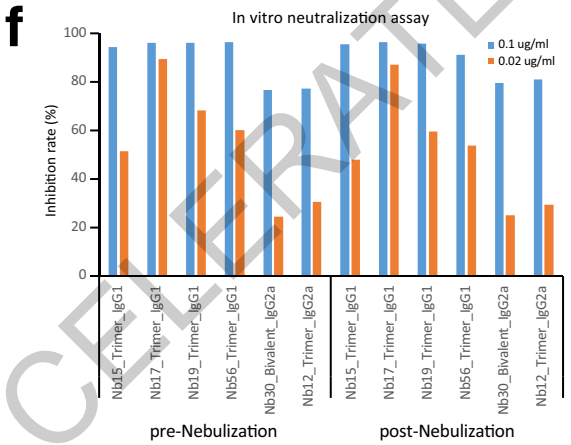
IC50 (ng/ml)

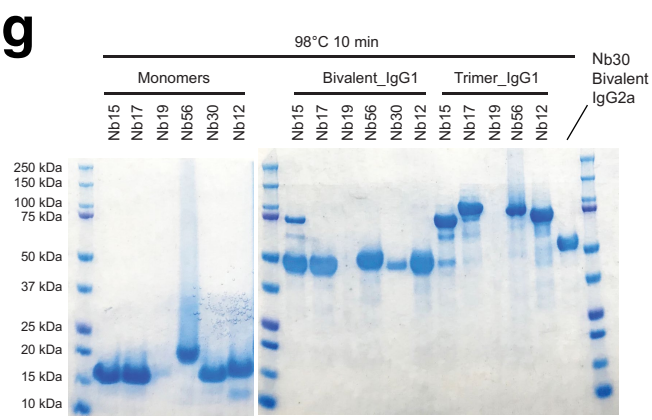
Nb15	1.1	0.3	7.5	2
Nb56	0.2	0.6	1.3	0.2
Nb12	18.1	4.7	20.8	4.7
Nb30	400.1	23	117.6	80





