Supplementary Material

Towards an optimal monoclonal antibody with higher binding affinity to the receptor-binding domain of SARS-CoV-2 spike proteins from different variants

Andrei Neamtu[a,i], *Francesca Mocci*[c], *Aatto Laaksonen*[b,c,d,e,f], *and Fernando L. Barroso da Silva*[g,h*]

[a] Department of Physiology, "Grigore T. Popa" University of Medicine and Pharmacy of Iasi, Str. Universitatii nr. 16, 700051 Iasi, România

[b] Centre of Advanced Research in Bionanoconjugates and Biopolymers, PetruPoni Institute of Macromolecular Chemistry AleeaGrigoreGhica-Voda, 41A, 700487 Iasi, Romania.

[c] University of Cagliari, Department of Chemical and Geological Sciences, Campus Monserrato, SS 554 bivio per Sestu, 09042 Monserrato, Italy

[d] Department of Materials and Environmental Chemistry, Arrhenius Laboratory, Stockholm University, SE-106 91 Stockholm, Sweden.

[e] State Key Laboratory of Materials-Oriented and Chemical Engineering, Nanjing Tech University, Nanjing, 210009, P. R. China.

[f] Department of Engineering Sciences and Mathematics, Division of Energy Science, Luleå University of Technology, SE-97187 Luleå, Sweden.

[g] Universidade de São Paulo, Departamento de Ciências Biomoleculares, Faculdade de Ciências Farmacêuticas de Ribeirão Preto, Av. café, s/no – campus da USP, BR-14040-903 – Ribeirão Preto – SP, Brazil.

[h] Department of Chemical and Biomolecular Engineering, North Carolina State University, Raleigh, North Carolina 27695, United States

[i] TRANSCEND Centre – Regional Institute of Oncology (IRO) Iasi, Str. General Henri Mathias Berthelot, Nr. 2-4 Iași, România

*Corresponding author (flbarroso@usp.br and/or flbarros@ncsu.edu)

List of acronymous

ACE2 – angiotensin-converting enzyme 2
Ag – antigen
Ab-Ag – antibody-antigen
CG – coarse-grain
CDRs – complementarity-determining regions
CpH – constant-pH
CPT – convalescent plasma therapy
FPTS – Fast proton titration scheme
FDA – Food and Drug Administration
FF – force field
FORTE – Fast cOarse-grained pRotein-proTein modEl
L – light chain
H – heavy chain
MD – Molecular Dynamics
mAb – monoclonal antibody
MC – Monte Carlo
PMF – Potentials of Mean Force
PDF – probability distribution functions
RBD – receptor-binding domain
RAbD – RosettaAntibodyDesign
REU – ROSETTA Energy Units
RMSD – Root-mean-square deviation
SARS-CoV-2 – severe acute respiratory syndrome coronavirus 2
SIRAH – Southamerican Initiative for a Rapid and Accurate Hamiltonian
S –spike
TAS – theoretical alanine scanning
US – Umbrella sampling
VOC – present and past variants of concern
VOI – present and past variants of interest
wt – wild-type

S1) Theoretical methods – Supplementary information

Following a biophysical approach, classical MD and MC methods are techniques that have a long history of successful applications in many scientific problems as standalone theoretical studies or in combination with experimental techniques [1–8]. Other bioinformatics algorithms are completing theoretical resources and helping to address challenging problems in the biomolecular world [9–12].

Some previous theoretical studies with CR3022

As mentioned in the introduction, CR3022 has been suggested as a promising therapeutic option to neutralize SARS-CoV-2 [13–16], and is often explored in theoretical studies [17–20]. In a pioneer computational study at the very beginning of the pandemic[17], using constant-pH MC simulations, Giron, Laaksonen, and Barroso da Silva showed that CR3022 - known to bind to SARS-CoV-1 RBD - could also bind to SARS-CoV-2 RBD. They also mapped the epitopes and identified the importance of electrostatic interactions for the corresponding Ab-Ag interface [17]. Following different routes, Ding et al. [18] have proposed an efficient and reliable computational screening method based on "Molecular Mechanics Poisson-Boltzmann surface area" (MM/PBSA) to estimate binding free energy between SARS-CoV-2 RBD and ACE2 together with CR3022 and CB6 in good agreement with reported experimental values. Their scheme identifies the key residues that increase hydrophobicity and indicates that changing the sign of charged residues from positive to negative can increase the binding affinity. In comparison with standard MM/PBSA, their method is more accurate due to the introduction of electrostatic energy in the scheme.

