Check for updates

SARS-CoV-2 variant prediction and antiviral drug design are enabled by RBD in vitro evolution

Jiří Zahradník [1], Shir Marciano[1], Maya Shemesh[1], Eyal Zoler[1], Daniel Harari[1], Jeanne Chiaravalli[2], Björn Meyer [3], Yinon Rudich [4], Chunlin Li[4], Ira Marton[1,4], Orly Dym[5], Nadav Elad[6], Mark G. Lewis [7], Hanne Andersen [7], Matthew Gagne[8], Robert A. Seder [8], Daniel C. Douek[8] and Gideon Schreiber [1]✉

SARS-CoV-2 variants of interest and concern will continue to emerge for the duration of the COVID-19 pandemic. To map mutations in the receptor-binding domain (RBD) of the spike protein that affect binding to angiotensin-converting enzyme 2 (ACE2), the receptor for SARS-CoV-2, we applied in vitro evolution to affinity-mature the RBD. Multiple rounds of random mutagenic libraries of the RBD were sorted against decreasing concentrations of ACE2, resulting in the selection of higher affinity RBD binders. We found that mutations present in more transmissible viruses (S477N, E484K and N501Y) were preferentially selected in our high-throughput screen. Evolved RBD mutants include prominently the amino acid substitutions found in the RBDs of B.1.620, B.1.1.7 (Alpha), B1.351 (Beta) and P.1 (Gamma) variants. Moreover, the incidence of RBD mutations in the population as presented in the GISAID database (April 2021) is positively correlated with increased binding affinity to ACE2. Further in vitro evolution increased binding by 1,000-fold and identified mutations that may be more infectious if they evolve in the circulating viral population, for example, Q498R is epistatic to N501Y. We show that our high-affinity variant RBD-62 can be used as a drug to inhibit infection with SARS-CoV-2 and variants Alpha, Beta and Gamma in vitro. In a model of SARS-CoV-2 challenge in hamster, RBD-62 significantly reduced clinical disease when administered before or after infection. A 2.9 Å cryo-electron microscopy structure of the high-affinity complex of RBD-62 and ACE2, including all rapidly spreading mutations, provides a structural basis for future drug and vaccine development and for in silico evaluation of known antibodies.

Humans can be infected by SARS-CoV-2 either through inhalation of airborne viral particles or by touching contaminated surfaces. Structural and functional studies have shown that a single RBD of the SARS-CoV-2 homotrimer spike glycoprotein interacts with ACE2, which serves as its receptor[1,2]. Binding of spike (S) protein to ACE2 and subsequent cleavage by the host protease transmembrane serine protease 2 (TMPRSS2) results in cell and virus membrane fusion and cell entry[1]. Blocking of the ACE2 receptor by specific antibodies prevents viral entry[1,3–5]. In vitro binding measurements have shown that SARS-CoV-2 S protein binds ACE2 with an affinity of around 10 nM, which is about tenfold tighter than the binding of the SARS-CoV S protein[2,4,6]. It has been suggested that this is, at least partially, responsible for the higher infectivity of SARS-CoV-2[7]. Recently, three major SARS-CoV2 variants of concern have emerged and mutations in the RBD of the spike proteins of these variants have further strengthened this hypothesis. Deep-mutational scanning of the RBD domain showed that the N501Y mutation in the Alpha variant to enhances binding to ACE2[7]. The Beta variant has three altered residues in the ACE2-binding site (K417N, E484K and N501Y), and has spread extremely rapidly, becoming the dominant lineage in the Eastern Cape and Western Cape Provinces within weeks[8]. The Gamma variant, with independent K417T, E484K and N501Y mutations, similar to the B.1.351 variant is spreading rapidly from the Amazon region[9]. Another S mutation associated with

increased SARS-CoV-2 infectivity is S477N, which became dominant in many regions[10].

Efficacious vaccines are now being administered[11]. However, especially when a large fraction of the global population remains unvaccinated, the potential of the continuously mutating virus to become at least partially resistant to vaccines means that drug development must continue. Potential therapeutic targets that block viral entry include molecules that block the spike protein, the TMPRSS2 protease or the ACE2 receptor[12]. Multiple high-affinity neutralizing antibodies have been developed[13]. Soluble forms of the ACE2 protein[14,15] or engineered parts or mimics have also shown efficacy[16,17]. In addition, previously developed TMPRSS2 inhibitors have been repurposed for treatment of COVID-19[1].

The development of molecules to block the ACE2 protein has not received much attention. One potential caveat with this approach is the importance of ACE2 biological activity, both as a carboxypeptidase removing a single C-terminal amino acid from angiotensin II to generate angiotensin-(1–7) and in the regulation of amino acid transport and pancreatic insulin secretion[18,19]. Dalbavancin is a drug that blocks the spike protein–ACE2 interaction, however it does so with low affinity[20] (approximately 130 nM).

We hypothesized that the RBD domain of SARS-CoV-2 could be used as a competitive inhibitor of the ACE2 receptor binding site. However, this would probably require an RBD with picomolar affinity for ACE2. We recently developed a yeast display strategy,

[1]Department of Biomolecular Sciences, Weizmann Institute of Science, Rehovot, Israel. [2]Chemogenomic and Biological Screening Core Facility, Institut Pasteur, Paris, France. [3]Viral Populations and Pathogenesis Unit CNRS UMR 3569, Institut Pasteur, Paris, France. [4]Department of Earth and Planetary Sciences, Weizmann Institute of Science, Rehovot, Israel. [5]Department of Life Sciences Core Facilities, Weizmann Institute of Science, Rehovot, Israel. [6]Department of Chemical Research Support, Weizmann Institute of Science, Rehovot, Israel. [7]Bioqual, Rockville, MD, USA. [8]Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA. ✉e-mail: gideon.schreiber@weizmann.ac.il
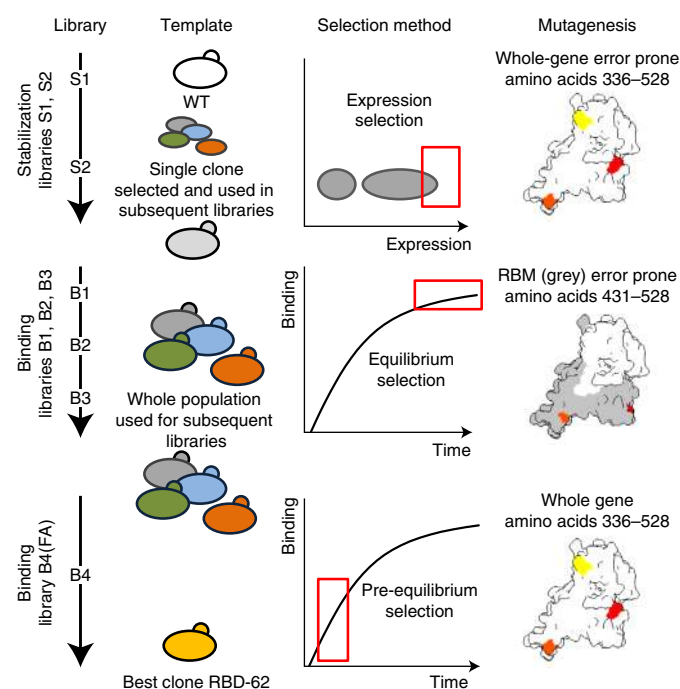
Fig. 1 | SARS-CoV2 RBD in vitro selection. Six consecutive libraries were created by error-prone PCR (inserting 1–5 mutations per gene) to increase RBD domain stability and binding to the ACE2 receptor. Libraries S1 and S2 were selected for higher expression at 30 °C and 37 °C, respectively. A single clone, I358F, was identified (Extended Data Fig. 1). Based on this stabilized RBD, we created libraries B3–B5, randomizing residues 431–528. From each selected library, the pool of enriched clones was used as a template for the subsequent library. The last library B6(FA) was created by pre-equilibrium selection of randomly mutated residues 336–528. Multiple clones from each library were sequenced and monitored for affinity and stability (Tables 1 and 2).

based on carboxy- and amino-terminal fusions of extremely bright fluorescent moieties, which can report expression at extremely low levels, enabling selection down to picomolar concentrations[21].

SARS-CoV-2 has sampled an enormous sequence space, owing to more than 180 million infections globally. Among the huge number of mutations, a very limited number have emerged multiple times and become fixed in the population, indicating that these mutations have a strong adaptive advantage. Here we mutated and affinity-matured SARS-CoV-2 S RBD to characterize the adaptive advantage of RBD mutations and inform epidemiology and risk assessments of variants, and to identify RBD variants that might function as effective drugs to treat COVID-19.

Results

RBD domain mutations recapitulate steps in SARS-COV-2 evolution. For in vitro evolution, we took advantage of an enhanced yeast surface display protocol we developed recently using two different detection strategies, eUnaG2 and DnbALFA[21]. Before library construction, we tested various RBD lengths for optimal surface expression and stability (Supplementary Table 1). Subsequently, we chose RBDcon2 for yeast display and RBDcon3 for protein expression (see Methods). Libraries were constructed in a stepwise manner. Variants from libraries S1 and S2 were screened for stability-enhancing mutations, followed by ACE2 binding (B3–B5), and finally for fast association (B6(FA)) (Fig. 1). All libraries were constructed by error-prone PCR random mutagenesis of the RBD, introducing 1–5 mutations per clone (Fig. 1). Libraries S1 and S2 converged towards the I358F mutation, with the phenylalanine resi-

due fitting into the hydrophobic pocket formed in the RBD domain (Fig. 1 and Extended Data Fig. 1). This mutation nearly doubled the fluorescence signal intensity and was used to construct library B3 (Fig. 1), in which mutations were restricted to the receptor binding motif (RBM) (residues G431–K528) (Fig. 1). The B3 library was expressed at 37 °C to maintain pressure on protein stability, and yeast cells were selected by fluorescence-activated cell sorting (FACS) against decreasing concentrations of ACE2 labelled with CF640R succinimidyl ester (1,000, 800 and 600 pM and 4 h of incubation). Library enrichment was achieved by selecting the top 3% of binding cells (Fig. 1 and Extended Data Fig. 2; Methods). Plasmid DNA was isolated from selected yeast cells in the sorted library and transformed into *Escherichia coli* for preparation of a new library (B4). This approach, which mimics natural virus evolution, enriches the subsequent library with multiple selected mutations and purges non-beneficial mutations, enabling the screening of a wider sequence space and epistatic mutations, with multiple trajectories being sampled. Thirty single-colony isolates were used for sequencing to monitor the enrichment process and subsequently for screening of binding affinity. Analysis of the selected B3 library yielded two dominant mutations appearing in more than 70 % of clones: E484K and N501Y. In addition, multiple minor mutations: V483E, N481Y, I468T, S477N, N448S and F490S were found (Table 1, Fig. 2a,b and Extended Data Fig. 3). The analysis of library B4 (selected against 600, 400 and 200 pM ACE2) showed absolute domination of E484K and N501Y mutations. Besides the dominant clones, mutations N460K, Q498R and S477N increased to frequencies above 20%, and new minor populated mutations were identified: G446R, I468V, T478S, F490Y and S494P (Table 1, Fig. 2 and Extended Data Fig. 3).

Among the mutations selected and fixed in the yeast population during these initial steps of affinity maturation were three mutations that came to prominence[8–10] in clinical samples of SARS-CoV-2: S477N, E484K and N501Y. In addition, the yeast selection probed the most abundant naturally occurring variants in positions 468, 490 and 494, which were not fixed in yeast selection (Table 1). Indeed, analysis of the I468T and S494P on their own show them to be nearly neutral with respect to ACE2 binding or yeast surface expression (Fig. 2b,d). The stringency of our selection process is likely to be responsible for the fixation of variants with the largest contribution to the binding affinity, whereas other factors are also important for viral evolution.

RBD affinity mutants compared with variants identified in GISAID. To evaluate why some mutations were prevalent in SARS-CoV-2 patient samples and also selected by yeast display, while others were not, we plotted the occurrence of all single-amino-acid mutations in the GISAID database[22] with respect to the apparent change in the RBD–ACE2-binding affinity ($K_{D,App}$) as estimated by the frequency of given amino acids within a mutant library at a given concentration (that is, using the deep-mutational scanning approach[7]). Figure 2e (black dots) shows that the more prevalent mutations in GISAID have a higher binding affinity in our assays. To quantify these results, we measured the binding of recloned isogenic variants of the most prevalent mutations from library B3 and the mutations that appear most frequently in SARS-CoV-2. The $K_{D,YD}$ values calculated here for the single mutations from full titration curves (Fig. 2 and Extended Data Fig. 4) are shown as blue squares in Fig. 2e. Overall, $K_{D,App}$ and $K_{D,YD}$ values for the different mutants change in a coordinated way—however, the fold-change relative to wild type (WT) is higher for $K_{D,YD}$ (Extended Data Fig. 5a), as indicated by a larger slope for the regression of the blue squares compared with the black dots.

The tightest binding clone of library B3 is the E484K/N501Y double mutant (RBD-33) (Fig. 2a), which has an affinity of 204 pM and includes the core mutations in the B.1.351, P.1 and P.3 variants.

Table 1 | RBD domain mutations selected by in vitro evolution for higher affinity versus those spreading in SARS-CoV-2

In vitro evolution																	
RBD position*	358	445	446	448	460	468	470	477	478	481	483	484	490	493	494	498	501
WT residue	I	V	G	N	N	I	T	S	T	N	V	E	F	Q	S	Q	N
Library S2	**F**																
Library B3	**F**		S		T			N		Y	E	K	S				Y
Library B4	**F**		R	K	V			N	S			**K**	Y		P	R	Y
Library B5	**F**		R	K	V			N				**K**	Y	H	P	**R**	Y
Library B6(FA)	**F**	**K**			**K**	**T**	**M**	**N**				**K**			P	**R**	Y
Clone B62	**F**	**K**			**K**	**T**	**M**	**N**				**K**				**R**	Y
SARS-CoV2 variant	Lineage[a]																
	B.1.160							**N**									
	B.1.1.317							**N**									
Alpha	B.1.1.7																Y
Beta	B.1.351											**K**					Y
Gamma	P.1											**K**					Y
Theta	P.3											**K**					Y
All mutations [b]					T, (I), **K**, S, C	F, **T**, L, J, M, (**V**)	I, V, K, (N), R, A	I, R, G, T, K, (**N**)	**S**, A, I, R, (K), H	**Y**, S, (K), D, T, H	(F), A, G, D, T, L, I	Q, A, D, G, R, V, L, (**K**)	**Y**, I, R, L, V, (**S**)	L, (K), R, I, E, P	(**P**), H, R, I, L, A, R, T	H, P, F, (**R**), M, P, L	R, T, S, K, E, H, (**Y**)

Amino acids in bold are dominant, that is, they account for more than 50% of selected clones; underlined amino acids are minor, that is, they account for fewer than 50% of clones at a given position. Emerging mutations that are present both in clinical samples and in yeast display experiments occur at positions 477, 484 and 501. [a]Lineage designation by Rambaut et al.[48]; alternative strain designation proposed by NextStrain initiative[49]: 20E/EU1 (B.1.160), Alpha (B.1.1.7, 20I/501Y.V1), Beta (B.1.351, 20I/501Y.V2) and Gamma (P.1, 20I/501Y.V3). [b]The most abundant mutation at a given position is shown in parentheses. Mutations that were detected by yeast display affinity maturation are in bold. Insertion–deletion mutations were not included (GISAID database (13 April 2021)). Biophysical and binding parameters of multiple selected mutants and RBDs from different virus lineages are shown in Table 2 and Fig. 2.

For comparison, the single N501Y mutant (B.1.1.7) has an affinity of 500 pM and the S477N mutant has an affinity of 963 pM (compared to 1.7 nM for the WT) (Fig. 2c and Extended Data Fig. 3). In addition to E484K/N501Y, the B.1.351 and P.1 variants include a mutation in residue K417 to N and T, respectively. As residue 417 is not in the RBM, it was not included in the region mutated in libraries B3–B5. Therefore, we introduced the K417N and K417T mutations (K417N/T) by site-directed mutagenesis in addition to E484K/N501Y and measured its contribution to binding and expression. On their own, both K417N and K417T decrease binding but increase expression on the yeast surface (Fig. 2c,f). However, on the background of E484K/N501Y, K417N/T does not affect binding, but increases the yeast surface expression of RBD E484K/N501Y to nearly that of the WT RBD (RBD-WT) (Fig. 2f). Calculating double-mutant cycles between E484K/N501Y and K417N/T indeed shows positive cooperativity between these residues (Fig. 2g). These data explain why K417N/T mutations convergently emerged in B.1.351 and P.1 variants.