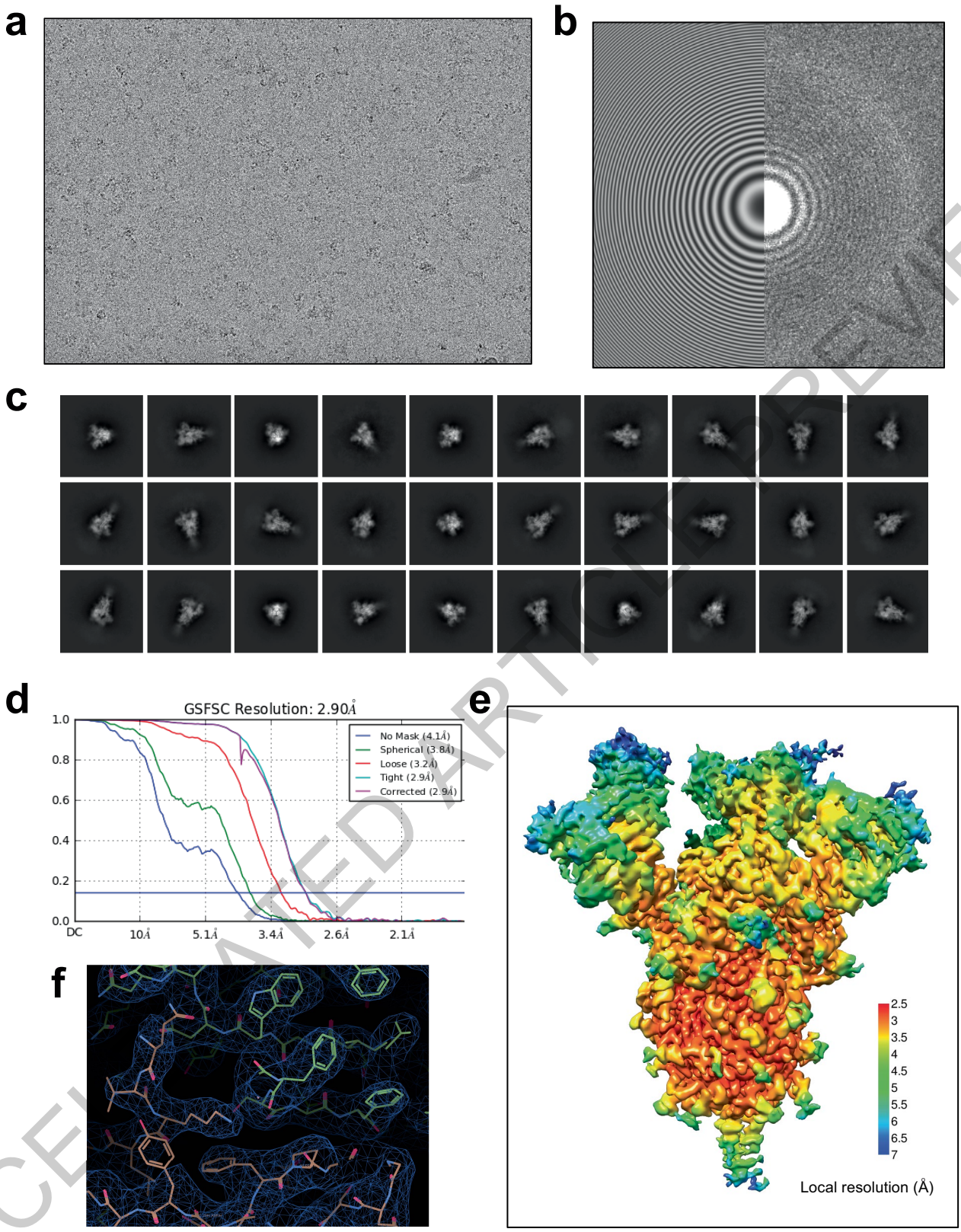




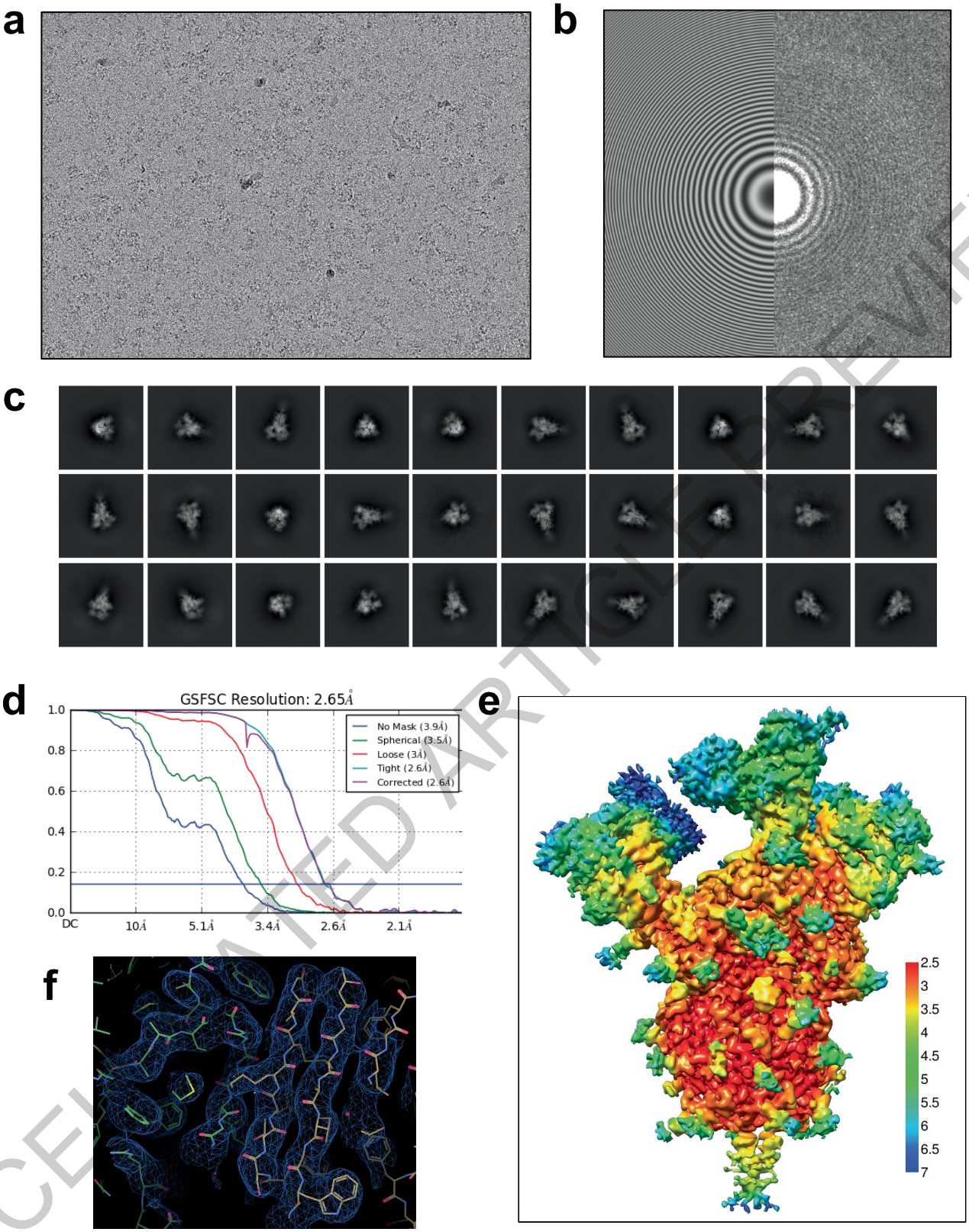


Extended Data Fig. 6 | See next page for caption.