Lagoumintzis et al. [21] used *in silico* methods in their studies of the recent hypothesis that SARS-CoV-2 would interact with nicotinic acetylcholine receptors (nAChRs) and disrupt the regulation of the nicotinic cholinergic system (NCS) and the cholinergic anti-inflammatory pathway. They used ROSETTA and multi-template homology modeling to study a sequence from a snake venom toxin to predict the structure of the extracellular domains of nAChRs ("toxin binding site"). Using the "High Ambiguity Driven protein-protein DOCKing" (HADDOCK) approach [22], they found a protective role of nicotine and other cholinergic agonists and observed that CR3022 and other similar mAbs show an increased affinity for SARS-CoV-2 Spike glycoprotein. To study the molecular mechanisms in SARS-CoV-2 S protein binding with several mAbs, Verkhivker and Di Paola [23] performed all-atom and CG simulations with mutational sensitivity mapping, using the BeAtMuSiC approach and perturbation response scanning profiling of SARS-CoV-2 receptor-binding domain complexed with CR3022 and CB6 antibodies, complementing it with a network modeling analysis of the residue interactions. Their results provide insight into allosteric regulatory mechanisms of SARS-CoV-2 S proteins, where the mAbs are modulating the signal communication. This provides a strategy to target specific regions of allosteric interactions therapeutically. Recently, Riahi et al. [24] presented a combined physics-based and machine learned-based computational mAb engineering platform to improve the binding affinity to SARS-CoV-2. They minimized (protonated, if needed) the Protein Data Bank (PDB) structures [25] using the "Molecular

Operating Environment" (MOE) program [26], and continued with ROSETTA [27,28] and FastRelax ROSETTA [29] to later apply machine learning. MOE was used for residue scanning in combination with two machine learning models: TopNETTree (for local geometry of protein complexes) and SAAMBE3D (for a variety of chemical, physical as well as sequential, and mutation properties). CR3022 and two other mAbs (m396 and 80R) were used as templates for their diversified epitopes, complexed with SARS-CoV-2 RBD. Their results suggest combining these three mAbs for higher neutralization activity. Nguyen et al. [19] confirmed that electrostatic interactions explain the higher binding affinity of CR3022 for SARS-CoV-2 RBD_{wt} than the 4A8 mAb in their all-atom and CG MD simulations (including steered MD). They used the Jarzynski equality to estimate the non-equilibrium binding free energy. They analyzed H-bonds and non-bonded contacts and used Debye-Hückel theory to model the electrostatic interactions for both RBD and N-terminal domain binding sites containing charged residues. Their results indicate that effective mAb candidates should contain many charged amino acids in the regions binding to spike protein. As the important residues of the spike protein involved in the binding are positively charged (Lys and Arg), the mAbs should correspondingly contain negatively charged residues (Asp and Glu) as anticipated before [17]. Martí et al. [30] have applied classical MD simulations and accelerated MD (aMD) for enhanced sampling. They have included the complexes between the RBD of SARS-CoV-2 spike (S) glycoprotein and CR3022 or S309 antibodies and the ACE2 receptor. Using MD simulations, they calculated the potential of mean force to obtain the free energy profiles for the complexes with the RBD. With their protocol, they could explore a large part of the conformational space accessible to RBD-ACE2/CR3022/S309 complexes. They found the affinity in protein-protein complexes to follow the decreasing order: S/CR3022 > S/309 > S/ACE2. Shariatifar and Farasat [20] have also performed MD simulations for SARS-CoV-2 RBD complexed with CR3022 and its modifications and calculated the free binding energies. They used the FastContact software to select mutations favorable for the wild type to produce two variants of CR3022 based on their amino acid binding conformations, showing a clear affinity enhancement compared to the wild type.

Structural-bioinformatics-based RAbD approach

The RAbD protocol consists of alternating outer and inner Monte Carlo design cycles. Each outer cycle consists of randomly choosing CDRs (L1, L2, L3, H1, H2, H3) from clusters in the RAbD database and then grafting that CDR's structure onto the antibody framework in place of the existing CDRs (GraftDesign). The program then performs *N* rounds of the inner cycle, consisting of sequence design (SeqDesign) followed by energy minimization. Each inner cycle introduces mutations and structurally optimizes the backbone and repacks side chains of the CDR chosen in the outer cycle to optimize interactions of the CDR with the antigen and other CDRs. Through all the steps the Ab-Ag complex is modeled at the atomistic level.

Three scenarios were used for antibody design: (A) all CDRs were modeled; (B) L1 was preserved from the original CR3022, the rest were subjected to full design. L1 in the original fragment of CR3022 mAb (for the sake of simplicity, we shall refer to the fragment of CR3022 simply as CR3022 from now on) is an extended loop that makes a large surface area contact

with the RBD_{wt} antigen (wildtype/Wuhan sequence) which stabilizes the Ab-Ag interaction. So, L1 was subjected only to SeqDesign, not to GraftDesign as the latter one also modifies the length of the CDR; (C) L1 and H3 were taken from CR3022, the rest being subjected to full design. L1 was modeled as in step (B). Among the 6 CDRs, H3 - located in between the H and L chains - is the only one not canonical, i.e., it does not adopt classifiable conformations (clusters of conformations). When visually analyzing the RAbD-generated conformations, it became obvious that some candidates with both very high scores and Ab-Ag interface surface areas were not realistic **(Figure S1).** For example, large Ab-Ag interface surface areas encountered in some cases were due to very long H3 CDRs. However, care should be taken when predicting H3 CDR, especially for long ones. Thus, H3 was not subjected to GraftDesign or SeqDesign in this scenario.