The most abundant naturally occurring mutations in the RBD[8–10,23] were selected by yeast display, and are present in library B3. We next explored whether higher affinity binding was achievable by applying additional selections.

Characterization of ACE2–RBD affinity limits. Based on the observation that tighter binding clones were selected in our assay, the selection of such binders might predict future SARS-CoV-2 variants. Furthermore, a tight binder could be used to block ACE2 and thereby inhibit SARS-CoV-2 infection. We therefore created library B5 (Fig. 1) and selected it using 200, 50 and 30 pM ACE2. Sorting with less than 100 pM bait was done after overnight incubation to reach equilibrium in 50 ml solution in order to prevent the ligand-depletion effect (as the number of ACE2 molecules becomes much lower than the number of RBD molecules). B5 resulted in the fixation of mutations E484K, Q498R and N501Y in all sequenced clones (Extended Data Fig. 6). Mutations N460K, S477N and S494P were present with frequencies higher than 20%. Additional mutations identified were G446R, I468V and F490Y. Representative clones with different mutational profiles were subjected for detailed analyses (Table 2 and Extended Data Figs. 3 and 5).

The final selection aimed to achieve faster association rates by using pre-equilibrium selection[24]. Library B6(FA) (Fig. 1) was created by randomizing the full-length RBD gene, using the enriched B5 library as template. B6 was pre-selected with 30 pM ACE2 for 8 h (reaching equilibrium after overnight incubation), followed by 90 min incubation before selection. This procedure resulted in the accumulation of additional mutations: V445K, I468T and T470M, and the fixation of the previously observed S477N mutation in all sequenced clones. Five minor mutations, N354E, K417T, V367W, S494P and S514T, with only a single sequence each, were also identified. Among them K417T was also present in the P.1 variant, indicating that it may confer an advantage for the virus, as well as in yeast display selection[9]. It should be noted that V445K/T470M requires two nucleotide changes, demonstrating the utility of using multiple rounds of library creation on top of previous libraries (and not single clones). Interestingly, these mutations were not located at the binding interface, but rather at the periphery, in line with previously described computational fast-association design, in which peripheral mutations had key roles[25]. From the B6(FA) library, we determined the isogenic binding for 4 different clones, with RBD-62 showing the highest affinity of 3.5 ± 1.8 pM (Fig. 2a and Table 2). To validate the $K_{D,YD}$ values, we purified RBD clones WT, 36, 48, 52, 521, 62, 614 and 71 expressed in Expi293F cells (Supplementary Fig. 1). Measuring the binding affinity to ACE2 using the Biacore S200 (GE Healthcare) resulted in an affinity of 17 nM for RBD-WT and 16 pM for RBD-62 binding to ACE2, a 1,000-fold increase in affinity over the WT (Fig. 2h and Table 2). While most of the
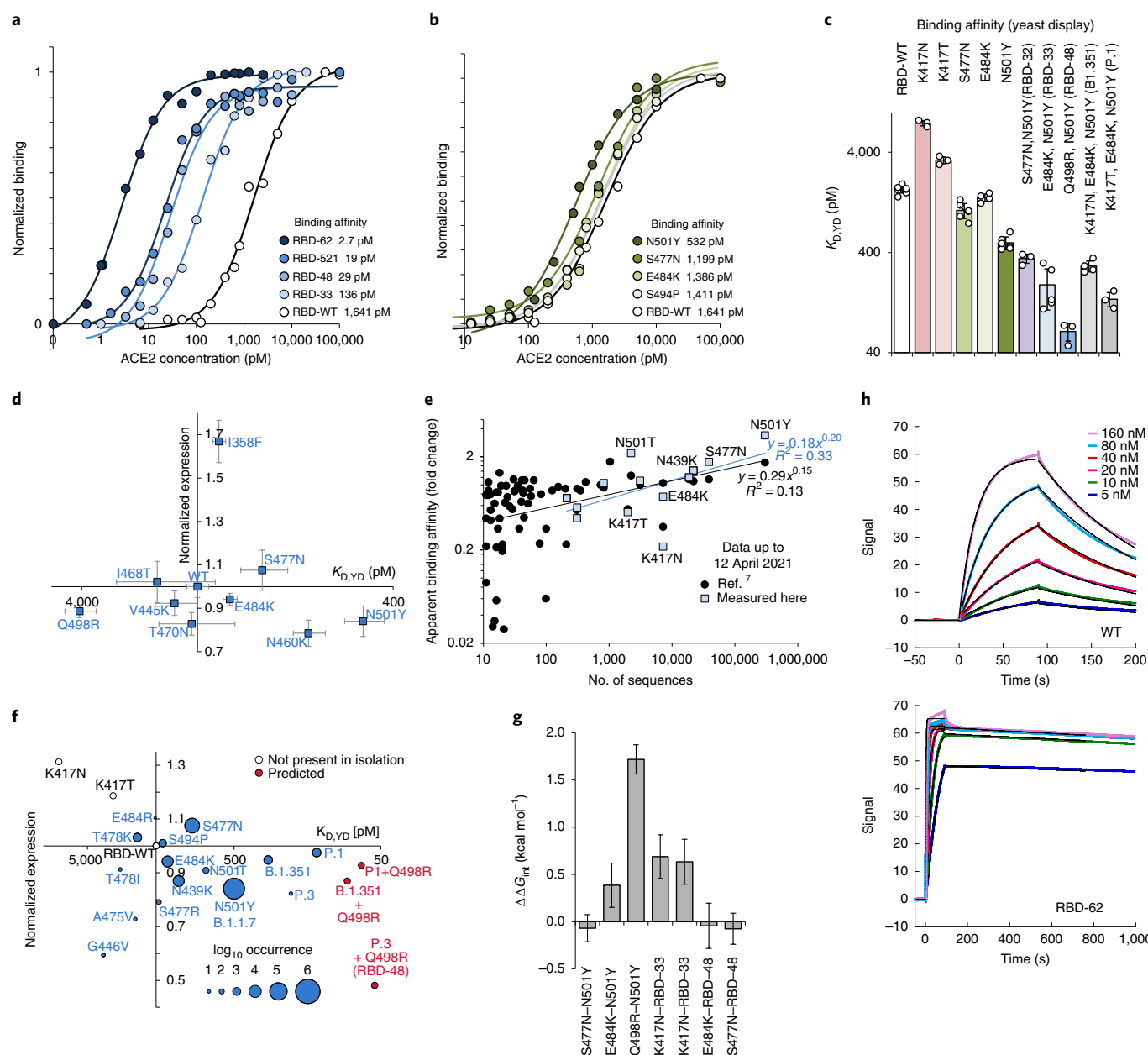
Fig. 2 | In vitro evolution of spike protein RBD. a, Binding titration curves for the highest affinity binding variant in each successive yeast library. The mutations in each of the variants are shown in Table 2. **b**, Binding titration curves for selected single mutations that occur frequently in both yeast display selection and in evolving SARS-CoV-2 variants. **c**, Comparison of binding affinities for selected mutants and RBD variants. Data are shown as individual data points and mean values ± s.e.m. **d**, The effect of the isogenic mutations identified in RBD-62 on binding to ACE2 and expression level on the yeast surface ($n \geq 3$; data are mean ± s.e.m.). **e**, Apparent binding affinities for ACE2 of single-amino-acid mutations as calculated from deep-mutational scanning of the RBD domain[7] ($K_{D,App}$, black dots) or from yeast titration done in this study ($K_{D,YD}$, blue squares) plotted against their frequency in the GISAID database[22]. Mutations with fewer than ten sequences were not included in the analysis. **f**, The effect of the most prevalent mutations in and variants of SARS-CoV-2 RBD on ACE2 binding and yeast surface expression combined with the frequency of variants. The K417N/T mutations occur in the B.1.351 and P.1 variants in combination with N501Y/E484K. The K417N/T occur independently only infrequently. Red circles show the affinity and expression of the single Q498R mutation on top of the B.1.351, P.1 and P.3 variants. **g**, Double-mutant cycle calculations showing interaction energies ($\Delta\Delta G_{int}$) between selected residues, suggesting independent or cooperative behaviour. Standard errors were calculated from 3–6 repeats of individual measurements (details in Methods). **h**, Surface plasmon resonance binding sensorgrams for RBD-WT (top) and RBD-62 (bottom) at six concentrations of analyte (5–160 nM). The black line is the global fit using the instrument's built-in function.

improvement came from reduced k_{off}, k_{on} increased tenfold, from 4.5×10^5 to 44×10^5 M^{-1}s^{-1} for RBD-62 (Table 2). The other variants had affinities in between these two extremes (Extended Data Fig. 5 and Table 2). The increase in affinity as determined using the Biacore S200 correlates with those measured by yeast titration

($R^2 = 0,89$; Extended Data Fig. 5c); however, absolute affinity values differed by tenfold on average (Table 2). Differences in absolute binding affinity (K_D) values are common between different techniques, whereas relative values are much more stable[26]. Measuring its thermostability showed that RBD-62 was more stable than the

Table 2 | Binding affinities and thermostability of the mutant clones selected by yeast display

Clone	Library	Plasmid[a]	Mutations	T_m (°C)	$K_{D,YD}$ (pM)[c]	K_D (M)[d]	k_{on} (M^{-1}s^{-1})	k_{off} (s^{-1})	Chi² (RU²)
RBD-WT		pJYDC1	Residues 336–528	53.9	1,700 ± 137 [8]	1.7×10^{-8}	4.5×10^5	7.6×10^{-3}	0.222
RBD-32	B3	pJYDC1	I358F, S477N, N501Y	ND	348 ± 35 [3]	ND			
RBD-33	B3	pJYDC1	I358F, E484K, N501Y	ND	204 ± 88 [4]	ND			
RBD-36	B3	pJYDC1	I358F, I468T, N481Y, N501Y	54.6	268 ± 35 [3]	1.0×10^{-8}	2.1×10^5	2.2×10^{-3}	0.119
RBD-48	B4	pJYDC3	I358F, E484K, Q498R, N501Y	54.8	55 ± 23 [4]	1.7×10^{-10}	1.0×10^6	1.7×10^{-4}	0.182
RBD-52	B5	pJYDC1	I358F, N460K, E484K, S494P, Q498R, N501Y, A520G	61.9	40 ± 16 [5]	2.4×10^{-10}	3.8×10^5	9.0×10^{-5}	0.0515
RBD-521	B5	pJYDC1	I358F, N460K, E484K, Q498R, N501Y	58.3	14 ± 3 [3]	1.7×10^{-10}	9.0×10^5	1.6×10^{-4}	0.0562
RBD-62	B6(FA)	pJYDC3	I358F, V445K, N460K, I468T, T470M, S477N, E484K, Q498R, N501Y	57.6	3.5 ± 1.8 [4]	1.6×10^{-11}	4.4×10^6	6.8×10^{-5}	0.42
RBD-614	B6(FA)	pJYDC3	I358F, V445K, N460K, E484K, S494P, Q498R, N501Y	58.5	21 ± 6.2 [3]	1.9×10^{-10}	6.8×10^5	1.3×10^{-4}	0.0488
RBD-71		pJYDC3	I358F, V367W, R408D, K417V, V445K, N460K, I468T, T470M, S477N, E484K, Q498R, N501Y	63.5	8.5 ± 2.5 [3]	1.6×10^{-10}	1.2×10^6	2.0×10^{-4}	0.104

Underlined clones have the highest affinity among the clones of a given library, and their representative binding curves are shown in Fig. 2a. ND, not determined; RU, response units. [a]pJYDC1 (plasmid 162458) uses the intrinsic eUnaG2 reporter; pJYDC3 (plasmid 162460) contains the DnbALFA reporter[21]. [b]Melting temperature (T_m) as measured by differential scanning fluorimeter Tycho NT.6 (NanoTemper Technologies). [c]K_D values measured between yeast surface-exposed RBD variants and the monomeric extracellular portion of the ACE2 receptor (residues G17–Y613). The number of biological replicates is shown in square brackets. [d]Measured by the Biacore S200 (GE Healthcare) using a series S sensor chip protein A (GE Healthcare) activated with Fc–ACE2[27] (Methods). The sensorgrams are shown in Fig. 2h (WT and RBD-62) and Extended Data Fig. 5. Values were calculated from global fitting of the data. k_{on} and k_{off}, association and dissociation rate constants, respectively; Chi², Chi squared test value.

WT by 4 °C, probably due to the introduction of the I358F, I468T and S477N stabilizing mutations (Extended Data Fig. 7).

To further increase RBD-62 affinity and stability, we prepared a site-directed mutagenesis library using RBD-62 as template, and included the 15 beneficial mutations suggested by deep-mutational scanning[7], most of which require more than one nucleotide change to be reached. None of these mutations alone increased the affinity towards RBD-62. However, a combination of three of mutations (V367W, R408D and K417V) stabilized RBD-62 by 5 °C, creating RBD-71, but at the cost of decreased binding affinity (Table 2 and Extended Data Fig. 7). This demonstrates the limitations of using single-amino-acid changes from deep-mutational scanning to obtain high-affinity RBD binders.

Cryo-electron microscopy structure of RBD-62 complexed with ACE2. We determined the cryo-electron microscopy (cryo-EM) structure of the N-terminal peptidase domain of ACE2 (G17–Y613) bound to RBD-62 (T333–K528) (Fig. 3a,b) (Electron Microscopy Data Bank accession EMD-12187, Protein Data Bank (PDB) ID: 7BH9). Details of sample preparation, data acquisition and structure determination are described in Methods. The cryo-EM data collection and refinement statistics are summarized in Supplementary Table 2 and Supplementary Figs. 2–4. Structure comparison of the ACE2–RBD-62 complex and the ACE2–WT complex (PDB ID: 6M0J) revealed their overall similarity, with a root mean squared deviation of 0.97 Å across 586 amino acids of ACE2 and 0.66 Å among 143 amino acids of the RBD (Extended Data Fig. 8). Three segments; R357–S371 (β2 and α2), G381–V395 (α3) and F515–H534 (β11) are disordered in RBD-62, and thus not visible in the density map (blue cartoon in Extended Data Fig. 8). These segments are situated opposite to the ACE2-binding interface

and are therefore not stabilized and rigidified by ACE2 contacts. All mutations (except I358F) are present in the density map. Mutations V445K, N460K, I468T, T470M, S477N, E484K, Q498R and N501Y are part of the receptor-binding motif (RBM) that interacts directly with ACE2 (Fig. 3a,b, orange spheres). The RBM, including residues S438–Q506, shows the most pronounced conformational differences compared with RBD-WT (Fig. 3c and Supplementary Fig. 5, black circle). Out of the nine mutations introduced in the RBM, four mutations create new intramolecular interactions, thus stabilizing the RBD-62 structure, including hydrogen-bond contacts between K460 and D420, and between T468 and R466. Interestingly, these are contacts that are gained in library B6(FA).

The mutations S477N, Q498R, N501Y form new contacts with ACE2. Y501 makes a π-interaction with Y41 and positions R498 to make a hydrogen bond and salt bridge with Q42 and D36 of ACE2, forming a strong network of new interactions, thus supporting the impact of these residues on affinity (Fig. 3d). Calculating the electrostatic potential of the RBD-62 in comparison with RBD-WT shows that RBD-62 has a much more positively charged surface, which is complementary to the negatively charged RBD-binding surface on ACE2 (Fig. 3e). In addition, mutation N477 interacts with S19 of ACE2 (Fig. 3c). The ACE2–RBD interface involves the interaction of amino acid residues from the N-terminal segment (Q24–Q42), K353 and D355 of the ACE2 domain and residues from the RBM domain of the RBD. The S477N, Q498R and N501Y mutations in RBD-62 are situated at the two extremes of the RBD–ACE2 interface, and thereby stabilize the complex (Fig. 3b).