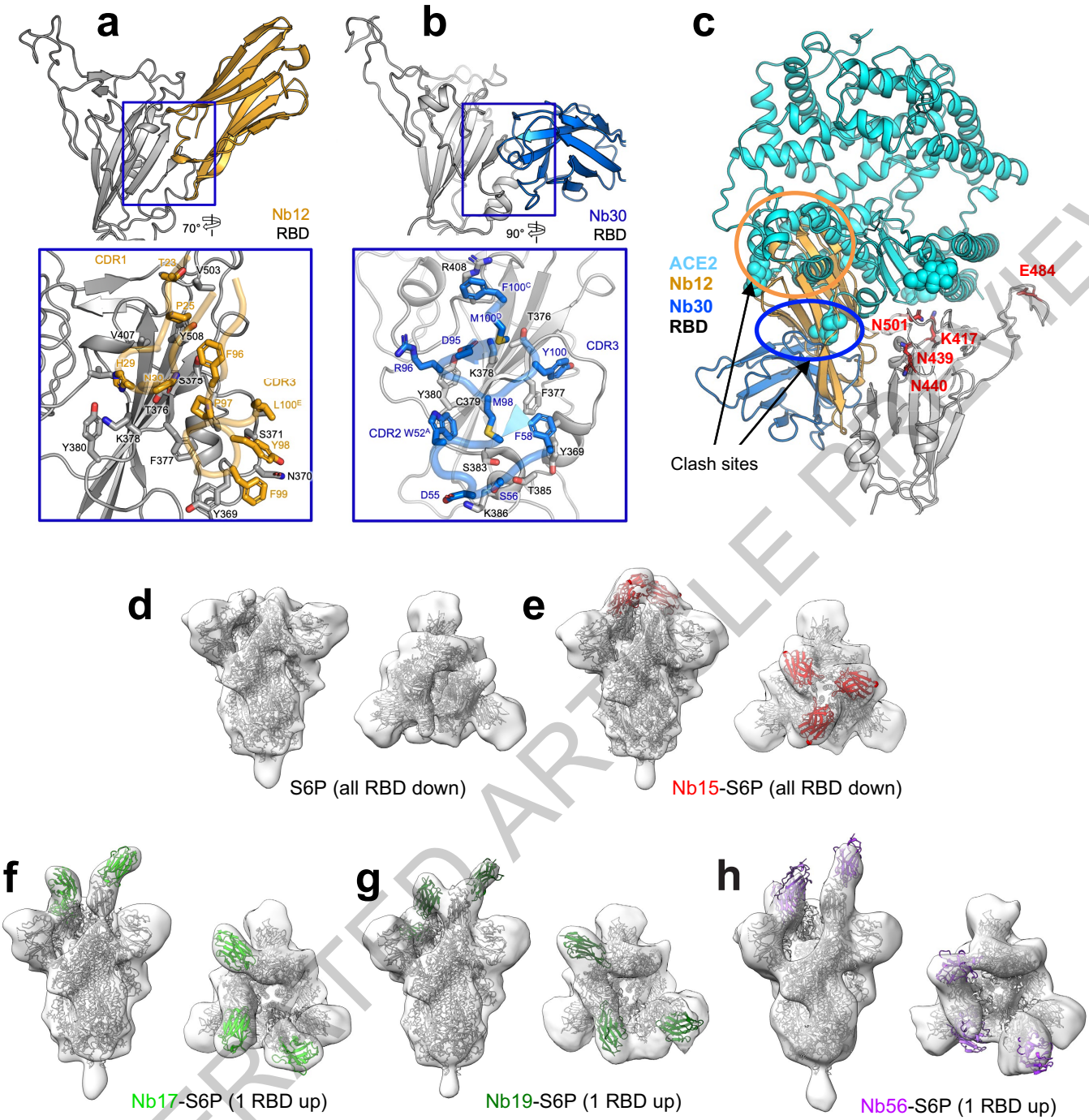
Extended Data Fig. 6 | Neutralization of pseudo- and authentic SARS-CoV-2 viruses. a, Comparison of neutralization activities of leading Nbs in monovalent, bivalent, or trivalent form (results for monovalent and trivalent reproduced from Fig. 3a). **b**, Neutralization assays for WT (WA1) and SARS-CoV-2 variants B.1.1.7 and P.1 for trivalent Nb56, Nb15, Nb12 and Nb30. Data are representative of two independent experiments. Data are the mean ± s.d. of triplicates. **c**, IC50 neutralization values for experiments shown in Fig. 3b and panel b of this figure. Upper table values are in pM, lower table in ng/ml. **d**, Additional panels related to Fig.3c-f. Immunocomplexes used were Nb17-RBD (left) and Nb19-RBD (right). **e**, Coomassie staining showing Nb integrity following nebulization. With the exception of Nb30 (bivalent) all Nbs were fused to Fcs as trimers. Data are representative of two independent experiments. **f**, Bar graph showing *in vitro* neutralization (percentage) of RBD-ACE2 interactions by the different trivalent Nb (bivalent for Nb30) before and after nebulization at two different concentrations (0.1μg/ml (blue) and 0.02 μg/ml). **g**, Coomassie staining showing integrity of Nb monomers (left) or multimers (right) following heat treatment (98 °C for 10 min). Data are representative of two independent experiments.