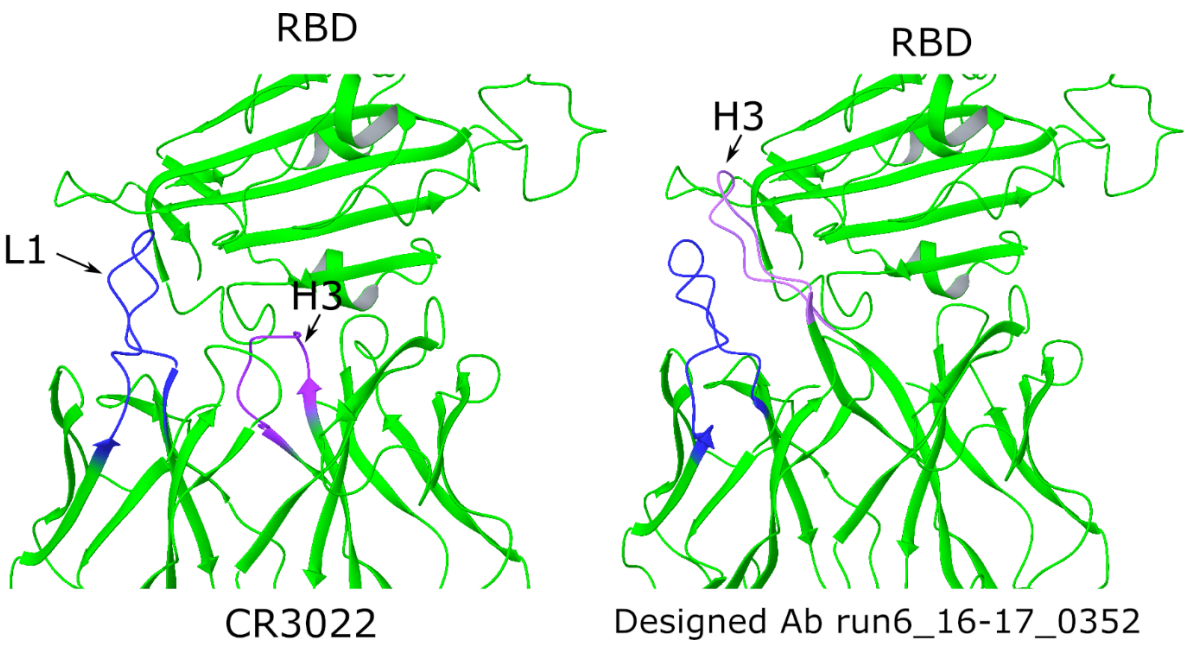

Figure S1. Example of a candidate mAb considered a false positive. Native CR3022/RBD complex (left) and the RAbD predicted pose, code 16-17_0352, of mAb-RBD complex (right) have been represented with the L1 and H3 CDRs highlighted. It can be seen how the designed long H3 (mauve) CDR actually prevents native L1 (blue) CDR from interacting with the antigen. The contact area between the 16-17_0352 mAb and RBD is higher (2339 $Å^2$) than in the native complex (2060 $Å^2$). However, H3 modeling should be regarded with care, especially for long H3 CDRs, for which the predictions have low reliability.

A total number of 235,800 complexes were designed using the RAbD approach. Choosing the best candidates for the subsequent evaluations from this very large number of antibody-antigen complexes is to a large extent arbitrary. However, the selection was based on two criteria that were chosen to include an interface interaction energy and contact area better or at least comparable with the ones of the native CR3022 - RBD complex: I) the candidate must have an

Ab-Ag interface score below –150 REU (ROSETTA Energy Units) (the native interface score is -65 REU), and II) an Ab-Ag interface surface area larger than 1900 $Å^2$ (the native complex interface surface area is 2060 $Å^2$).

CG molecular dynamics (MD) with an enhanced sampling approach for free energy calculations

During the US procedure, the geometry of the mAb was restrained by applying weak harmonic position restraints (20 kJ $mol^{-1}nm^{-2}$ force constant) on the CG beads corresponding to the protein backbone. The distance between the center of mass (COM) of mAb and the RBD was considered as the reaction coordinate. Cylindrical positional restraints (as defined in the GROMACS 2019 suite [31,32]). The simulation time was enough for convergence (see Figure S3).

A fast constant-pH coarse-grained (CG) simulation approach for free energy calculations on a large scale

This approach is aimed to capture only the main features of the complexation phenomena with a clear emphasis on the electrostatic interactions. Different studies have highlighted the importance of these interactions for the host-pathogen and antigen-specific antibody interfaces [17,19,33–39]. Successful applications of this simplified model to several biomolecular systems have been previously reported in the literature including viruses proteins and the SARS-CoV-2 RBD-mAb interactions [17,40–45]. In the FORTE model, charged and neutral spherical beads of different radii mimicking titratable and non-titratable amino acids, respectively, interact via Coulombic and van der Waals terms [40,41]. Protein coordinates given by the structural-bioinformatics-based approach were directly converted into this amino acid model. Atoms not belonging to the protein and the solvent were removed from the input structures.