RBD-62 inhibits SARS-CoV-2 infection without affecting enzymatic activity of ACE2. One aim of our study was to generate a tight inhibitor drug to compete with host cell ACE2 binding by
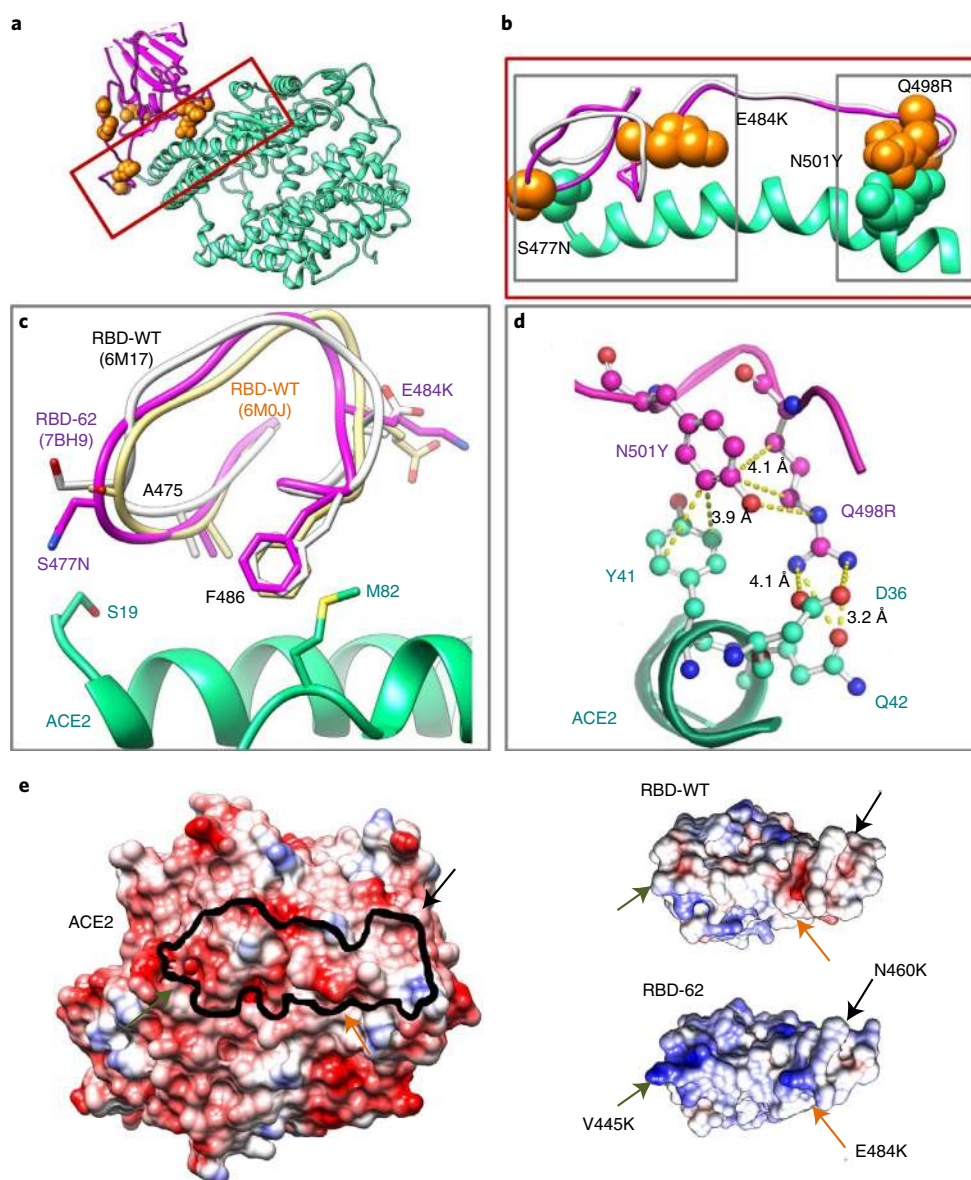
Fig. 3 | Cryo-EM structure of the ACE2–RBD-62 complex at 2.9 Å resolution. a, Cartoon representation of the Cryo-EM structure of ACE2 (cyan) in complex with RBD-62 RBM (magenta) with the RBD-62 mutations resolved in the density (orange spheres). **b**, The S477N, Q498R and N501Y mutations depicted in the RBM (orange spheres) interact with S19, Q42 and K353 of ACE2 (cyan spheres), respectively. These interacting residues are located at the two extremes of the RBD–ACE2 interface, stabilizing the complex. **c,d**, Molecular details of selected interactions depicted in the left (**c**) and right (**d**) grey boxes in **b** contributing to higher affinity. **c**, The presence of E484K and S477N causes repositioning of the RBD-62 loop between amino acids 475 and 487. The loop of RBD-WT is shown for comparison (PDB ID 6M17 in white and 6M0J in pale yellow). The change of main chain positions moves F486 to the optimum position for methionine–aromatic interaction[50]. **d**, The interaction network formed between RBD-62 mutated residues and ACE2. **e**, Open book presentation of ACE2 and RBD-WT and RBD-62. The electrostatic potential was calculated using the program UCSF Chimera v1.13, with negative potential in red and positive in blue. Electrostatic complementarity between RBD and ACE2 is strengthened in RBD-62 by positive charges at V445K, N460K, E484K and Q498R. The black line on ACE2 indicates the RBD-binding site.

SARS-CoV-2. Any ACE2-binding moiety might affect the enzymatic activity of ACE2, which is important for the renin–angiotensin system[18,19]. To determine whether RBD-62 affects host cell ACE2, we assayed the impact of RBD-WT and RBD-62 proteins on ACE2 activity. We found no effect, using either in vitro experiments or assays using various cell lines expressing ACE2 (Fig. 4a).

The effect of RBD-WT and RBD-62 on viral entry was first explored using lentivirus pseudotyped with spike protein variant[27] SΔC19. Results in Supplementary Fig. 6a show that the half-maximum effective concentration (EC$_{50}$) was reduced from 88 nM for RBD-WT to 5.1 nM for RBD-62. Next, RBD-WT and

RBD-62 were evaluated for their potency in inhibiting SARS-CoV-2 infection of Vero E6 cells (Fig. 4b). We determined inhibition of the WT strain (BetaCoV/France/IDF0372/2020), as well as B.1.1.7, B.1.351 and P.1 virus variants with RBD-62 and RBD-WT. The half-maximum inhibitory concentration (IC$_{50}$) and 90% maximum inhibitory concentration (IC$_{90}$) values were 100- to 1,000-fold lower for RBD-62, with concentrations in the pM range for all variants (whereas high nM concentrations were required using RBD-WT). RBD-62 completely blocked viral entry and replication at concentrations in the 10 nM range (Fig. 4b, grey zone), whereas RBD-WT reduced viral load by only 2–3 orders of magnitude at a
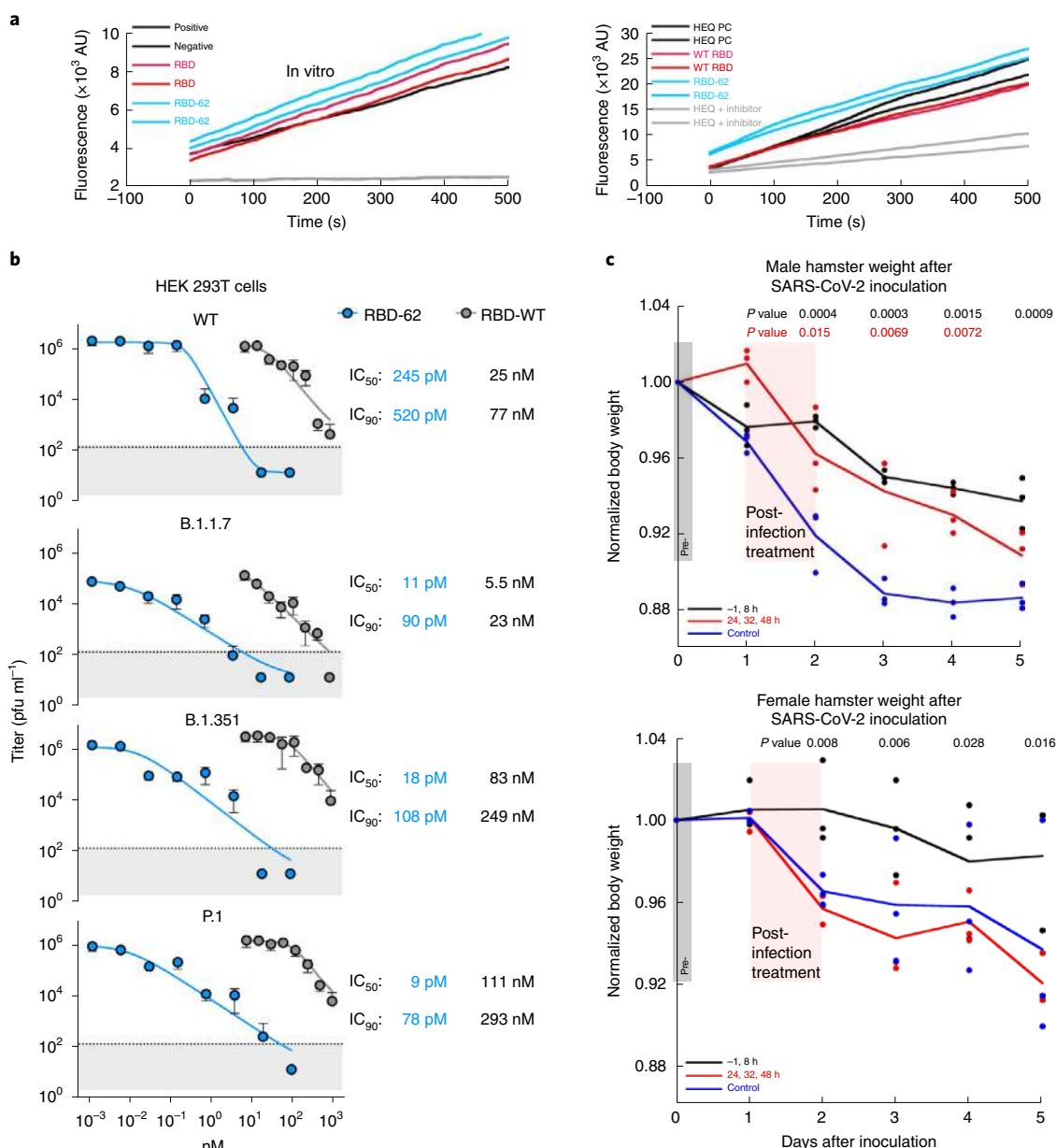
Fig. 4 | RBD-62 treatment inhibits SARS-CoV-2 entry and reduces infection in a hamster model. a, Enzymatic activity of ACE2 (in vitro or on cells) in the presence of RBD-WT and RBD-62 (100 nM) or absence (positive control) as assayed using the SensoLyte 390 ACE2 Activity Assay Kit, which measures fluorogenic peptide cleavage. The activity rate is indicated by the slope of the plot (product/time). An ACE2 inhibitor, provided with the kit, was used as the negative control. AU, arbitrary units. **b,** Inhibition of infection of Vero E6 cells by SARS-CoV-2 WT, B.1.1.7, B.1.351 and P.1 variants by RBD-WT and RBD-62 proteins. Data are mean ± s.e.m. of 6 repeats. The number of viruses was determined by plaque assay (Methods). Curves were fitted to the [agonist] versus response variable slope function in Prism. The grey area is below the limit of detection. **c,** Syrian hamsters (*Mesocricetus auratus*) were infected with SARS-CoV-2 at $t = 0$. RBD-62 (100 µg) was administered by aerosol inhalation, either 1 h before or 8 h after viral challenge (grey band) or 24, 32 and 48 h after viral challenge (pink area). Each group comprised three male and three female hamsters. Control hamsters were treated with PBS. Body weight is shown normalized to the body weight at day zero. Differences were tested with the two-sided linear mixed-effects model, and P values were corrected for multiple comparisons by Tukey's test. P values are shown.

much higher concentration (1 µM). RBD-62 is therefore a promising drug candidate.

RBD-62 efficacy in Syrian hamster challenge model. We measured the capacity of RBD-62 to block SARS-CoV-2 infection in vivo using a highly pathogenic challenge model in golden Syrian hamsters (*M. auratus*)[28]. Hamsters were infected intranasally with 5×10^3 plaque-forming units (PFU) of SARS-CoV-2 USA-WA1/2020, and RBD-62 was administered via inhalation using an Aerogen Solo

nebulizer. Hamsters were treated either twice on the day of challenge (1 h before and 8 h after challenge (−1, 8 h regimen)) or 3 times over the following 48 h (24, 32 and 48 h after challenge). One-hundred micrograms of RBD-62 was delivered intranasally by aerosol at each time point. Figure 4c shows that the −1, 8 h regimen resulted in a significant reduction in weight loss for both male and female hamsters and that this protection lasted for the entire 5 d time course compared to the PBS controls. The post-challenge-only treatment regimen (24, 32 and 48 h) also resulted in a significant reduction

in weight loss, but only in males. The difference in protection between male and female hamsters is probably due to documented gender-specific differences, with more severe disease in male hamsters after SARS-CoV-2 infection[28]. We also determined viral loads using oral swabs 2 d after infection (Supplementary Fig. 6b,c), and found a reduction (although below statistical significance) for the post-infection treatment. Oral swabs may not be ideal to determine viral loads in the lungs, which are mostly inhibited by inhalation of RBD-62. The improved infection outcome in hamsters following treatment with RBD-62 is supportive of our in vitro data and suggests that this approach may be of therapeutic value to block infection in humans.

Discussion

During the first year of the SARS-CoV-2 pandemic, a small subset of mutations have been shown to have an adaptive advantage. Here we report the application of an in vitro evolution method to select for higher affinity binding of the SARS-CoV-2 spike RBD to the host cell receptor ACE2. We found mutations selected by natural infection, S477N, E484K and N501Y, among the first selected by yeast surface display affinity maturation of RBD against ACE2 (Table 1). Moreover, the E484K/N501Y variant was the tightest binding mutation emerging from the B3 library. These two mutations evolved independently, but merged quite rapidly into single clones (Table 1 and Extended Data Fig. 6). The S477N mutation was selected in B3, but was fixed only in B6(FA). This is related to the smaller gain in affinity with S477N compared with N501Y (Fig. 2c). In vitro selection is expected to follow the affinity gain, but it was surprising to observe that natural SARS-CoV-2 selection follows the same path.

In 2020, the SARS-CoV-2 D614G mutation emerged, and is now prevalent. Then, the more infectious B.1.1.7, B.1.351 and P.1 variants were detected[29]. Among 42 non-synonymous mutations delineating the latter variants (PANGO lineages (https://cov-lineages.org/)), N501Y is fixed, whereas the K417N/T and E484K mutations are only found in the B.1.351 and P.1 variants. Recently, an additional virus variant named P.3 was detected in the Philippines with the same combination of E484K and N501Y mutations. Although our results (Fig. 2c,g) showed the cooperative effect of the K417N/T, reversions of K417N/T are also frequent in both the B.1.351 (for example, Finland/THL-202109511/2021) and P.1 strains (for example, Brazil/SP-770679-NBD14/2021) suggesting ongoing pressure for strain optimization. Extended Data Fig. 9 shows how B.1.1.7 (Alpha) outcompeted the WT (20A), D614G (20B) and S477N (20C) variants in the UK and how B.1.351 (Beta) outcompeted 20B in South Africa. Most interesting is the P.1 (Gamma) lineage, which outcompeted all the other variants, including B.1.1.7 in Brazil, strongly suggesting that its higher affinity provides the variant with an advantage in infectivity.

The main advantage of in vitro evolution is that it can mimic natural evolution but on a much faster timescale. We pooled the selected clones from each library to construct a new library, resulting in parallel sampling of a large sequence space in multiple trajectories (Fig. 1). This strategy enabled the emergence of epistatic mutations. The use of random, rather than saturation mutagenesis, mimics virus evolution for tighter binding during infection of naive hosts.

RBD-62, which has an affinity 1,000-fold higher than that of RBD-WT, was the tightest binding clone that we identified. As shown in Table 1 and Extended Data Fig. 6, many mutations were transient during the selection process, whilst a small number persisted (as is also observed in SARS-CoV-2 isolates). Q498R was first observed in library B4, but established itself only in B5. This mutation emerges only together with N501Y, suggesting that there may be an epistatic effect between the two mutations. Indeed, deep-mutational scanning of single RBD mutations[7] showed Q498R to negatively affect both protein stability and binding, while we show that in combination

with N501Y binding, affinity is increased by approximately fourfold above that of N501Y alone. The epistatic effect is also confirmed by calculating the interaction energy between the two residues using a double-mutant cycle (Fig. 2g) and by the RBD-62–ACE2 complex structure, where Y501 positions R498 to form multiple interactions (Fig. 3d). Single-mutation scanning does not account for this cooperativity, which has a crucial role at protein–protein interfaces[30]. The rapid spread of N501Y in the population increases the likelihood of the emergence of the Q498R mutant, which may have higher infectivity. Unlike Q498R, the N501Y and E484K mutations established themselves independently (their coupling energy is not significant; Fig. 2g), allowing for their rapid in vitro selection and emergence in SARS-CoV-2. Owing to the spread of the B.1.1.7, B.1.351, P.1 and P.3 variants, the Q498R mutation may emerge in future, on top of these mutations; the synergism of Q498R with N501Y and E484K in our assays increased ACE2 binding by around 50-fold compared with the WT. Another mutation that emerged late in yeast display is N460K. Both Q498R and N460K are located in a hypervariable region of the RBD (Supplementary Fig. 5), suggesting that their emergence is not constrained.