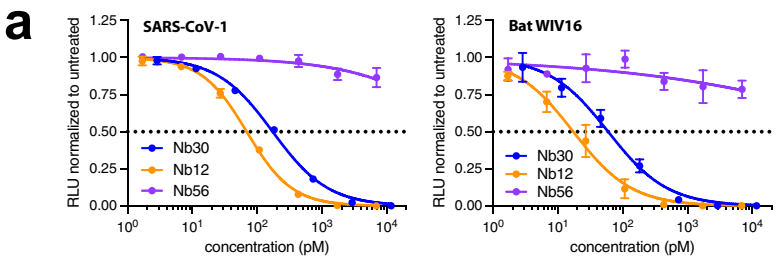
Extended Data Fig. 7 | Cryo-EM data processing and validation for Nb12-spike complex. a, A representative cryo-EM micrograph showing Nb12-spike complex embedded in vitreous ice. **b**, A CTF fit of the micrograph. **c**, Representative 2D average classes. **d**, Overall Resolution estimation (FSC, 0.143). **e**, Local resolution estimation of the cryo-EM map. **f**, Cryo-EM density and models for an interface region between RBD and Nb12 after local refinement.

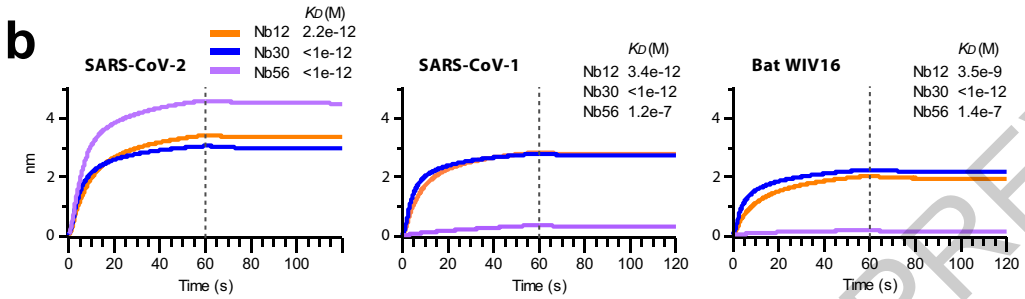


Extended Data Fig. 8 | Cryo-EM data processing and validation for Nb30-spike complex. a, A representative cryo-EM micrograph showing Nb30-spike complex embedded in vitreous ice. **b**, A CTF fit of the micrograph. **c**, Representative 2D average classes. **d**, Overall Resolution estimation (FSC, 0.143). **e**, Local resolution estimation of the cryo-EM map. **f**, Cryo-EM density and models for an interface region between RBD and Nb30 after local refinement.



Extended Data Fig. 9 | Structural analysis of nanomice and llama nanobody interface with the SARS-CoV2 spike. a, Structure of Nb12 and RBD region (inset, Interface between Nb12 and RBD with contact resides shown in stick representation). **b**, Structure of Nb30 and RBD region (inset, Interface between Nb30 and RBD with contact resides shown as stick representation). **c**, Cryo-EM defined structures of nanomouse Nbs recognize regions on RBD distal from emerging mutations 417, 484 and 501. **d**, SARS-CoV2 spike (Hexpro) structure in two perpendicular views. **e**, Spike-Nb15 (red) complex structure in two perpendicular views. **f**, Spike-Nb17 (light green) complex structure in two perpendicular views. **g**, Spike-Nb19 (dark green) complex structure in two perpendicular views. **h**, Spike-Nb56 (purple) complex structure in two perpendicular views.