Early ideas of such a model are rooted in the works of Marcus [46] and Jönsson and co-authors [47–49]. The reduction in the degrees of freedom together with the description of the proteins at the mesoscopic level in a continuum solvent model is a clear advantage from a computational point of view. For instance, the smaller amino acid (glycine) is modeled by four sites in the SIRAH ff while a single bead is used in FORTE. The reduction is even more significant for larger residues such as aspartic acid where the drop is from 11 to 1. This can result in a decrease in the computing time by a factor of 11^2 making it possible to apply it in a large-scale scenario. On the other hand, the main drawbacks are the assumption of a rigid body description to model the macromolecules and the ambiguity involved in the choices of the van der Waals contributions [17,40]. Yet, to form the RBD-mAb complex, the two molecules have to come close before conformational changes happening at short-range separation distances can be important to add additional attraction and/or stabilize the formed complex [50]. These limitations have not been critical in previously studied cases. The present outcome also contributes to this direction as discussed in the results section below. More details of this electrostatic protein-protein model are given elsewhere [17,40,41].

The relatively lower computation costs in comparison with other theoretical approaches allow the repetition of the calculations at different physical-chemical conditions (e.g., different solution pHs) and macromolecular systems (e.g., RBDs with several different mutations) using high-performance computers. This is a key aspect to investigate the binding affinities for a set of several RBDs (with all possible mutations of interest), and different binder candidates and perform further optimization of the best ones. Many runs are needed which are prohibitive with more elaborated molecular models.

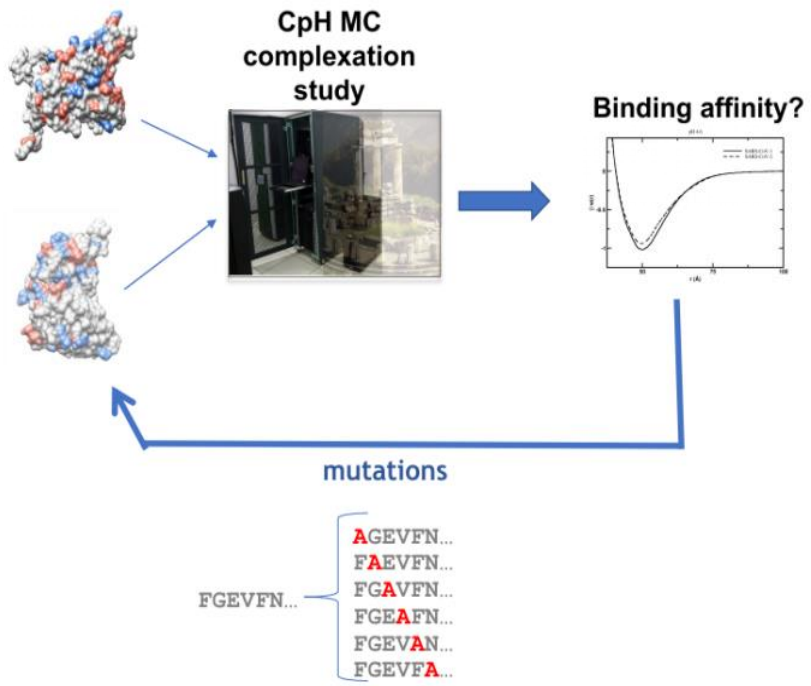

Figure S2. Scheme for the procedure used for the theoretical electrostatic mAb optimization using TAS. A pair RBD-mAb is submitted to a constant-pH (CpH) complexation study using FORTE where the two macromolecules can titrate, translate and rotate in all directions. After each simulation run, the free energy of interaction is saved for comparison at the end of the full cycle. A substitution of an amino acid by ALA is introduced in the wt mAb. A new complexation run is carried out with this new binder (i.e., the wild type with a new single ALA mutation). The process is systematically repeated for all residues. Only one residue is replaced by ALA per time. This procedure corresponds to step 8 in Figure 1 (see the main text).

We included in our evaluation different mutations present on RBD for different past and present variants of concern (VOC): (**a**) N501Y (Alpha/B.1.1.7), (**b**) K417N, E484K and N501Y (Beta/B.1.351), (**c**) K417T, E484K and N501Y (Gamma/P.1), (**d**) L452R, T478K and E484Q (Delta/B.1.617.2), (**e**) G339D, S371L, S373P, S375F, K417N, N440K, G446S, S477N, T478K, E484A, Q493R, G496S, Q498R, N501Y, Y505H (Omicron-BA.1/B.1.1.529), and variants of interests (VOI) [51]: (**a**) L452R (Epsilon/B.1.427/B.1.429), (**b**) E484K (Eta/B.1.525), (**c**) E484K and N501Y (Iota/B.1.526 NY), (**d**) L452R and E484Q (Kappa/B.1.617.1), and (**e**)

Y453F (mink) [52–56]. The input structures with the mutations in the SARS-CoV-2 RBDs were prepared with "UCSF Chimera 1.15" [57] by the simple replacement of the amino acid followed by a minimization with default parameters and considering the H-bonds [57]. Structures obtained by this simple procedure seem to be equivalent to other available ones [17]. For instance, a comparison between a recent RBD structure predicted by AlphaFold2 for the Omicron variant [55] with the one generated using "UCSF Chimera 1.15" gives a root-mean-square difference of 0.5 Å. This suggests that the mutations seen so far do not have a significant effect on the overall folded structure of the RBD.