Infectivity is one concern for emerging mutations, but equally important is their potential for immune evasion, both from the resistance provided by previous infections of the WT virus and from vaccination. To evaluate the effect of these RBD mutations on antibody binding, we inspected 92 antibody–RBD (nanobody–S) structures for clashes in silico, replacing RBD-WT with the RBD-62 structure, as determined in complex with ACE2 using Cryo-EM. Twenty-eight of the antibodies bind outside the RBM, and 8 interactions are not affected by the mutations on RBD-62. However, interaction was compromised for 56 antibodies, and major clashes with RBD-62 were identified for 9 antibodies (Extended Data Fig. 10). Notably, E484K and Q498R caused most of the observed effects on predicted antibody binding.

RBD-62 has potential as a drug, as it blocks ACE2 with very high affinity, while leaving ACE2 functionality unaffected. It completely inhibited SARS-CoV-2 infection of Vero E6 cells, including by the most common SARS-CoV-2 variants, at a picomolar concentration. Furthermore, we tested the efficacy of RBD-62 in blocking COVID-19 in a highly pathogenic hamster challenge model, and found a reduction in disease (with weight loss as the proxy) when RBD-62 was administered either before or after infection. With post-infection treatment, significant differences were observed only for male hamsters, which exhibit more severe disease than female hamsters[28]. The main advantage of blocking host cell ACE2 is that it is not affected by viral escape mutations, which evolve continuously. A follow-up primate experiment using RBD-62 to block SARS-CoV-2 was approved by the Animal Care and Use Committee of the Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes for Health.

Knowing the enemy provides an advantage to the defender. Using in vitro evolution, we have successfully delineated the pathway of SARS-CoV-2 variant emergence. Although SARS-CoV-2 must evolve under multiple selection pressures (such as shown for the D614G mutation of the spike protein), we found that an increase in RBD–ACE2 binding affinity can be replicated using affinity-binding selection pressure alone. We advise that sequencing of the RBD of SARS-CoV-2 natural isolates will be a crucial tool in the early identification of variants with greater infectivity. Moreover, such knowledge may enable shortened timelines to vaccine and drug countermeasure development and deployment. We note that the methods applied in our study could be generalized for other viruses once their evolutionary path is identified.

Methods

Cloning and DNA manipulations. The RBD domain variants (Supplementary Table 1) were PCR amplified (KAPA HiFi HotStart ReadyMix, Roche, catalogue

(cat.) no. KK3605) from codon-optimized SARS-CoV-2 Spike protein gene (Sino Biological, SARS-CoV-2 (2019-nCoV) cat. no. VG40589-UT; GenBank: QHD43416.1) using appropriate primers (Supplementary Table 3). Amplicons were purified using NucleoSpin Gel and PCR Clean-up kit (Nacherey-Nagel, cat. no. 740609.50) and eluted in double-distilled water. Yeast surface display plasmid pJYDC1 (Adgene ID: 162458) and pJYDC3 (162460) were cleaved by NdeI and BamHI (NEB, cat. no. R0111 and R0136) restriction enzymes, purified, and tested for non-cleaved plasmids via transformation to *E. coli* Cloni 10 G cells (Lucigen). Each amplicon was mixed with cleaved plasmid in the ratio: 4 µg insert:1 µg plasmid per construct, electroporated in *Saccharomyces cerevisiae* EBY100[31], and selected by growth on SD-W plates (for detailed composition see ref. [21]).

Cloning of the ACE2 extracellular domain (amino acids G17–Y613) and RBDs into vectors pHL-sec[32] were done in two steps. Initially, the RBD gene was inserted in helper vector pCA by restriction-free cloning[33]. pCA is a pHL-sec derivative lacking 862 bp in the GC rich region (nucleotides 672–1534). In the second step, the correctly inserted, verified by sequencing, RBD coding sequence with flanking sequences was cleaved using restriction enzymes XbaI and XhoI (NEB, cat. nos R0145 and R0146) and ligated (T4 DNA ligase, NEB, cat. no. M0202) in cleaved full-length plasmid pHL-sec.

Site-directed mutagenesis of RBDs was performed by restriction-free cloning[33]. Megaprimers were amplified by KAPA HiFi HotStart ReadyMix (Roche), purified with NucleoSpin Gel and PCR Clean-up Kit (Nachery-Nagel), and subsequently inserted by PCR in the destination using high fidelity Phusion (NEB) or KAPA polymerases. The parental plasmid molecules were inactivated by DpnI treatment (1 h, NEB, cat. no. R0176) and the crude reaction mixture was transformed into electrocompetent *E. coli* Cloni 10 G cells (Lucigen). The clones were screened by colony PCR and their correctness was verified by sequencing. All primers used in this study are listed in Supplementary Table 3.

DNA libraries preparation. SARS-CoV-2 RBD gene libraries were prepared by MnCl$_2$ error-prone mutagenesis[34] using Taq Ready-mix (Hylabs, cat. no. EZ3006). The mutagenic PCR reactions (50 µl) were supplemented with increasing MnCl$_2$ concentrations: 0.05, 0.1, 0.2, 0.4, 0.6, 0.8 and 1.0 nM. Template DNA concentration ranged between 100 and 400 ng per reaction and 20–30 reaction cycles were applied. The amplified DNA was purified, pooled and used directly for yeast transformation via electroporation. The whole gene randomization amplicon comprised the whole RBD gene (amino acids 336–528, pJYDC1 vector). Libraries B3, B4 and B5 were prepared by homologous recombination of an invariant fragment of RBD with necessary overlaps (amino acids 336–431) and the mutagenized library fragment (amino acids 431–528). The mutagenic fragments were prepared by the same error-prone PCR procedure (20 cycles).

Yeast transformation, cultivation and expression procedures. The procedures and the enhanced yeast display platform have been described in detail elsewhere[21]. In brief, plasmids were transformed into EBY100 *S. cerevisiae*[31,35]. Single colonies were inoculated into 1.0 ml liquid SD-CAA medium[21] and grown overnight at 30 °C (220 rpm). The overnight cultures were centrifuged (3,000 g, 3 min) and the exhausted culture medium was removed before 1:9 dilution in the expression medium[21] to an optical density of approximately 1. Expression cultures were grown at 20, 30 or 37 °C for 8–24 h at 220 rpm, depending on the experimental setup. The expression co-cultivation labelling was achieved by the addition of 1 nM DMSO solubilized bilirubin (pJYDC1, eUnaG2 reporter holo-form formation, green/yellow fluorescence (excitation 498 nm, emission 527 nm), Sigma, cat. no. B4126) or 5 nM ALFA-tagged mNeonGreen (pJYDC3, DnbALFA). Aliquots of cells (100 µl) were collected by centrifugation (3,000 g, 3 min) resuspended in ice-cold PBS with 1 g l^{-1} BSA (PBSB), passed through a nylon membrane cell strainer (40 µM, SPL Life Sciences) and analysed.

Binding assays, affinity determination and selection for tighter binding clones using yeast surface display. Aliquots of yeast expressed and labelled cells ready for flow cytometry analysis were resuspended in analysis solution with a series of labelled ACE2 concentrations. The concentration range of CF640 R succinimidyl ester (Biotium, cat. 92108) labelled ACE2 extracellular domain (amino acids Q18–S740) was dependent on the protein analysed (0.1 pM – 50 nM). The volume of analysis solution (between 1 and 100 ml to avoid ligand depletion (>10%)) and the time needed to reach the equilibrium (1 h – 12 h, 5 rpm, 4 °C)[36] were adjusted for each experiment according to the expected K_D. After the incubation, samples were collected (3,000 g, 3 min), resuspended in 200 ul of ice-cold PBSB (200 µl), passed through a cell strainer, and analysed. The expression and binding signals were determined by flow cytometry using BD Accuri C6 Flow Cytometer (BD Biosciences). The cell analysis and sorting were done by S3e Cell Sorter (BioRad, USA). The analysis was done by single-cell event gating (Extended Data Fig. 2), the green fluorescence channel (FL1-A) was used to detect RBD-expressing cells (RBD$^+$) via eUnaG2 or DnbALFA, and the far-red fluorescent channel (FL4-A) recorded CF640 R-labelled ACE2-binding signals (CF640$^+$). The eUnaG2 signals were automatically compensated by the ProSort Software and pJYDNp positive control plasmid (Adgene ID 162451)[21]. The mean FL4-A fluorescence signal values of RBD$^+$ cells, subtracting the fluorescence of RBD$^-$ cells, was used to determine the binding constant $K_{D,YD}$. The standard non-cooperative Hill equation was fitted

using a nonlinear least-squares regression in Python 3.7. The total concentration of yeast exposed protein was fitted together with two additional parameters describing the given titration curve[7].

The FACS selection and gating strategy is shown in Fig. 1 and Extended Data Fig. 2. Every library was selected by three consecutive rounds of FACS. In the first selection round, less stringent conditions were applied, with the 3% best binding cells sorted. Selected cell populations were cultivated and expressed before the next sorting round. In the next two rounds, the sorting threshold was strengthened to select the top 0.1–1% of yeast cells. A minimum of 20,000 cells was sorted.

Double-mutant cycle calculations. $K_{D,YD}$ values were used to calculate coupling free energy between pairs of amino-acids, using the double-mutant cycle method using the following equation, $\Delta\Delta G_{int} = \Delta G_{ab} + \Delta G_{cd} - \Delta G_{ad} - \Delta G_{cb}$, where a and b are the WT residues and c and d are the mutant residue. Each of the ΔG values was calculated from $K_{D,YD}$ according to $\Delta G = -RT\ln K_{D,YD}$, where R is the gas constant and T the absolute temperature. The standard errors were calculated from the 3–6 repeats of independent $K_{D,YD}$ measurements, from which the standard error shown in Fig. 2g was calculated.

Production and purification of RBD and ACE2 proteins. The extracellular part of ACE2 (Q18–S740) and RBD protein variants were produced in Expi293F cells (ThermoFisher). Pure DNA was transfected using ExpiFectamine 293 Transfection Kit (ThermoFisher) using the manufacturer's protocol. Seventy-two hours after transfection, the cells were centrifuged at 1,500 rpm for 15 min. The supernatant was filtered using 0.45 µm Nalgene Rapid-Flow sterile single-use vacuum filter units (ThermoFisher, cat. no. 564–0020) and the pellet was discarded. The filtered supernatant was loaded onto a 5 ml of HisTrap Fast Flow column (Cytivia (GE Healthcare), cat. no. 17–5255–01). ÄKTA pure (Cytivia) was used to purify the protein. The column was washed in 25 mM Tris, 200 mM NaCl 20 mM imidazole pH 7.2, then the protein was eluted using gradient elution with elution buffer containing 25 mM Tris, 200 mM NaCl 1 M imidazole pH 7.2. Buffer exchange to PBS and the concentration of the protein were done by using Amicon concentrators (3 K MWCO, Merck Millipore, cat. no. UFC900324).

Cryo-electron microscopy. *Sample preparation.* ACE2–RBD-62 complex (2.5 µl at 3.5 mg ml^{-1}) was transferred to glow-discharged UltrAuFoil R 1.2/1.3 300 mesh grids (Quantifoil, cat. 4976G-FA), blotted for 2.5 s at 4 °C, 100% humidity, and plunge frozen in liquid ethane cooled by liquid nitrogen using a Vitrobot plunger (ThermoFisher Scientific).

Cryo-EM image acquisition. Cryo-EM data were collected on a Titan Krios G3i transmission electron microscope (ThermoFisher Scientific) operated at 300 kV. Movies were recorded on a K3 direct detector (Gatan) installed behind a BioQuantum energy filter (Gatan), using a slit of 20 eV. Movies were recorded in counting mode at a nominal magnification of ×165,000, corresponding to a physical pixel size of 0.53 Å. The dose rate was set to 16.2 e$^-$ pixel^{-1} s^{-1}, and the total exposure time was 1.214 s, resulting in an accumulated dose of 70 e$^-$ Å^{-2}. Each movie was split into 57 frames of 0.021 s. The nominal defocus range was −0.7 to −1.1 µm, however, the actual defocus range was larger. Imaging was done using an automated low-dose procedure implemented in SerialEM[37]. A single movie was collected from the centre of each hole using image shift to navigate within hole arrays and stage shift to move between arrays. The 'multiple record setup' together with the 'multiple hole combiner' dialogues were used to map hole arrays of up to 3×3 holes. Beam tilt was adjusted to achieve coma-free alignment when applying image shift.

Cryo-EM image processing. Image processing was performed using CryoSPARC software[38] v3.0.1. The processing scheme is outlined in Supplementary Fig. 2. A total of 4,470 acquired movies were subjected to patch motion correction, followed by patch contrast transfer function estimation. Of these, 3,357 micrographs having contrast transfer function fit resolution better than 5 Å and relative ice thickness lower than 1.07, were selected for further processing. Initial particle picking was done using the 'blob picker' job on a subset of 100 micrographs. Extracted particles were iteratively classified in 2D and their class averages were used as templates for automated particle picking from all selected micrographs, resulting in 2,419,995 picked particles. Particles were extracted, binned 6×6 (60-pixel box size, 3.18 Å per pixel), and cleaned by multiple rounds of 2D classification, resulting in 1,649,355 particles. These particles were used for ab initio 3D reconstruction with 5 classes. Out of the 5 classes only one, containing 552,575 particles, refined to high resolution. Two additional classes may show ACE2 in a closed conformation (containing 249,841 and 503,670 particles), however, they did not refine, partially because of preferred orientation. The 3D class containing 552,575 particles was refined as follows: particles were re-extracted only from micrographs with defocus lower than 1.7 µm, binned 2×2, and subjected to homogeneous refinement (355,891 particles, 200-pixel box size, 1.06 Å per pixel). The particles were then sub-classified into 2 classes, and particles from the higher-resolution class were re-extracted without binning in 680-pixel boxes, subjected to per-particle motion correction, followed by non-uniform refinement[39] with per-particle defocus optimization. The final map, at a resolution of 2.9 Å (Supplementary Fig. 3), was

sharpened with a B-factor of −83 before atomic model building. In the final map, the RBD is only partially resolved at the distal region from the ACE2 interface. To better understand the reason for the missing density, we subjected the particles from the well-refined 3D class (355,891 particles) to variability analysis[40], with a binary mask imposed on the RBD region (Supplementary Fig. 4). Classification into five distinct classes based on three eigenvectors, revealed variable density at the RBD distal region, which could not be modelled reliably. The cryo-EM data collection process and refinement statistics are summarized and visualized in Supplementary Figs. 2–4 and Supplementary Table 2.

Model building. The atomic model of the ACE2–RBD-62 was solved by docking into the Cryo-EM maps the homologous refined structure of the SARS-CoV-2 spike receptor-binding domain bound with ACE2 (PDB ID 6M0J) as a model, using the Dock-in-Map program in PHENIX[41]. All steps of atomic refinements were carried out with real-space refinement in PHENIX[42]. The model was built into the cryo-EM map by using the COOT program[43]. The ACE2–RBD-62 model was evaluated with the MolProbity program[44]. ACE2 (G17–Y613) contains one zinc ion linked to H374, H378 and E402 and three *N*-acetyl-β-glucosaminide glycans linked to N53, N90 and N546. In the RBD-62 structure (T333–K528) three fragments–R357–S371 (β2, α2), G381–V395 (α3) and F515–H534 (β11) are disordered, and thus not visible in the density map. Details of the refinement statistics of the ACE2–RBD-62 structure are described in Supplementary Table 2. Three-dimensional visualization and analyses were performed using UCSF Chimera[45] and PyMol (Schrödinger, 2.4.0).

Analysis of RBD circulating virus variants. All amino acid substitutions in the RBD (116) were downloaded from the GISAID database (13 April 2021)[22] with the corresponding numbers of sequences and regions and plotted against the binding ($\Delta\log_{10}(K_{D, App})$ or $K_{D,YD}$), or expression ($\Delta\log_{10}MFI$) extracted from the RBD deep-mutational scanning dataset[7]. We gratefully acknowledge all GISAID contributors and Starr et all for sharing their data.