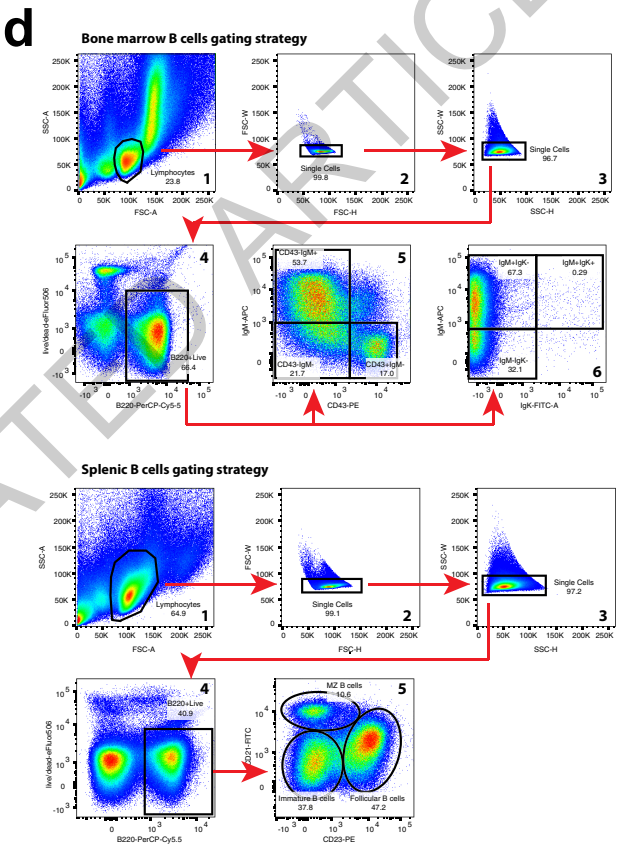




c

IC50 (pM)

	HIV-1 based pseudovirus		VSV based pseudovirus										
	SARS-CoV-1	BatCoV WIV16	SARS-CoV-2	SARS-CoV-2 like bat/pangolin CoVs			SARS-CoV-1	SARS-CoV-1-like bat CoVs					
				RaTG13	GDPangolin	GXPangolin		WIV1	SHC014	LYRa11	Rs7327	Rs4084	Rs4231
Nb12	69	18	146	22	40	229	88	15	17	7	10	19	151
Nb30	174	61	1805	13	27	423	350	7	7	10	19	31	50000
Nb56	50000	50000	2	335	2	50000	50000	50000	50000	50000	50000	50000	3



Extended Data Fig. 10 | Binding and neutralization of sarbecoviruses by nanomouse Nbs and gating strategy for nanomouse and WT B cells. **a**, Neutralization using trivalent Nb56, Nb12, and bivalent Nb30 against pseudoviruses carrying SARS-CoV-1(left) or bat WIV16 (right) spikes. Data are representative of two independent experiments and the error bars represent the mean ± s.d. of triplicates. **b**, BLI analysis of trivalent Nb56, Nb12, and bivalent Nb30 binding to immobilized RBD from SARS-CoV2 (left), SARS-CoV-1 (middle), and Bat WIV16 coronavirus (right). Equilibrium (*KD*) constants are provided. **c**, IC50 (pM) values for neutralization in culture assays, showing the sensitivity of HIV-1 and VSV pseudotyped viruses containing 13 Sarbecoviral spike proteins. **d**, Analysis of bone marrow (upper) or splenic (lower) B cells was done by gating lymphocytes (first plot), avoiding aggregates (plots 2 and 3), $B220^{+}$apoptotic^{-}gaiting (plot 4), and visualization with cell surface makers as indicated (plots 5 and 6).

nature research

Corresponding author(s): Jianliang Xu, Michel Nussenzweig, Peter D Kwong, Rafael Casellas

Last updated by author(s): Apr 29, 2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	BD FACSDiva Software version 9.0 for FACS; Modulus II Microplate Reader User interface version 2.1.0 by TURNER BioSystems; BioTek Gen 5 software for ELISA; Leginon system installed on Titan Krios electron microscopes for CryoEM; MiSeq software v3.1 for Miseq.
Data analysis	FlowJo 10.7.1 for FACS analysis; GraphPad Prism 8.4.2; Microsoft Excel 16.36; Snapgene 4.2.11 for sequence analysis; Fortebio Octet Data Analysis Software 8.0; cryoSPARC 2.15 for EM analysis; NGmerge version 0.2, pTrimmer version 1.3.3, fastp version 2020, BLAST+ version 2.11.0, IgBLAST version 1.17.0, Sickle v1.33, FLASH v1.2.11, ANARCI version 2019 and CD-HIT v4.6.8 for deep sequencing analysis.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All sequencing data are deposited at GEO (GSE167310). All structural data are deposited at EMDB and PDB (EMD-24078, EMD-24077, PDB-7MY3 and PDB-7MY2).