In silico alanine scanning was used to determine the contribution of Ab-specific amino acids to the Ab-Ag binding. In this process, an amino acid from a given binder was replaced by ALA, and the complexation simulation was repeated for this new possible binder. This mutation was done directly at the mesoscopic level of the proteins without the need to pre-generate its coordinates by the above-described procedures. It is assumed that the tested point mutations will have a minor effect on the overall folded structure of the mAb. The minima values of $\beta w(r)$ for each new system (RBD_{wt} interacting with the mAbs classified as P_{best} or P_{best-1} carrying the ALA mutation) are recorded for comparison after all possible single replacements were fully explored one by one. This theoretical lead optimization protocol is schematically illustrated in Figure S2.

A three-cycle process was followed in this electrostatic optimization pipeline for mAb engineering with higher affinity. After all possible ALA substitutions were tested one per time, the best single mutation (i.e., the one with higher binding affinity) obtained from this first cycle was incorporated into the binder resulting in a binder'. This new protein binder' contains a single ALA mutation in a specific position that resulted in the higher RBD_{wt} affinity among all tested possibilities. The process was restarted with this new macromolecule (i.e., binder') being subjected to another TAS loop. After this second loop, the new binder (binder'') has a new amino acid substitution in its sequence together with the previous one already incorporated by the ALA single mutation in the first cycle (i.e., at this point, the binder'' has two ALA replacements [A-A]). The third cycle of TAS was also done using the binder'' with the double ALA mutation to explore the effect of an additional ALA replacement in its sequence (A-A-A). Besides TAS, equivalent tests were also done with the substitution of any residue by GLU in a "theoretical glutamic acid scanning". This acid residue was chosen based on a previous work where we could observe some dependency on an increase in the binding affinities with a decrease in the total net charges of the binders [17]. This process with three cycles was repeated for GLU as done with ALA. All these replacements by either ALA or GLU were combined in different ways (A-A-A, A-A-E, A-E-E, and E-E-E). Each possible combination was evaluated to guide the selection of the systems that improved the binding affinity for the RBD_{wt}. Only up to three simultaneous substitutions in the wildtype mAb were tested to avoid a complete mischaracterization of the original template protein (for both mAbs candidates classified as P_{best} or P_{best-1}). This could result in a complete unfolding of the macromolecule and/or an increase in the chances for Ab-Ag aggregation. Moreover, at this point, the number of different simulated systems was too large, and the computational costs started to be prohibitive even for a relatively cheaper CG model. Finally, the most promising mAb (identified as P_{best} in our

classification) designed with these three engineering cycles, i.e., the one with the higher RBD_{wt} binding affinity, was tested with the other RBDs from the main VOCs and VOIs. Additionally, titration simulations employing the FPTS[58] were used to provide the main physical-chemical properties of the individual macromolecules (all from the wild-type to P10 and each new macromolecule produced by the binding optimization technique with a single, double and triple mutation). In these runs, the total net charge numbers of these binders were computed in the absence of the RBD, i.e., the FTPS was employed for a single binder in an electrolyte solution. Such a large set of net charges data was used to further investigate the possible correlations between the binding affinities and the mAb´s net charge. This was useful to complement the previous initial analysis performed with a quite smaller number of pairs of RBD-binders in comparison with this larger data set [17].

Table S1. mAb candidates that meet the selection criteria described in the "Theoretical methods" section of the paper's main body (ΔG < - 150 REU, SASA > 1900 Å^2). Red is for the (A) approach; blue for (B) and black for (C) approaches as described in the main text.