Biacore S200 ACE2–RBD binding measurements. The Biacore S200 (GE Healthcare) was used for real-time binding determination. In brief, 10 μg ml⁻¹ of Fc–ACE2, provided by R. Diskin (Weizmann Institute)[27] was bound to a series S sensor chip protein A (Cytivia (GE Healthcare), cat. no. 29127557) in PBST and 0.02% (w/v) sodium azide buffer, leaving a free channel for non-specific binding subtraction. Fusion proteins were first immobilized at a coupling density of ~700 RU; the RBD was then injected at 0, 5, 10, 20, 40, 80 and 160 nM concentrations at a flow rate of 30 μl min⁻¹. The sensor chip was regenerated using 10 mM glycine-HCl pH 1.5 buffer. For data evaluation, the no Fc–ACE2 channel and the zero RBD measurement were subtracted from the sensorgram. The six concentrations were fitted using the built-in global fitting routine, using a 1:1 model using S200 Evaluation Software 1.1.

Pseudovirus production and inhibition of infection by RBD. Pseudovirus production: SARS-CoV-2-Spike pseudotyped lentivirus was produced by co-transfection of HEK 293T cells pCMV ΔR8.2, pGIPZ-GFP, (26) and pCMV3 SΔC19 at a ratio of 1:1:1. Twenty-four hours before the transfection, 1×10⁶ cells were seeded into a 10 cm plate. On the day of the transfection, cells were washed with Dulbecco's Modified Eagle's Medium (DMEM, Gibco, cat. no. 11965092) and 5 ml of Opti-MEM (Gibco, cat. no. 11058021) was added to the plate. Ten micrograms of plasmids mix was transfect using Lipofectamine 2000 transfection reagent (ThermoFisher, cat. 11668027) according to the manufacturer's instructions. After 4 h, the medium was replaced with 9 ml of fresh medium. The supernatant was collected 72 h after transfection, centrifuged (1,000*g*, 5 min), and filtered to remove all residual debris (Millex-HV Syringe Filter Unit, 0.45 μm, Merck, cat. no. SLHVM33RS).

RBD inhibition assay: HEK 293T cells stably expressing hACE2 (GenScript, cat. no. M00770) were seeded into a 24-well plate at an initial density of 6×10⁴ cells per well. The following day, cells were pre-incubated with serial dilutions of RBDs (1 h) and then the pseudotyped lentivirus was added. After 24 h, the cell medium was replaced with fresh DMEM, and cells were grown for an additional 24 h. After this procedure, cells were collected and the GFP signal was analysed by flow cytometry (BD Accuri C6 Plus Flow Cytometer, BD Biosciences).

Inhibition of SARS-CoV-2 infection. The WT and three major variants of SARS-CoV2 were used: strain 2019-nCoV/IDF0372/2020 (WT) was supplied by the National Reference Centre for Respiratory Viruses hosted by Institute Pasteur (Paris, France) and headed by S. van der Werf. The human sample from which strain 2019-nCoV/IDF0372/2020 was isolated was provided by X. Lescure and Y. Yazdanpanah from Bichat Hospital. The B.1.1.7: hCoV-19/France/IDF-IPP11324/2021 and B.1.351: hCoV-19/France/PDL-IPP01065/2021 were supplied through the European Virus Archive goes Global (EVAg) platform. The P.1: hCoV-19/Japan/TY7-501/2021 virus was provided by J. Vanhomwegen (Cellule d'Intervention Biologique d'Urgence (CIBU), Institut Pasteur).

Plaque assay was performed to determine the viral load in each sample (six replicates). Vero E6 cells were seeded in 24-well plates at a concentration of 7.5×10⁴ cells per well. The following day, serial dilutions were performed in

serum-free Minimum Essential Medium (MEM, Merck, cat. no. M5650) medium. After 1 h absorption at 37 °C, 2× overlay medium was added to the inoculum to give a final concentration of 2% (v/v) FBS in MEM and 0.4% (w/v) SeaPrep Agarose (Lonza, cat. no. 50302) to achieve a semi-solid overlay. Plaque assays were incubated at 37 °C for 3 d. Samples were fixed using 4% formaldehyde solution (approximately 10% formalin solution) for histology (Merck, cat. no. 1.00496) and plaques were visualized using crystal violet solution (Merck, cat. no. C0775).

ACE2 activity assay. Human ACE2 activity was evaluated using SensoLyte 390 ACE2 Activity Assay Kit (AnaSpec, cat. no. 72086) according to the manufacturer's protocol, with the following changes: the assay was performed in 384-well plates with a ratio of 1:5 of the recommended volume of buffer, substrate and inhibitor. The activity was measured on either purified ACE2 (0.75 ng; Abcam, cat. no. aB351852) or the following cell lines—HeLa cells transiently transfected with ACE2 (6,000 cells per assay), HEK 293T stable transfected with ACE2 (GenScript, cat. no. M00770, 8,000 cells per assay), Caco2 cells (40,000 cells per assay). RBD-WT or RBD-B62 were added at a concentration of 10 nM before activity measurement. The activity rate is indicated by the slope of the plot (product/time).

Golden Syrian hamster challenge experiments. Golden Syrian hamsters (*M. auratus*) were housed at Bioqual. Hamsters were 6–8 weeks old on the day of infection and were randomized into groups of 3 female and 3 male hamsters each per group. Hamsters were infected with 100 μl of 5×10³ PFU (USA-WA1/2020; BEI: NR-52281) intranasally on day 0. Hamsters were weighed daily and weight loss was measured as percent change from the baseline weight on day 0. Hamsters were treated with RBD-62 via an Aerogen Solo Nebulizer (Aerogen). RBD-62 was prepared in 100 μg aliquots in 1.5 ml of 0.5× PBS with 2 mg ml⁻¹ gelatin (Merck, cat. no. G1890). Gelatin (2 mg ml⁻¹) in 0.5× PBS was administered via inhalation to hamsters in the PBS control group. Oral swabs were collected 2 d after challenge and immediately placed in 1 ml PBS and frozen at −80 °C until RNA extraction. Specimens were thawed at 50 °C, and the swab was removed. The remaining PBS was mixed with 2 ml RNAzol BD (Molecular Research Center, cat. no. RB 192) and 20 μl acetic acid. Total RNA was extracted as described previously[46].

Defining the RBD domain for yeast display and protein expression. To optimize the RBD for yeast display affinity maturation, we screened multiple different constructs for yeast surface expression. RBDs with different starting and termination positions were cloned in a pJYDC1 vector and their impact on expression, stability and ACE2 binding were determined (Supplementary Table 1). RBDcon1 was the shortest construct, lacking the last C-terminal loop of the RBD domain (516–528) and including one unpaired cysteine. This resulted in poor expression and binding. RBDcon2 and RBDcon3 included this loop, resulting in domain stabilization and an increase in both binding and expression. Although the RBDcon4 (published by Kramer and colleagues[47]) construct demonstrated high expression yields both in yeast and Expi293F cells, as well as good thermostability, we decided to not use it in yeast display, since one unpaired cysteine (C538) is close to its C terminus and the construct contains part of the neighbouring domain. We continued with the RBDcon2 and RBDcon3 constructs for yeast display and protein expression in Expi293F cells, respectively.

Statistics and reproducibility. Unless stated otherwise, three independent biological replicates were used for measurement. Data are reported as mean values ± s.e.m.. Differences in Syrian hamster weight were tested with two-sided linear mixed-effects model, with treatment, sex, day, treatment × sex interaction and treatment × day interaction as fixed factors, and animal ID as a random factor. Specific contrasts were tested by estimated marginal means, and *P* values were corrected for multiple comparisons by Tukey's test. Analysis was done in R, v. 4.0.4, using the packages lme4, lmerTest and emmeans.

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

Data availability
Cryo-EM data have been deposited in the Electron Microscopy Data Bank with accession codes EMD-12187 and PDB ID: 7BH9. Source data are provided with this paper.

References
1. Hoffmann, M. et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* **181**, 271–280.e8 (2020).
2. Lan, J. et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **581**, 215–220 (2020).

3. Chen, Y. et al. ACE2-targeting monoclonal antibody as a "pan" coronavirus blocker in vitro and in a mouse model. Preprint at *bioRxiv* https://doi.org/10.1101/2020.11.11.37597 (2020).

4. Wrapp, D. et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**, 1260–1263 (2020).

5. Tai, W. et al. Characterization of the receptor-binding domain (RBD) of 2019 novel coronavirus: implication for development of RBD protein as a viral attachment inhibitor and vaccine. *Cell. Mol. Immunol.* **17**, 613–620 (2020).

6. Walls, A. C. et al. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* **181**, 281–292.e6 (2020).

7. Starr, T. N. et al. Deep mutational scanning of SARS-CoV-2 receptor binding domain reveals constraints on folding and ACE2 binding. *Cell* **182**, 1295–1310.e20 (2020).

8. Tegally, H. et al. Detection of a SARS-CoV-2 variant of concern in South Africa. *Nature* **592**, 438–443 (2021).

9. Vasques Nonaka, C. K. et al. Genomic Evidence of SARS-CoV-2 reinfection involving E484K spike mutation, Brazil. *Emerg. Infect. Dis.* **27**, 1522–1524 (2021).

10. Chen, J., Wang, R., Wang, M. & Wei, G.-W. Mutations strengthened SARS-CoV-2 infectivity. *J. Mol. Biol.* **432**, 5212–5226 (2020).

11. Dai, L. & Gao, G. F. Viral targets for vaccines against COVID-19. *Nat. Rev. Immunol.* **21**, 73–82 (2021).

12. Nile, S. H. et al. COVID-19: Pathogenesis, cytokine storm and therapeutic potential of interferons. *Cytokine Growth Factor Rev.* **53**, 66–70 (2020).

13. Barnes, C. O. et al. SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. *Nature* **588**, 682–687 (2020).

14. Guo, L. et al. Engineered trimeric ACE2 binds viral spike protein and locks it in "three-up" conformation to potently inhibit SARS-CoV-2 infection. *Cell Res.* **31**, 98–100 (2021).

15. Abd El-Aziz, T. M., Al-Sabi, A. & Stockand, J. D. Human recombinant soluble ACE2 (hrsACE2) shows promise for treating severe COVID19. *Signal Transduct. Target. Ther.* **5**, 258 (2020).

16. Schütz, D. et al. Peptide and peptide-based inhibitors of SARS-CoV-2 entry. *Adv. Drug. Deliv. Rev.* **167**, 47–65 (2020).

17. Cao, L. et al. De novo design of picomolar SARS-CoV-2 miniprotein inhibitors. *Science* **370**, 426–431 (2020).

18. Lelis, D. D. F., Freitas, D. F. D., Machado, A. S., Crespo, T. S. & Santos, S. H. S. Angiotensin-(1-7), adipokines and inflammation. *Metabolism* **95**, 36–45 (2019).

19. Zhang, H., Penninger, J. M., Li, Y., Zhong, N. & Slutsky, A. S. Angiotensin-converting enzyme 2 (ACE2) as a SARS-CoV-2 receptor: molecular mechanisms and potential therapeutic target. *Intensive Care Med.* **46**, 586–590 (2020).

20. Wang, G. et al. Dalbavancin binds ACE2 to block its interaction with SARS-CoV-2 spike protein and is effective in inhibiting SARS-CoV-2 infection in animal models. *Cell Res.* **31**, 17–24 (2021).

21. Zahradník, J., Dey, D., Marciano, S. & Schreiber, G. An enhanced yeast display platform demonstrates the binding plasticity under various selection pressures. Preprint at *bioRxiv* https://doi.org/10.1101/2020.12.16.423176 (2020).

22. Elbe, S. & Buckland-Merrett, G. Data, disease and diplomacy: GISAID's innovative contribution to global health. *Glob. Chall.* **1**, 33–46 (2017).

23. Meng, B. et al. Recurrent emergence of SARS-CoV-2 spike deletion H69/V70 and its role in the Alpha variant B.1.1.7. *Cell Rep.* **35**, 109292 (2021).

24. Cohen-Khait, R. & Schreiber, G. Selecting for fast protein–protein association as demonstrated on a random TEM1 yeast library binding BLIP. *Biochemistry* **57**, 4644–4650 (2018).

25. Selzer, T., Albeck, S. & Schreiber, G. Rational design of faster associating and tighter binding protein complexes. *Nat. Struct. Biol.* **7**, 537–541 (2000).

26. Albeck, S. & Schreiber, G. Biophysical characterization of the interaction of the β-lactamase TEM-1 with its protein inhibitor BLIP. *Biochemistry* **38**, 11–21 (1999).

27. Cohen-Dvashi, H. et al. Coronacept—a potent immunoadhesin against SARS-CoV-2. Preprint at *bioRxiv* https://doi.org/10.1101/2020.08.12.247940 (2020).

28. Yuan, L. et al. Gender associates with both susceptibility to infection and pathogenesis of SARS-CoV-2 in Syrian hamster. *Signal Transduct. Target. Ther.* **6**, 136 (2021).

29. Volz, E. et al. Transmission of SARS-CoV-2 Lineage B.1.1.7 in England: insights from linking epidemiological and genetic data. Preprint at *medRxiv* https://doi.org/10.1101/2020.12.30.20249034 (2021).

30. Reichmann, D. et al. The modular architecture of protein–protein binding interfaces. *Proc. Natl. Acad. Sci. USA* **102**, 57–62 (2005).

31. Benatuil, L., Perez, J. M., Belk, J. & Hsieh, C.-M. An improved yeast transformation method for the generation of very large human antibody libraries. *Protein Eng. Des. Sel.* **23**, 155–159 (2010).

32. Aricescu, A. R., Lu, W. & Jones, E. Y. A time- and cost-efficient system for high-level protein production in mammalian cells. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 1243–1250 (2006).

33. Peleg, Y. & Unger, T. Application of the restriction-free (RF) cloning for multicomponents assembly. *Methods Mol. Biol.* **1116**, 73–87 (2014).

34. Wilson, D. S. & Keefe, A. D. Random mutagenesis by PCR. *Curr. Protoc. Mol. Biol.* https://doi.org/10.1002/0471142727.mb0803s51 (2001)..

35. Gietz, R. D. Yeast transformation by the LiAc/SS carrier DNA/PEG method. *Methods Mol. Biol.* **1163**, 33–44 (2014).

36. Chao, G. et al. Isolating and engineering human antibodies using yeast surface display. *Nat. Protoc.* **1**, 755–768 (2006).

37. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005).

38. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).

39. Punjani, A., Zhang, H. & Fleet, D. J. Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nat. Methods* **17**, 1214–1221 (2020).

40. Punjani, A. & Fleet, D. J. 3D Variability analysis: resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM. *J. Struct. Biol.* **213**, 107702 (2021).

41. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).

42. Klaholz, B. P. Deriving and refining atomic models in crystallography and cryo-EM: the latest Phenix tools to facilitate structure analysis. *Acta Crystallogr. D Struct. Biol.* **75**, 878–881 (2019).

43. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).

44. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010).

45. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

46. Corbett, K. S. et al. Immune correlates of protection by mRNA-1273 vaccine against SARS-CoV-2 in nonhuman primates. *Science* **29**, eabj0299 (2021).

47. Amanat, F. et al. A serological assay to detect SARS-CoV-2 seroconversion in humans. *Nat. Med.* **26**, 1033–1036 (2020).

48. Rambaut, A. et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat. Microbiol.* **5**, 1403–1407 (2020).

49. Hadfield, J. et al. Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* **34**, 4121–4123 (2018).

50. Weber, D. S. & Warren, J. J. The interaction between methionine and two aromatic amino acids is an abundant and multifunctional motif in proteins. *Arch. Biochem. Biophys.* **672**, 108053 (2019).

Acknowledgements

Author contributions

J.Z. and G.S. conceived the project; J.Z., S.M., M.S., E.Z., J.C., B.M., M.G.L., H.A., M.G., R.A.S., D.C.D., Y.R., C.L., I.M., D.H. and G.S. performed experiments; N.E. prepared cryo-EM samples, built atomic models and refined structures with O.D.; J.Z, S.M., N.E., O.D, Y.R., M.G.L, D.C.D. and G.S. wrote the manuscript.

Competing interests

J.Z., G.S., Y.R., C.L., I.M., D.H, M.S., S.M. and E.Z. declare the Israel patent application no. 23/09/2020—277546 and US patent application no. 16/12/2020—63/125,984 with the title *Methods and Compositions for Treating Coronaviral Infections*. J.C., B.M., O.D., N.E., M.G.L., H.A., M.G., R.A.S. and D.C.D. declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41564-021-00954-4.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41564-021-00954-4.