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	To ensure good immune response be obtained, five and six nanomice were immunized with SARS-CoV-2 RBD and Spike recombinant protein, respectively. Best responders were used for nanobody library construction. Ten llamas were screened for serum antibody titer against rabies and clostridium vaccine, and then one llama with highest titer was immunized with RBD and Spike recombinant protein.
Data exclusions	Best responder mice, one from RBD immunized group, two from Spike immunized group, were picked for nanobody isolation. The rest of mice were not used for further analysis. Exclusion criteria were not pre-established.
Replication	All experiments successfully repeated at least twice.
Randomization	This is not relevant as this is an observational study and there is no selection or accidental bias introduced in the study.
Blinding	Real SARS-CoV-2 virus neutralization assay on WA1, B.1.1.7, B.1.351 and P.1 was performed double blinded. For all other experiments, investigators were not blinded during data collection and analysis, as blinding was not relevant in those observational studies where no bias was introduced.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	anti-B220-PerCP-Cy5.5 (eBioscience, 45-045-82), anti-B220-APC (Invitrogen, 17-0452-83), anti-IgM-APC (eBioscience, 17-5790-82), anti-Igk-PE (BD Pharmingen, 559940), anti-Igk-FITC(BD Pharmingen, 550003), anti-Igl-FITC (BD Pharmingen, 553434), anti-IgG1-PE (BD Pharmingen, 550083), anti-IgG1-APC (BD Pharmingen, 550874), anti-IgD-FITC (BD Pharmingen, 553439), anti-CD95-PE (BD Pharmingen, 554258), anti-CD43-PE (BD Pharmingen, 553271), anti-CD23-PE (BD Pharmingen, 553139), anti-CD21-FITC (Biolegend, 123408), Viability Dye eFluor506 (Invitrogen, 1923275), anti-VHH (Jackson ImmunoResearch, 128-035-232), anti-CD180 (BD Pharmingen, 552128)
Validation	All antibodies are commercially available with at least one reference citation.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	E14 ES cells (ATCC CRL-1821) 293T (ATCC CRL-11268) 293TAce2 (derived from 293T); was generated in the Laboratory of Retrovirology, Rockefeller University (Dr. Paul D. Bieniasz) VeroE6 (ATCC CRL-1586) Expi293 (Thermo Fisher Scientific, A14528) WK6 cells (ATCC, 47078)

Authentication	Not authenticated after purchase.
Mycoplasma contamination	All cell lines tested negative for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Five (2 male and 3 female, 10 weeks old) and six (4 male and 2 female, 8 weeks old) nanomice were immunized with SARS-CoV-2 RBD and Spike recombinant protein, respectively. One llama (male, 2 years old) was immunized with RBD and Spike recombinant protein.
Wild animals	The study did not involve wild animals.
Field-collected samples	The study did not involve samples collected from the field.
Ethics oversight	Animal study was approved by NIAMS ACUC at ht the NIH.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Cells isolated from bone marrow, spleen or peritoneal cavity of nanomice were washed with PBS before staining. Cells stimulated in culture medium for 72-96 hours were also washed with PBS before staining.
Instrument	FACSCanto (Becton Dickinson)
Software	BD FACSDiva Software and FlowJo
Cell population abundance	Cells were stained and analyzed on FACSCanto directly, no sorting applied.
Gating strategy	Cells were first gated for lymphocytes in FSC-A (x-axis) versus SSC-A (y-axis). We identify single cells in FSC-W versus FSC-H, and then SSC-W versus SSC-H. We then select B220+ Live B Cells (Viability Dye eFluor506 negative, Invitrogen) for further analysis.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.