	ΔG (REU)	SASAh (Å^2)	SASA (Å^2)	SASAp (Å^2)	nHB	nres	Candidate code	
1	-362	1078	2009	932	9	80	INTERFACE_rescored_01.pose.6w41-FabOut.10_0119_0001	**P01**
2	-357	1022	1979	957	8	78	INTERFACE_rescored_01.pose.6w41-FabOut.05_0483_0001	**P02**
3	-298	1141	1955	814	3	86	INTERFACE_rescored_02.pose.6w41-FabOut.05_0427_0001	**P03**
4	-282	1058	2029	971	10	94	INTERFACE_rescored_03.pose.6w41-FabOut.24_0219_0001	**P04**
5	-272	1166	1941	775	7	80	INTERFACE_rescored_04.pose.6w41-FabOut.05_0291_0001	**P05**
6	-251	1087	1995	908	6	83	INTERFACE_rescored_03.pose.6w41-FabOut.03_0220_0001	**P06**
7	-248	1190	2065	875	6	93	INTERFACE_rescored_08.pose.6w41-FabOut.08_0886_0001	**P07**
8	-242	1155	1994	839	7	81	INTERFACE_rescored_10.pose.6w41-FabOut.21_00364_0001	**P08**
9	-230	1110	1966	856	8	81	INTERFACE_rescored_18.pose.6w41-FabOut.06_00190_0001	**P09**
10	-227	1092	1932	840	8	81	INTERFACE_rescored_07.pose.6w41-FabOut.03_0550_0001	**P10**
11	-226	1135	1919	785	5	82	INTERFACE_rescored_12.pose.6w41-FabOut.08_0558_0001	
12	-220	977	1916	939	9	86	INTERFACE_rescored_01.pose.6w41-FabOut.01_0110_0001	
13	-201	986	1928	942	8	81	INTERFACE_rescored_14.pose.6w41-FabOut.19_0918_0001	
14	-198	1048	1999	952	9	86	INTERFACE_rescored_11.pose.6w41-FabOut.03_0437_0001	
15	-186	1156	1951	795	5	83	INTERFACE_rescored_13.pose.6w41-FabOut.24_0566_0001	
16	-185	1100	2005	905	11	84	INTERFACE_rescored_11.pose.6w41-FabOut.20_0689_0001	
17	-184	1053	1968	915	7	80	INTERFACE_rescored_17.pose.6w41-FabOut.09_0149_0001	
18	-181	1102	1995	893	9	82	INTERFACE_rescored_19.pose.6w41-FabOut.02_0806_0001	
19	-178	1099	2082	982	9	85	INTERFACE_rescored_18.pose.6w41-FabOut.17_0043_0001	
20	-177	997	1954	957	11	78	INTERFACE_rescored_16.pose.6w41-FabOut.19_0451_0001	
21	-177	1016	1940	924	9	81	INTERFACE_rescored_13.pose.6w41-FabOut.15_0068_0001	
22	-176	1115	2052	937	10	85	INTERFACE_rescored_15.pose.6w41-FabOut.04_0788_0001	
23	-175	1013	1907	893	8	79	INTERFACE_rescored_10.pose.6w41-FabOut.08_0719_0001	
24	-175	1040	1930	890	9	85	INTERFACE_rescored_14.pose.6w41-FabOut.13_0109_0001	
25	-175	1184	2079	895	8	81	INTERFACE_rescored_20.pose.6w41-FabOut.02_0069_0001	
26	-175	1067	1937	870	7	83	INTERFACE_rescored_12.pose.6w41-FabOut.02_0701_0001	
27	-175	1133	2000	867	8	85	INTERFACE_rescored_17.pose.6w41-FabOut.08_0226_0001	
28	-174	1117	1911	794	8	82	INTERFACE_rescored_20.pose.6w41-FabOut.23_0912_0001	
29	-173	1042	1959	917	8	83	INTERFACE_rescored_06.pose.6w41-FabOut.14_0537_0001	
30	-172	1004	1989	984	10	81	INTERFACE_rescored_14.pose.6w41-FabOut.12_0003_0001	
31	-170	1072	1925	853	8	82	INTERFACE_rescored_05.pose.6w41-FabOut.18_0099_0001	
32	-170	959	1903	944	10	78	INTERFACE_rescored_18.pose.6w41-FabOut.16_0459_0001	
33	-165	1047	1968	921	11	83	INTERFACE_rescored_16.pose.6w41-FabOut.13_0289_0001	
34	-165	1105	2189	1084	8	80	INTERFACE_rescored_20.pose.6w41-FabOut.10_0302_0001	
35	-165	1119	1974	855	8	86	INTERFACE_rescored_19.pose.6w41-FabOut.20_0138_0001	
36	-164	1043	2237	1194	5	105	INTERFACE_rescored_11.pose.6w41-FabOut.20_0897_0001	
37	-161	1144	2157	1012	10	85	INTERFACE_rescored_18.pose.6w41-FabOut.19_0255_0001	
38	-159	1130	1932	801	8	84	INTERFACE_rescored_17.pose.6w41-FabOut.24_0100_0001	
39	-159	1054	2064	1010	12	85	INTERFACE_rescored_20.pose.6w41-FabOut.14_0421_0001	
40	-155	1028	1933	905	10	78	INTERFACE_rescored_07.pose.6w41-FabOut.19_0010_0001	
41	-151	1113	2002	889	6	82	INTERFACE_rescored_09.pose.6w41-FabOut.06_0158_0001	

SASA: solvent-accessible surface area of the mAb-RBD interface
SASAh: hydrophobic SASA
SASAp: polar SASA
nHB: number of hydrogen bonds at the interface
nres: number of residues participating in the interface