**Correspondence and requests for materials** should be addressed to G.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
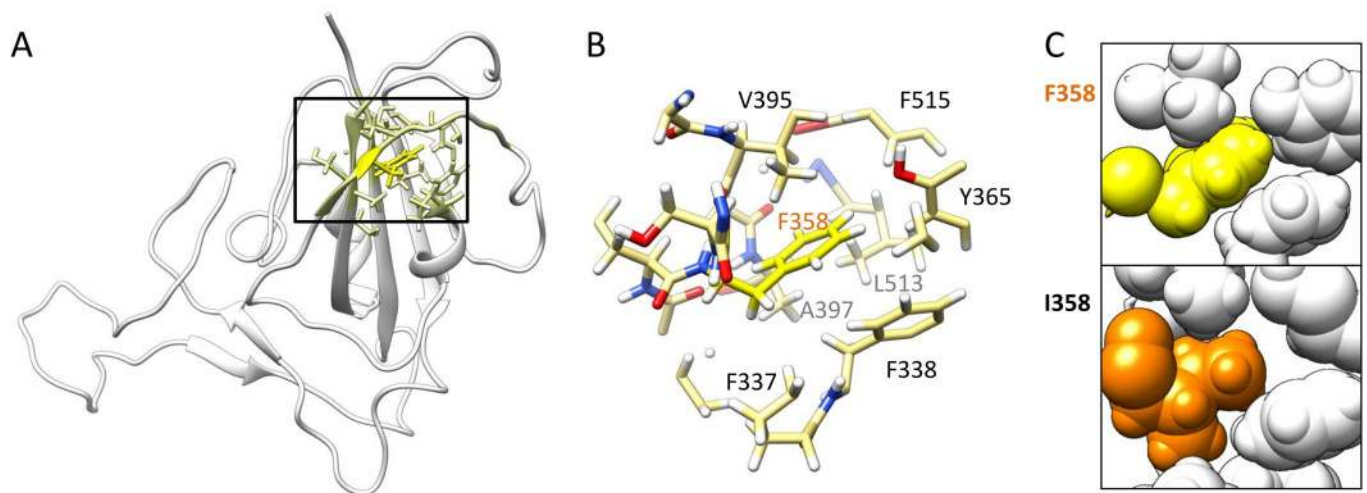
Extended Data Fig. 1 | The I358F mutation, selected by yeast surface display, increases protein stability and expression. (**A**) The position of the I358F (bright yellow) mutation in the RBD structure (PDB ID 6M17) and the neighboring residues within 5 Å distance (pale yellow). (**B**) Shows the residues involved in the formation of the hydrophobic cavity around the I358F mutation predicted from the X-ray structure. Additional residues that are involved: K356, R357, S359, V395, Y396. (**C**) The spherical representation of the predicted position of the phenylalanine mutation inside the hydrophobic cavity (yellow) and wild-type residue (isoleucine, orange).
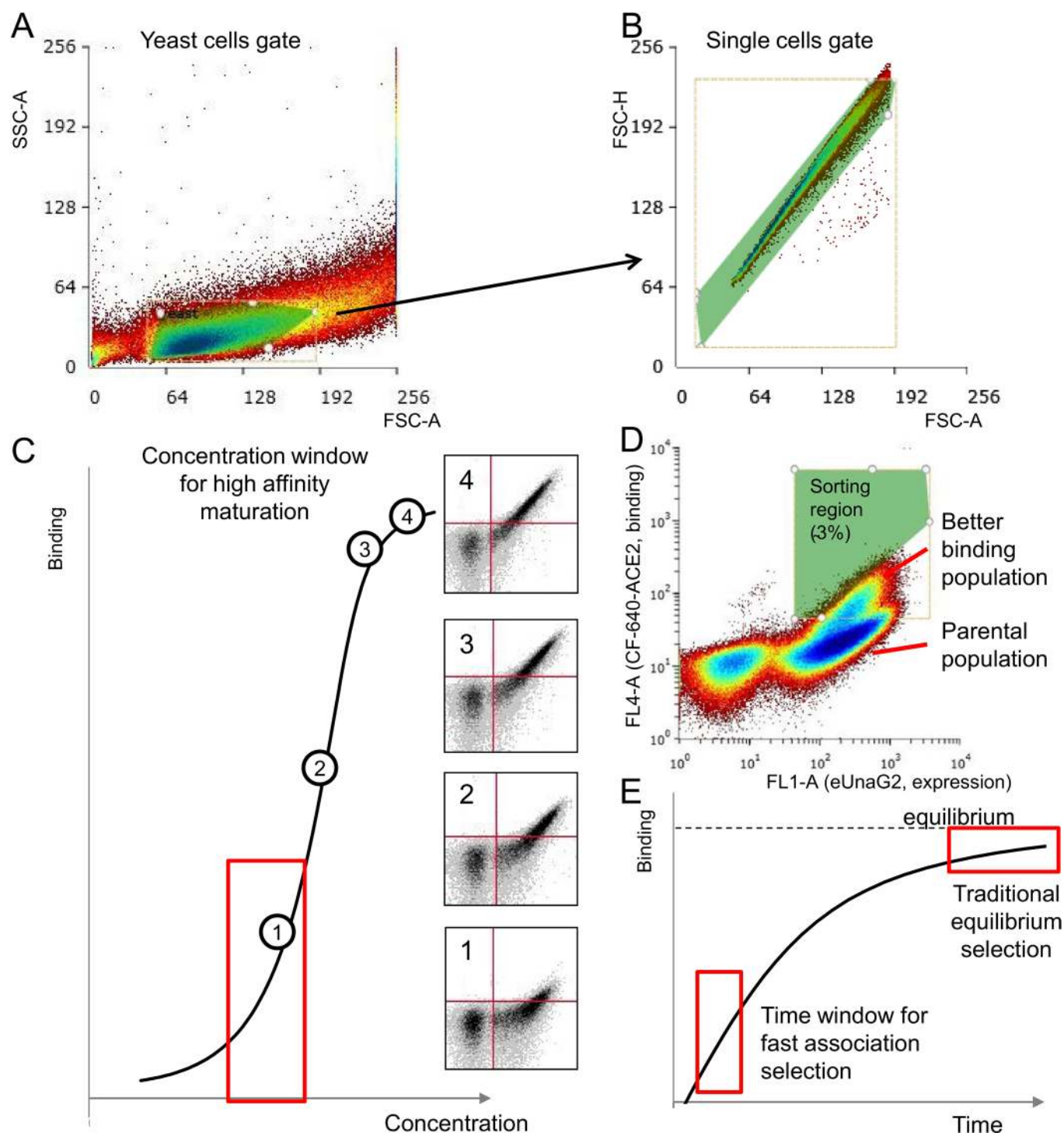
Extended Data Fig. 2 | Gaiting and selection strategies for *in vitro* evolution of the SARS-CoV-2 RBD domain. (**A**, **B**) Gating strategy for FACS sorting. In the first step, yeast cells were isolated by their FSC-A and SSC-A properties (A). In the second step (B), single cells were isolated by their FSC properties (area and height) on the diagonal plot. The Green area represents the gated region. (**C**) Selection strategy for affinity maturation. The library was titrated with a range of ACE2 concentrations to select the concentration with limited signal (inset 1). Under such conditions, the tighter binding clones gain the highest advantage over the parental population. Using less stringent selection (insets 2 – 4) reduces the advantage of the tighter binders. Using sub-optimal concentrations of ACE2 protein will also result in loss of selectivity. (**D**) Affinity maturation library after 3 sorts, where the separation between parental and tighter binding population is well defined. The top 0.1 – 0.3 % of cells were sorted – green region. (**E**) Fast-association selection strategy. The library was incubated with a constant concentration (30 pM) of ACE2 for different times. The time with minimal signal was determined and used for the selection of clones with faster association (library B6(FA)). The same shape of the sorting region as in (**D**) was applied.

Library	Clone	no. mut.	Plasmid	Clone specific mutations	$K_{D,YD}$ [pM]* [replicates]
	WT	0	pJYDC1	-	1700 ± 137 [8]
B3	RBD-31	3	pJYDC1	I358F , N448S, N501Y	907 ± 112 [3]
B3	RBD-32	3	pJYDC1	I358F, S477N, N501Y	348 ± 35 [3]
B3	RBD-33	3	pJYDC1	I358F, E484K, N501Y	204 ± 88 [4]
B3	RBD-34	4	pJYDC1	I358F, I468T, E484K, N501Y	644 [1]
B3	RBD-312	3	pJYDC1	I358F, V483E, N501Y	442 [1]
B3	RBD-36	4	pJYDC1	I358F, I468T, N481Y, N501Y	268 ± 35 [3]
B3	RBD-310	4	pJYDC1	I358F, N448S, E484K, N501Y	607 [1]
B3	RBD-316	5	pJYDC1	I358F, N448S, E484K, F490S, N501Y	722 [1]
B4	RBD-41	4	pJYDC3	I358F, S477N, E484K, N501Y	104 ± 28 [3]
B4	RBD-47	4	pJYDC3	I358F, N460K, E484K, N501Y	96 ± 14 [3]
B4	RBD-48	4	pJYDC3	I358F, E484K, Q498R, N501Y	55 ± 23 [4]
B5	RBD-51	6	pJYDC1	I358F, N460K, S477N, E484K, Q498R, N501Y	11 ± 4 [3]
B5	RBD-52	7	pJYDC1	I358F, N460K, E484K, S494P, Q498R, N501Y, A520G	40 ± 16 [5]
B5	RBD-53	6	pJYDC1	I358F, N460K, T478S, E484K, Q498R, N501Y	22 ± 4 [3]
B5	RBD-57	6	pJYDC1	I358F, G446R, E484K, F490Y, Q498R, N501Y	35 [1]
B5	RBD-520	5	pJYDC1	I358F, I468V, E484K, Q498R, N501Y	31 [1]
B5	RBD-521	5	pJYDC1	I358F, N460K, E484K, Q498R, N501Y	14 ± 3 [3]
B6(FA)	RBD-62	9	pJYDC3	I358F, V445K, N460K, I468T, T470M, S477N, E484K, Q498R, N501Y	3.5 ± 1.8 [4]
B6(FA)	RBD-610	7	pJYDC3	I358F, N460K, S477N, T478S, E484K, Q498R, N501Y	10 ± 5 [3]
B6(FA)	RBD-611	8	pJYDC3	I358F, N460K, S477N, E484K, S494P, Q498R, N501Y, S514T	26 ± 8 [3]
B6(FA)	RBD-614	7	pJYDC3	I358F, V445K, N460K, E484K, S494P, Q498R, N501Y	21 ± 6.2 [3]
RBD mutant clones not from library selection					
-	RBD-BRY	5	pJYDC1	I358F, K417T,E484K, Q498R, N501Y	84 ± 28 [5]
-	RBD-MBRY	6	pJYDC1	I358F, K417T, T470M, E484K, Q498R, N501Y	56 ± 8 [3]
-	RBD-477RY	4	pJYDC1	I358F, S477N, Q498R, N501Y	46 ± 6 [3]
-	RBD-478RY	4	pJYDC1	I358F, T478I, Q498R, N501Y	93 ± 8 [3]
-	RBD-SARY	5	pJYDC1	I358F, K417N, E484K, Q498R, N501Y	68 ± 22 [3]
-	RBD-477KRY	5	pJYDC1	I358F, S477R, E484K, Q498R, N501Y	30 ± 3 [3]
-	RBD-RYM	5	pJYDC1	I358F, T470M, E484K, Q498R, N501Y	39 ± 2 [3]
-	RBD-RY	3	pJYDC1	I358F, Q498R, N501Y	65 ± 13 [3]
-	RBD-71	12	pJYDC3	I358F, V367W, R408D, K417V, V445K, N460K, I468T, T470M, S477N, E484K, Q498R, N501Y	8.5 ± 2.5 [3]

Extended Data Fig. 3 | ACE2 binding affinity of yeast display selected clones and designed variants. * The affinity was determined by CF640-ACE2 titration (Fig. 2a) at 12-16 concentrations, binding to yeast expressing clones. Shading tones are according to library progression. The standard error values calculated from multiple measurements (number of repeats in brackets) are shown. Curve fitting errors did not exceed ± 10%.
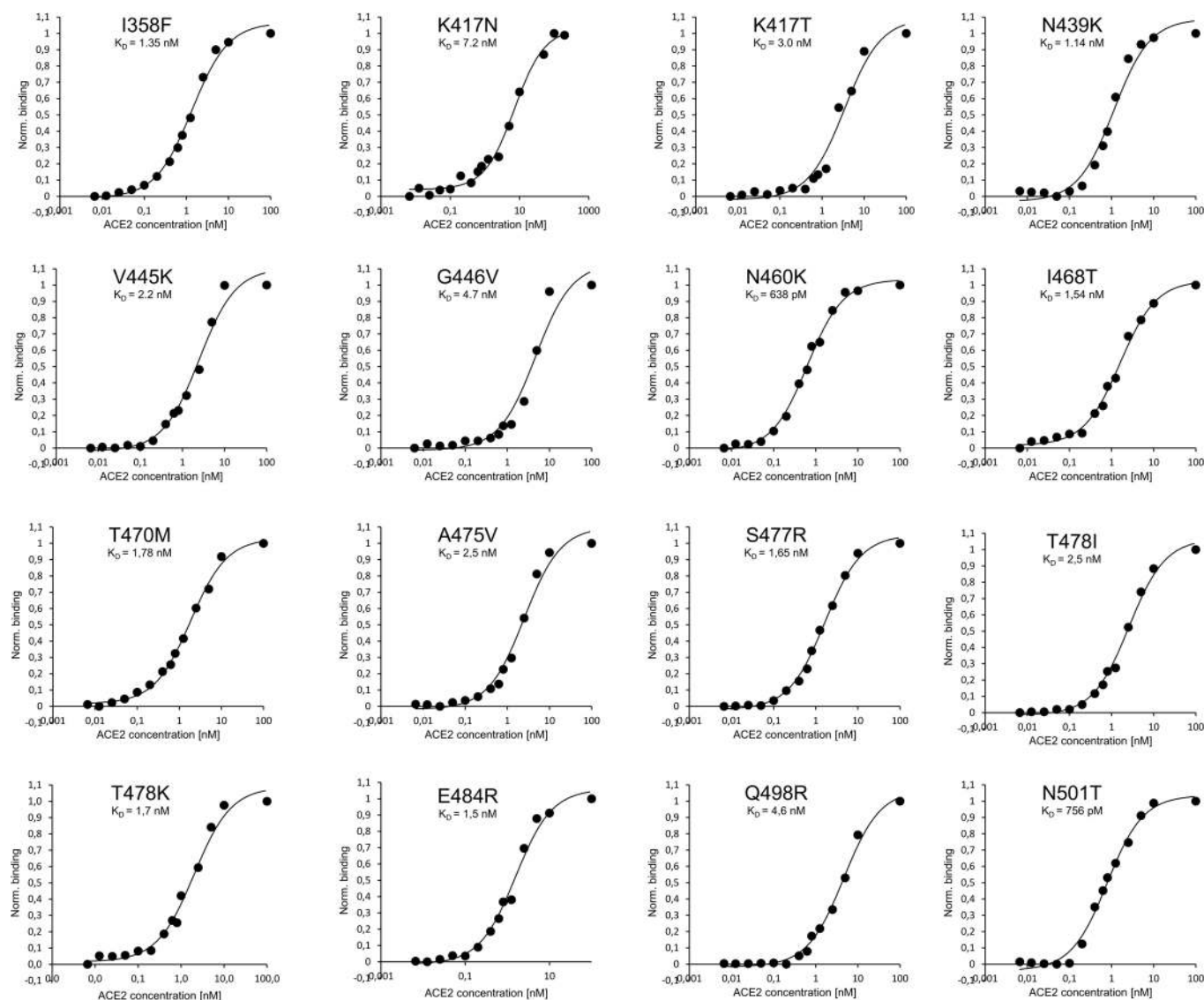
Extended Data Fig. 4 | Binding titration curves and binding affinities of isogenic variants. A single binding curve is shown here. The mean values and number of replicates are given in Fig. 2 and Source Data Fig. 2.
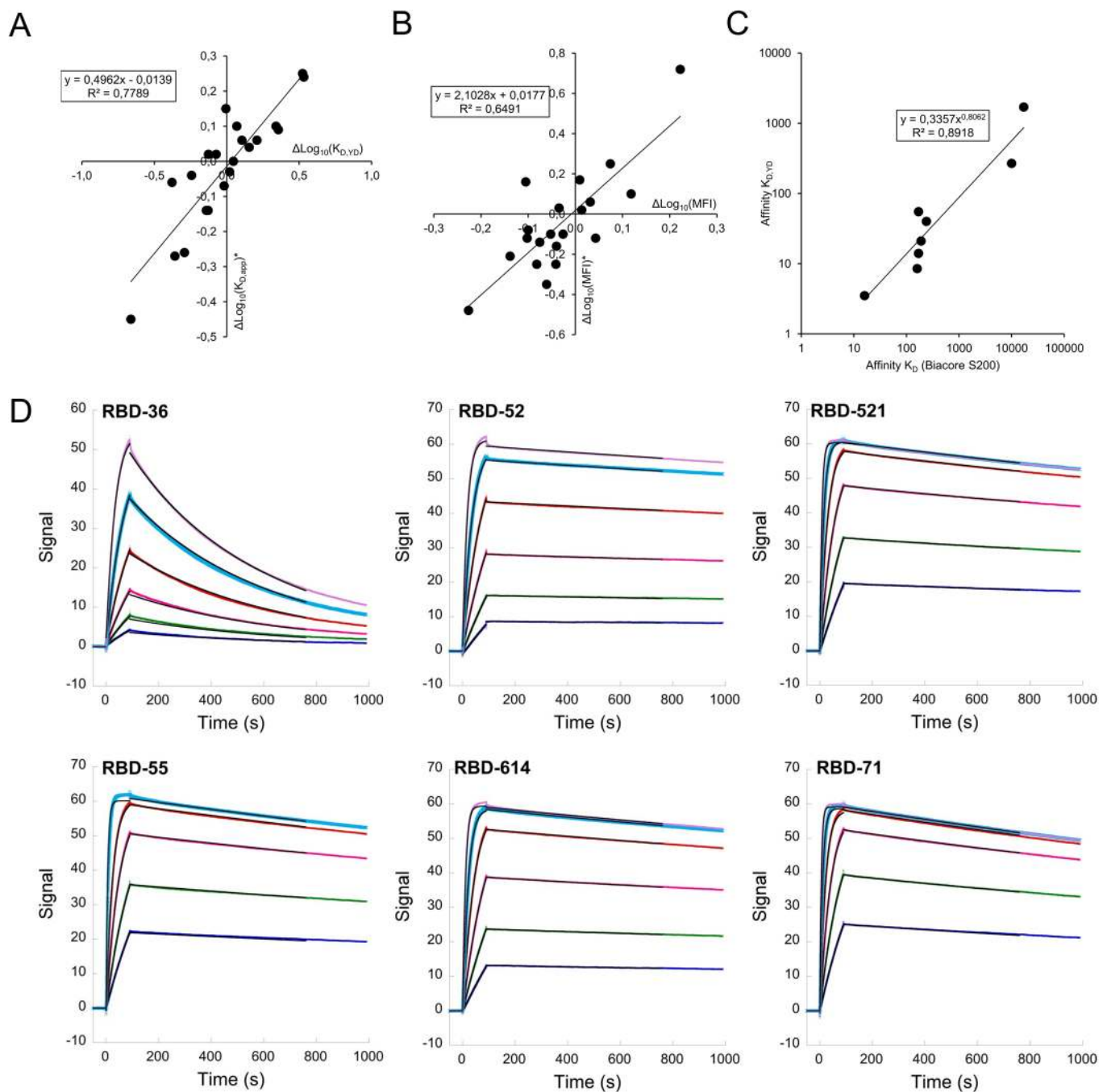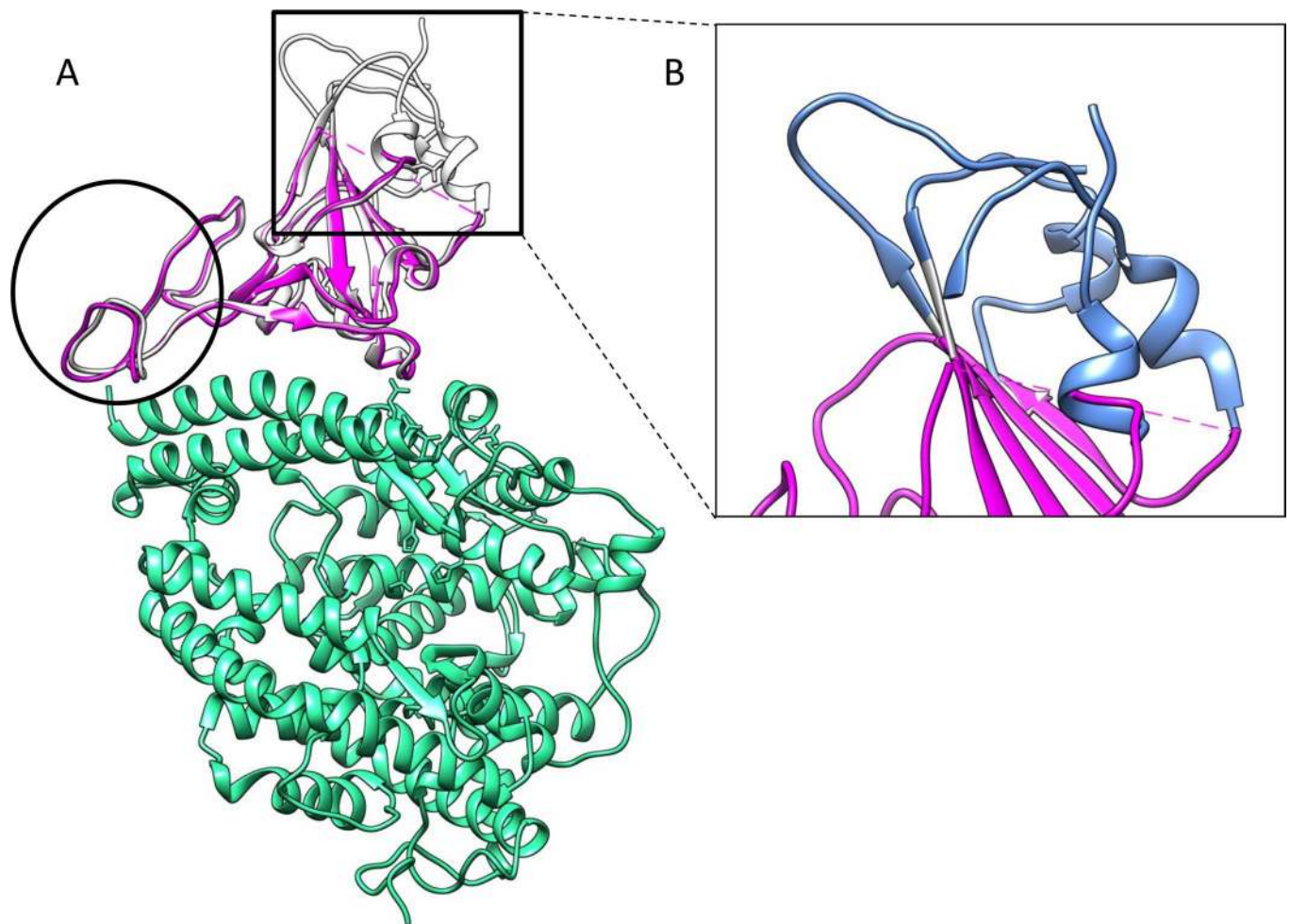
Extended Data Fig. 5 | Comparison between different methods for KD assessment. (**a**) Correlation between the herein measured mean $K_{D,YD}$ (Fig. 2 and Extended Data Fig. 3) and deep-mutational scanning data $K_{D,App}$[7]. (**b**) Correlation between mean fluorescent intensity MFI obtained here (Fig. 2) by using enhanced yeast display[21] and obtained from deep-mutational scanning ($\Delta Log10(MFI)^*$)[7]. (**c**) Correlation between mean $K_{D,YD}$ values obtained by yeast surface display titration (Fig. 2 and Extended Data Fig. 3) and SPR Biacore S200. (**d**) Biacore S200 sensorgrams for different RBD variants at 5, 10, 20, 40, 80, 160 nM concentration binding to immobilized ACE2. The black line is the global fit using the build-in function.

Position: 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528

```
>WT       D S K V G G N Y N Y L Y R L F R K S N L K P F E R D I S T E I Y Q A G S T P C N G V E G F N C Y F P L Q S Y G F Q P T N G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K
>RBD-31   D S K V G G S Y N Y L Y R L F R K S N L K P F E R D I S T E I Y Q A G S T P C N G V E G F N C Y F P L Q S Y G F Q P T Y G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K
>RBD-32   D S K V G G N Y N Y L Y R L F R K S N L K P F E R D I S T E I Y Q A G N T P C N G V E G F N C Y F P L Q S Y G F Q P T Y G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K
>RBD-33   D S K V G G N Y N Y L Y R L F R K S N L K P F E R D I S T E I Y Q A G S T P C N G V K G F N C Y F P L Q S Y G F Q P T Y G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K
>RBD-34   D S K V G G N Y N Y L Y R L F R K S N L K P F E R D T S T E I Y Q A G S T P C N G V K G F N C Y F P L Q S Y G F Q P T N G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K
>RBD-36   D S K V G G N Y N Y L Y R L F R K S N L K P F E R D I S T E I Y Q A G S T P C N G V E G F N C Y F P L Q S Y G F Q P T Y G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K
>RBD-39   D S K V G G N Y N Y L Y R L F R K S N L K P F E R D I S T E I Y Q A G S T P C N G V E G F N C Y F P L Q S Y G F Q P T Y G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K
>RBD-310  D S K V G G S Y N Y L Y R L F R K S N L K P F E R D I S T E I Y Q A G S T P C N G V K G F N C Y F P L Q S Y G F Q P T Y G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K
>RBD-312  D S K V G G N Y N Y L Y R L F R K S N L K P F E R D I S T E I Y Q A G S T P C N G E E G F N C Y F P L Q S Y G F Q P T Y G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K
>RBD-316  D S K V G G S Y N Y L Y R L F R K S N L K P F E R D I S T E I Y Q A G S T P C N G V K G F N C Y S P L Q S Y G F Q P T Y G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K
>RBD-317  D S K V G G N Y N Y L Y R L F R K S N L K P F E R D I S T E I Y Q A G S T P C N G V K G F N C Y F P L Q S Y G F Q P T N G V G Y Q P Y R V V V L S F E I L H A P A T V C G P K
>RBD-318  D S K V G G N Y N Y L Y R L F R K S N L K P F E R D I S T E I Y Q A G S T P C N G V K G F N C Y F P L Q S Y G F Q P T N G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K
>RBD-323  D S K V G G N Y N Y L Y R L F R K S N L K P F E R D I S T E I Y Q A G S T P C N G V K G F N C Y F P L Q S Y G F Q P T N G V G Y Q P Y R V V V L S F E L L H A P A T V C A P K
>RBD-41   D S K V G G N Y N Y L Y R L F R K S N L K P F E R D I S T E I Y Q A G N T P C N G V K G F N C Y F P L Q S Y G F Q P T Y G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K
>RBD-42   D S K V G G S Y N Y L Y R L F R K S K L K P F E R D I S T E I Y Q A G S T P C N G V K G F N C Y S P L Q S Y G F Q P T Y G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K
>RBD-43   D S K V G G N Y N Y L Y R L F R K S N L K P F E R D T S T E I Y Q A G S T P C Y G V E G F N C Y F P L Q S Y G F Q P T Y G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K
>RBD-47   D S K V G G N Y N Y L Y R L F R K S K L K P F E R D I S T E I Y Q A G S T P C N G V K G F N C Y F P L Q S Y G F Q P T Y G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K
>RBD-48   D S K V G G N Y N Y L Y R L F R K S N L K P F E R D I S T E I Y Q A G S T P C N G V K G F N C Y F P L Q S Y G F R P T Y G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K
>RBD-49   D S K V G G N Y N Y L Y R L F R K S N L K P F E R D I S T E I Y Q A G N T P C N G V K G F N C Y F P L Q S Y G F R P T Y G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K
>RBD-410  D S K V G G N Y N Y L Y R L F R K S N L K P F E R D I S T E I Y Q A G S T P C N G V K G F N C Y F P L Q S Y G F R P T Y G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K
>RBD-411  D S K V G G N Y N Y L Y R L F R K S K L K P F E R D V S T E I Y Q A G S T P C N G V K G F N C Y F P L Q S Y G F Q P T Y G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K
>RBD-414  D S K V G G N Y N Y L Y R L F R K S N L K P F E R D I S T E I Y Q A G S T P C N G V K G F N C Y F P L Q S Y G F Q P T Y G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K
>RBD-418  D S K V G G N Y N Y L Y R L F R K S N L K P F E R D I S T E I Y Q A G S T P C N G V K G F N C Y F P L Q P Y G F Q P T Y G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K
>RBD-419  D S K V G G N Y N Y L Y R L F R K S N L K P F E R D I S T E I Y Q A G S S P C N G V K G F N C Y F P L Q S Y G F R P T Y G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K
>RBD-420  D S K V G G N Y N Y L Y R L F R K S N L K P F E R D I S T E I Y Q A G N T P C N G V K G F N C Y F P L Q S Y G F R P T Y G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K
>RBD-52   D S K V G G N Y N Y L Y R L F R K S K L K P F E R D I S T E I Y Q A G S T P C N G V K G F N C Y F P L Q P Y G F R P T Y G V G Y Q P Y R V V V L S F E L L H G P A T V C G P K
>RBD-53   D S K V G G N Y N Y L Y R L F R K S K L K P F E R D I S T E I Y Q A G S S P C N G V K G F N C Y F P L Q S Y G F R P T Y G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K
>RBD-57   D S K V R G N Y N Y L Y R L F R K S K L K P F E R D I S T E I Y Q A G S T P C N G V K G F N C Y P L Q S Y G F R P T Y G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K
>RBD-520  D S K V G G N Y N Y L Y R L F R K S N L K P F E R D V S T E I Y Q A G S T P C N G V K G F N C Y F P L Q S Y G F R P T Y G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K
>RBD-521  D S K V G G N Y N Y L Y R L F R K S K L K P F E R D I S T E I Y Q A G S T P C N G V K G F N C Y F P L Q S Y G F R P T Y G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K
>RBD-522  D S K V G G N Y N Y L Y R L F R K S K L K P F E R D I S T E I Y Q A G S T P C N G V K G F N C Y F P L Q S Y G F R P T Y G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K
>RBD-534  D S K V G G N Y N Y L Y R L F R K S K L K P F E R D I S T G I Y Q A G S T P C N G V K G F N C Y F P L Q S Y G F R P T Y G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K
>RBD-547  D S K V G G N Y N Y L Y R L F R K S K L K P F E R D I S T E I Y Q A G S T P C N G V K G F N C Y F P L H S Y G F R P T Y G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K
>RBD-62   D S K K G G N Y N Y L Y R L F R K S K L K P F E R D T S T M I Y Q A G N T P C N G V K G F N C Y F P L Q S Y G F R P T Y G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K
>RBD-610  D S K V G G N Y N Y L Y R L F R K S K L K P F E R D I S T E I Y Q A G N S P C N G V K G F N C Y F P L Q S Y G F R P T Y G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K
>RBD-611  D S K V G G N Y N Y L Y R L F R K S K L K P F E R D I S T E I Y Q A G N T P C N G V K G F N C Y F P L Q P Y G F R P T Y G V G Y Q P Y R V V V L T F E L L H A P A T V C G P K
>RBD-614  D S K K G G N Y N Y L Y R L F R K S K L K P F E R D I S T E I Y Q A G N T P C N G V K G F N C Y F P L Q P Y G F R P T Y G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K
```