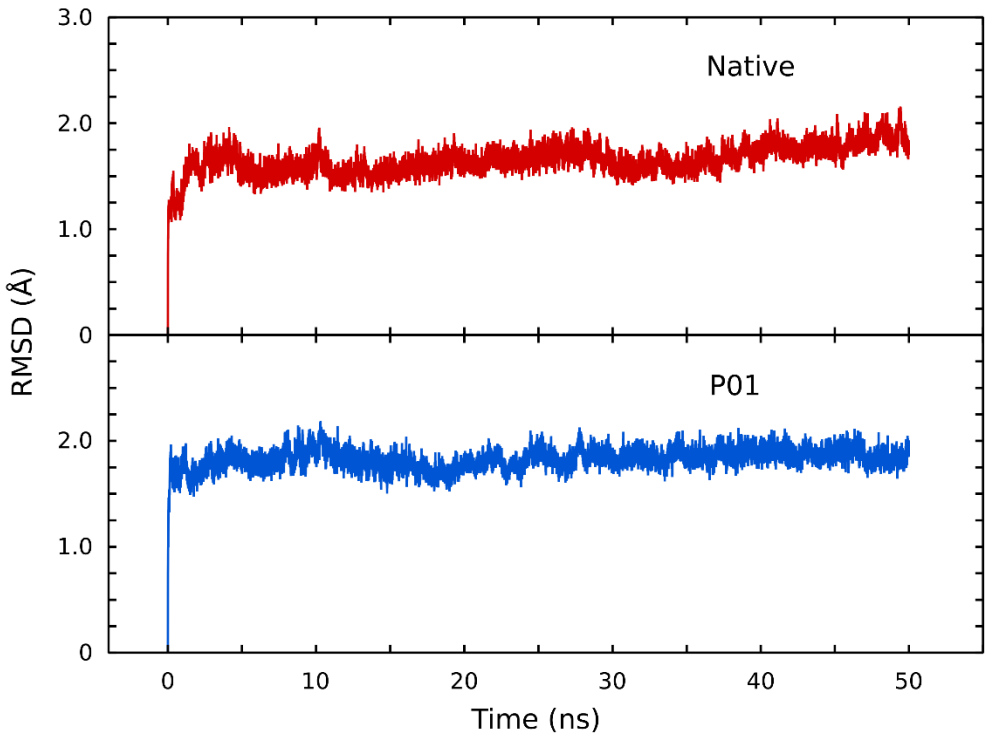

Figure S3. Root-mean-square deviation (RMSD) of the residues situated at the interaction interface between the antibodies (native and P01) and RBD during the atomistic molecular dynamics' simulations. A cut-off of 3 Å was used to define the antibody-antigen interaction interface.

S2) Supplementary discussions on atomistic MD simulations results

Mutations designed in the P01 candidate were located in all six CDRs on both the light (L) and heavy (H) chains. These mutations improved specificity and affinity by favoring the formation of supplementary salt bridges and hydrogen bonds as well as hydrophobic interactions. Protein-protein interactions are complex processes that, at short distances, are dependent on the interface properties consisting of shape complementarity, size, short-range electrostatic interactions, and polar and hydrophobic interactions. Specificity in protein-protein binding is mainly given by electrostatic interactions [59] while non-specific and van der Waals interactions are the driving force for increased affinity [60]. Thus, for each saved MD trajectory frame the number of salt bridges formed between charged residues of the mAb and RBD_{wt} was calculated. A salt bridge was considered to exist between acidic/basic residues if the distance between the heavy atoms of the charged moieties (i.e., side-chain carboxyl and amino) was shorter than 3.5 Å. Using the entire pool (full trajectory) of calculated salt-bridge numbers, a histogram was extracted concerning the number of salt bridges (i.e., the distribution of several frames with a particular number of salt bridges). The histogram was then normalized considering the area under the curve to be 1 to obtain the PDF of the mAb-RDB_{wt} salt-bridges number. The number of non-polar contacts was obtained by counting, for each frame, the hydrophobic residue atom pairs (mAb-to-RBD_{wt}) within a distance shorter than 3.5 Å. A similar procedure as above was applied to calculate the histogram and PDF of non-polar contacts. For the polar interaction histogram and PDF, we counted the number of hydrogen bonds between polar residues and polar to charged residues in the mAb-RBD_{wt} pair. It can be seen from the plots in Figure 8 that the main contributions to the increased affinity of P01 to RBD_{wt}, compared with the wild-type CR3022, come from an increased number of salt bridges and a higher hydrophobic contact area to RBD_{wt}. The average number of salt bridges at the P01 interface to RBD_{wt} is ~4 while for the wild-type CR3022 is ~3. The Lys35 residue, which replaced a Ser35 in the CR3022, is an important player in short-range electrostatic interactions at the P01-RBD_{wt} interface as it is strategically placed on the tip of the L1 loop and has the optimum side-chain length to form a strong salt-bridge with the acidic Glu in position 516 on the RBD_{wt}. The average number of hydrophobic contacts between P01 and RBD_{wt} (~95) is significantly bigger than for the CR3022-RBD_{wt} pair (~125), which will result in a higher binding affinity. Four mutations in P01, that replace original residues in CR3022 with more hydrophobic ones, contribute to this increased hydrophobic contact area: W40F, S110A, G111A, and I112L. They form a hydrophobic patch in the middle of the mAb–RBD_{wt} interface that interacts with Lys378(4 methylene groups)-Tyr380-Pro-384-Tyr369-Val382 on the RBD_{wt} counterpart.