Extended Data Fig. 6 | Amino-acid sequences of non-redundant clones isolated during the RBD affinity-maturation process. Rare/transient mutations are in light blue. Important changes are highlighted in dark blue. Except the stabilizing mutation I358F, no additional mutations were identified outside the mutagenized region AA442 – 528.
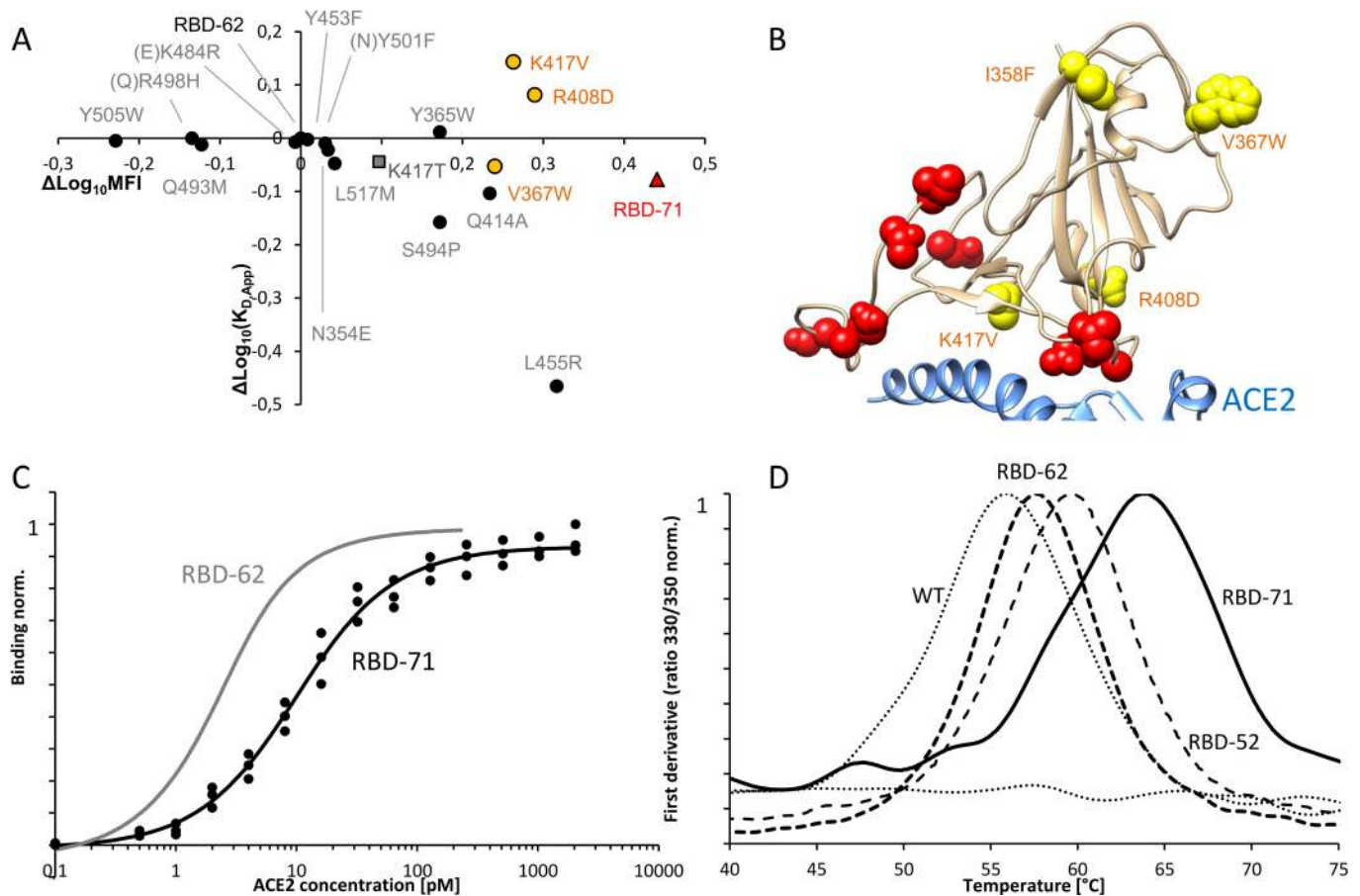
Extended Data Fig. 7 | Site−directed mutagenesis of RBD-62, using the affinity-enhancing mutations. Fifteen mutations were predicted to enhance RBD-ACE2 binding or stability[6]. Mutation K417T (grey) was introduced from the B6(FA) library. These mutations were evaluated for enhancing the affinity of RBD-62 towards ACE2 and their effect on stability. Single measurements only. (**A**) Impact of mutations, on top of RBD-62, on ACE2 binding (y-axis) and yeast surface expression. Three mutations (orange circles), which had the highest impact on expression, were combined in RBD-71 (red triangle). (**B**) Localization of stabilizing (yellow) and binding-enhancing mutations depicted in the RBD structure (PDB ID 6m17, best rotamer is shown). (**C**) Binding curve of RBD-71 with RBD-62 binding curve fit (in grey) for comparison. All data points, taken in three independent repeats, are shown. (**D**) Normalized protein melting curves for RBD-WT, RBD-62, RBD-52, and RBD-71 measured using the Tycho NT.6 (NanoTemper).

Extended Data Fig. 8 | Global comparison between RBD-WT and RBD-62 shows overall similarity. (**A**) RBD-62 preserves its typical twisted five‑stranded antiparallel β sheet (β1, β3-β5, and β10) with an extended insertion containing the short β5-β9 strands, α4, and η3 helices and loops. The biggest differences are pronounced between M470 and F490 (black circle). (**B**) The upper part comprising of three segments: R357–S371 (β2, α2), G381-V395 (α3), and F515-H534 (β11**)** is not resolved in the density map (blue ribbon, added from PDB ID: 6M0J).
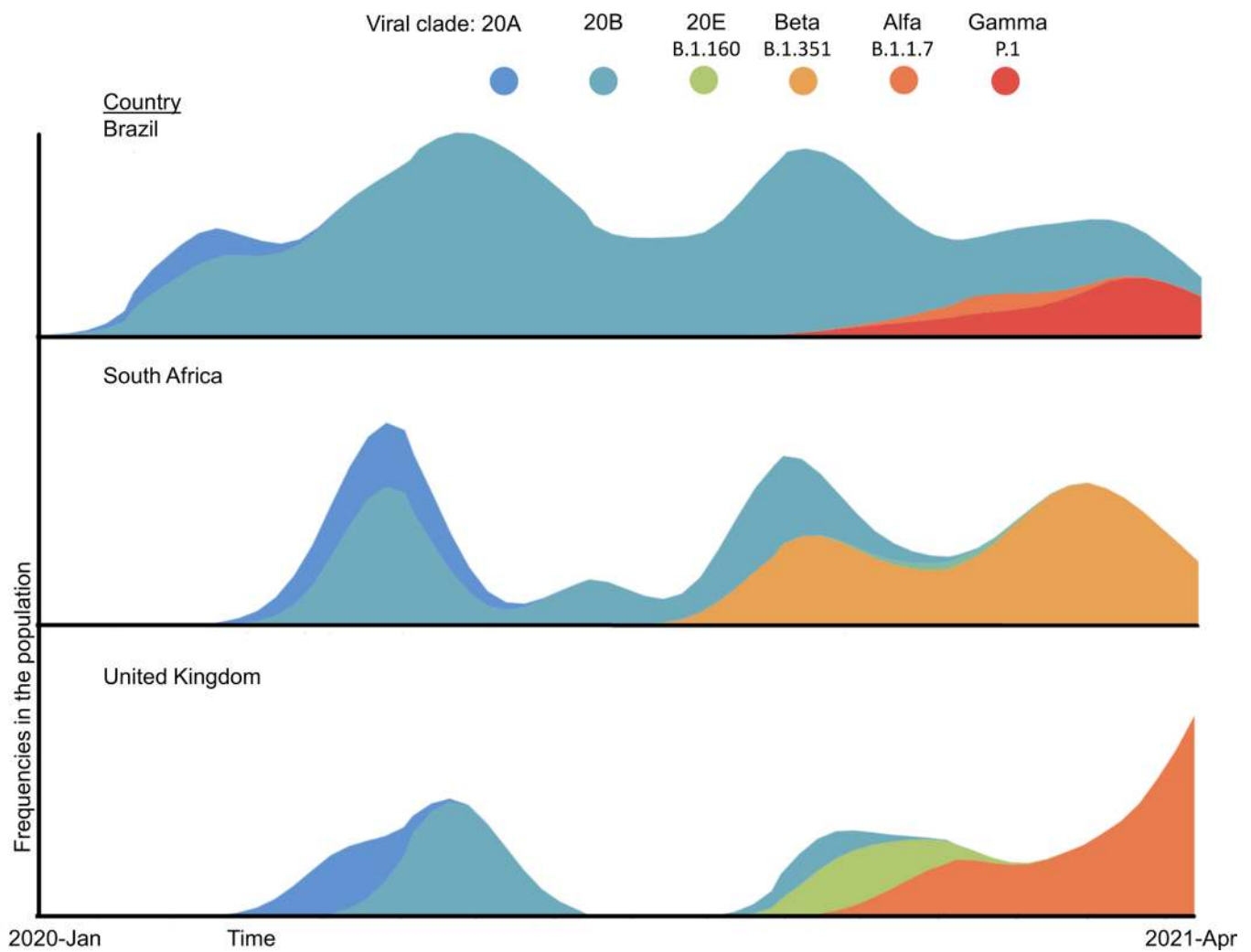
Extended Data Fig. 9 | Emergence and frequencies over time of new SARS-CoV2 virus variants in different countries. Data were reproduced from https://nextstrain.org/ (16.4.2021)[49]. The lineage designation by Rambaut et al.[50] is given below the strain designation proposed by NextStrain or WHO.
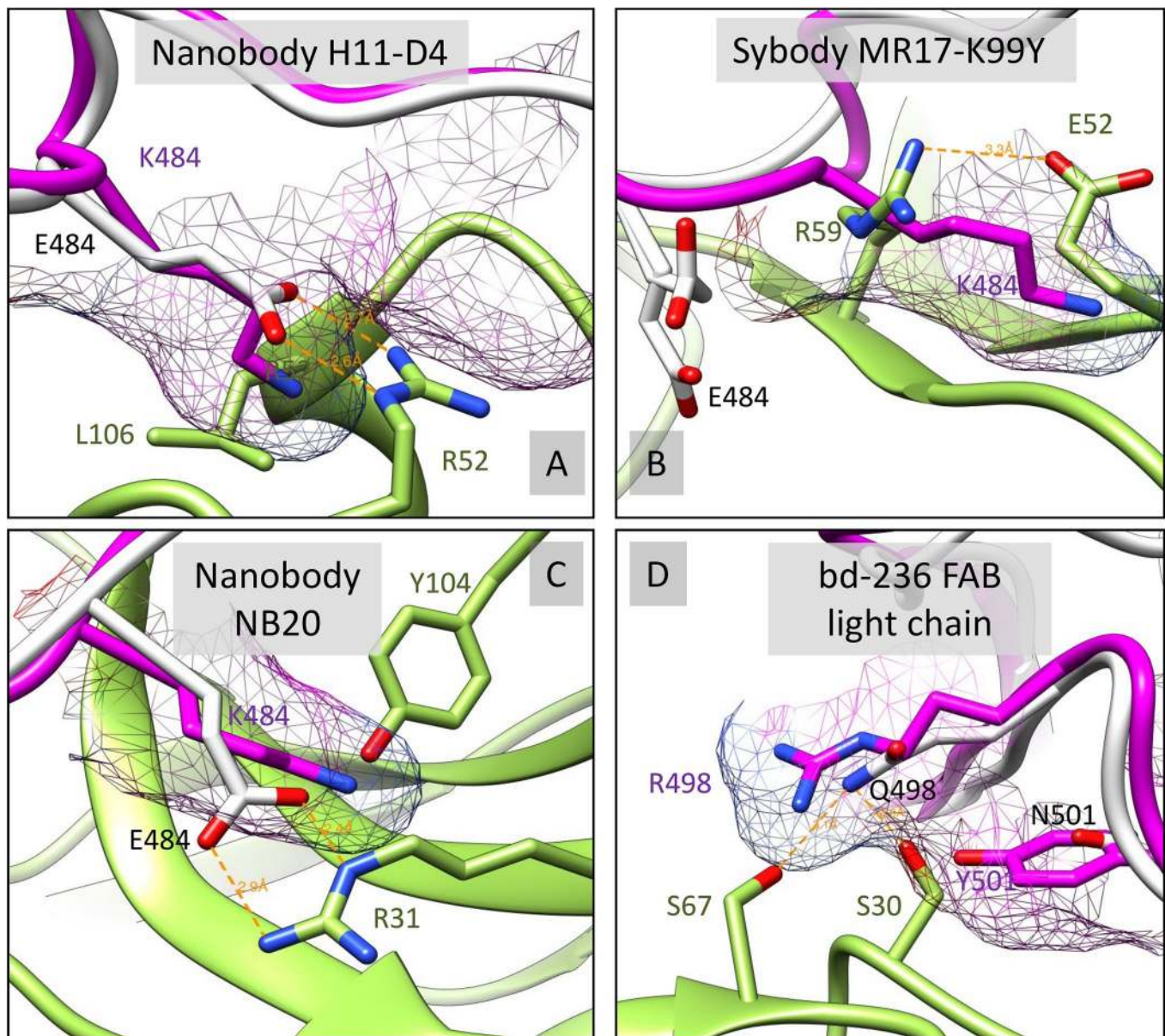
Extended Data Fig. 10 | RBD-62 mutations interfere with binding to multiple antibodies. The RBD-62 (magenta, mesh surface) was structurally overlaid with RBD-WT (white). S477N, E484K, Q498R, and N501Y RBD mutated residues were analyzed for disruption of wild-type contacts (orange dashed line) and clashes with corresponding binding antibody/nanobody (green) in relation to RBD-WT. Four examples, A) PDB ID: 6YZ5, B) PDB ID: 7CAN, C) PDB ID: 7JVB, D) PDB ID: 7CHE, where the mutations resulted in a dramatic impact, are shown.

nature research

Corresponding author(s): *Double-blind peer review submissions: write DBPR and your manuscript number here instead of author names.*

Last updated by author(s): *YYYY-MM-DD*

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 | *Provide a description of all commercial, open source and custom code used to collect the data in this study, specifying the version used OR state that no software was used.*

Data analysis | Yeast display titration curve fitting were done by the standard non-cooperative Hill equation, fitted by nonlinear least-squares regression with two additional parameters discribing the curve by using Python 3.7 similarly common methods eg doi.: 10.1016/j.cell.2020.08.012; 10.1038/nprot.2006.94;

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

The RBD-62 Cryo-EM data are available at EMD-12187 and PDB 7BH9.

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	*Describe how sample size was determined, detailing any statistical methods used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.*
Data exclusions	*Describe any data exclusions. If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.*
Replication	The number of replicates is clearly identified when relevant. There was a high level of reproducibility among yeast display expression levels and binding affinities.
Randomization	*Describe how samples/organisms/participants were allocated into experimental groups. If allocation was not random, describe how covariates were controlled OR if this is not relevant to your study, explain why.*
Blinding	*Describe whether the investigators were blinded to group allocation during data collection and/or analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.*

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

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	HEK293T cells (Weizmann Institute of Science cell line repository), HEK-293T-ACE2 (GenScript M00770); VeroE6 (C1008, Institute Pasteur, France), Expi293F cells (ThermoFisher), S.cerevisiae EBY100 (Invitrogen)
Authentication	Provided by the supplier.
Mycoplasma contamination	All used cell lines were negatively tested for mycoplasma contamination by Hy-mycoplasma detection PCR kit (HyLabs, Ky 50341)
Commonly misidentified lines (See ICLAC register)	*Name any commonly misidentified cell lines used in the study and provide a rationale for their use.*

Animals and other organisms

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

Laboratory animals	Golden Syrian hamsters (Mesocricetus auratus) were housed at BIOQUAL, Inc. Hamsters were 6-8 weeks old on day of infection and were randomized into groups of 3 female and 3 male hamsters each per group.

Wild animals	*Provide details on animals observed in or captured in the field; report species, sex and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.*
Field-collected samples	*For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.*
Ethics oversight	Animal studies were reviewed and approved by the Animal Care and Use Committee of the Vaccine Research Center, National Institute of Allergy and Infectious Disease, National Institutes of Health, Bethesda USA

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	Sample preparation is described in detail in the Method section. S.cerevisiae EBY100 cells expressing the RBD variant/s were washed by an ice-cold PBSB buffer (PBS with 1 g/L BSA) and incubated with various concentrations of CF-640 labeled ACE2 protein. The sample volume was adjusted (1 – 100 ml) to avoid the ligand depletion as well as an equilibrium incubation time (1 – 12 h). The cells were subsequently washed (PBSF), passed through the cell strainer, and analyzed by cytometry. The HEK293T cells were washed twice with PBS (Biological Industries, REF 02-023-1A) and directly analyzed on BD Accuri™ C6 Plus Flow Cytometer, BD Biosciences, USA. The gating strategy is shown in Figure S2.
Instrument	The sorting experiments were performed on S3e Cell Sorter (BioRad, USA, serial no 777BR2079). The binding data were acquired on S3e Cell Sorter (BioRad, USA) and BD Accuri™ C6 Flow Cytometer (BD Biosciences, USA).
Software	Analyses were done by ProSort™ Software (v. 1.6) or by CFlow Plus Software (v. 1.0.227.4)
Cell population abundance	*Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.*
Gating strategy	Gating strategies are described in detail in the Supplementary material Figure S2.

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