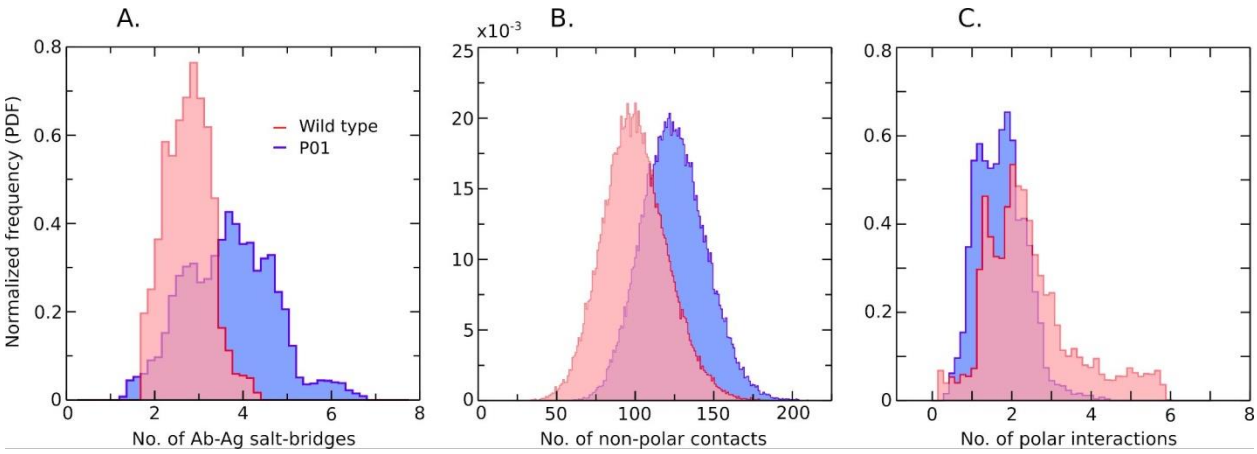

Figure S4. Molecular descriptors were used to analyze the contributions of different interaction types to RBD_{wt}-P01 binding: normalized probability distribution functions (PDF) of Ab-Ag salt-bridges (A), non-polar contacts (B), and hydrogen bond (C) numbers.

To get a detailed picture of the RBD_{wt}-P01 interaction interface, we have analyzed the impact of the introduced mutations on the local molecular environment of the complex, which is discussed below for each CDR separately.

L1 CDR design. The tyrosine residue in position 31 in L1 CDR was replaced by Asn in P01, which easily participates in hydrogen bonding with both Thr430 and Asp428 on RBD_{wt}. The interaction of the native Tyr31 with Thr430 is impeded due to the longer side chain of tyrosine compared with asparagine. One key mutation present in L1 is S35K which replaces a neutral serine residue with a positively charged lysine. The original Ser35 residue is located on the tip of the L1 loop that penetrates deep into the Ab-Ag interface stabilizing the complex. Replacing the serine with lysine in this position provides a strong salt bridge formation between the long basic side chain of Lys35 and the acidic Glu in position 516 on the RBD_{wt}, an interaction not present in the native CR3022. Replacement of Asn with Gln in position 37 provided further stabilization as the longer side chain of glutamine compared with asparagine allowed for a new interaction of Gln37 on the antibody with His519 on the RBD_{wt}.

L2 CDR design. The replacement of the neutral hydrophilic Ser72 with an acidic Glu residue in P01 introduces a strong salt bridge formation between Glu72 and the basic Lys386 on the RBD_{wt}.

L3 CDR design. Another mutation that increases binding specificity and affinity was introduced in the L3 loop replacing Tyr110 with Arg in P01. The substitution of neutral tyrosine with the positively charged arginine allows for a salt bridge formation with the neighboring Asp428 on RBD_{wt}. The P01-designed L3 CDR has two supplementary residues Pro135 and His136 compared with the native CR3022. This makes the L3 loop more extended towards RBD_{wt}, increasing the Ab-Ag contact surface area. The L3 loop also contains the mutation of Thr112 to Tyr112 which, due to the new conformation of L3, allows for hydrogen bond formation between Tyr112 and three partners on RBD_{wt}: the side chain of Glu414, the backbone carbonyl

oxygen atom of Pro412, and the side chain of the acidic Asp427, interactions not present in the native CR3022.

H1 CDR design. The Tyr residue in position 28 was replaced by Gly. This mutation seems advantageous because the smaller side chain of glycine and the particular conformation of the designed H1 loop allowed the side chain of Tyr369 from the RBD_{wt} to become buried on the mAb surface. The hydroxyl moiety of Tyr369 can also participate in hydrogen bonding with the side chain -NH group of Trp39, a residue that represents another mutation in the Rosetta-designed H1 loop: Y39W. Trp40 was replaced by a more hydrophobic Phe residue, whose side-chain inserts between the cycle of Tyr380 and the long hydrocarbon chain of Lys387, thus hydrophobically stabilizing the Ab-Ag interaction.

H3 CDR design. Serine in position 133 was mutated to a His residue which inserts in between G381 and Asp427 on the RBD_{wt}, forming a hydrogen bond-mediated bridge between the -NH group of G381 and the backbone -C=O group of Asp427.

S3) Supplementary discussions on the total net charge of the mAbs and their binding affinities

In summary, more negatively charged mAbs tend to have a higher RBD affinity[34]. Any negatively charged binder will have a stronger attraction for the newer variants as they become more positively charged (see ref. [61]). P01''' has a net charge number of -4.7 at pH 7 while the RBD from the wt and Omicron have +2.2 and +5.2, respectively. In Figure S5a, this charge-free energy correlation is explored for the WT and Px (x: 01 to 10). Although there is no perfect linear behavior, it can be seen that there is indeed this tendency. This is more clearly seen in Figure S5b, where the data from all single, double and triple mutations investigated by the theoretical ALA/GLU scanning is shown together. This data suggests that Coulombic forces are the main driving forces for the association of this antigen with the studied mAbs. There are now accumulative data on the importance of electrostatic interactions for this system [17,19,33,34]. Certainly, other physical interactions can participate in this associative mechanism that is not solely but largely controlled by Coulombic forces. Therefore, the total net charge of the mAb is an important physical-chemical parameter to design better binders with higher affinity. This is what was next explored towards a more specific and efficient macromolecule capable of preventing viral infection.

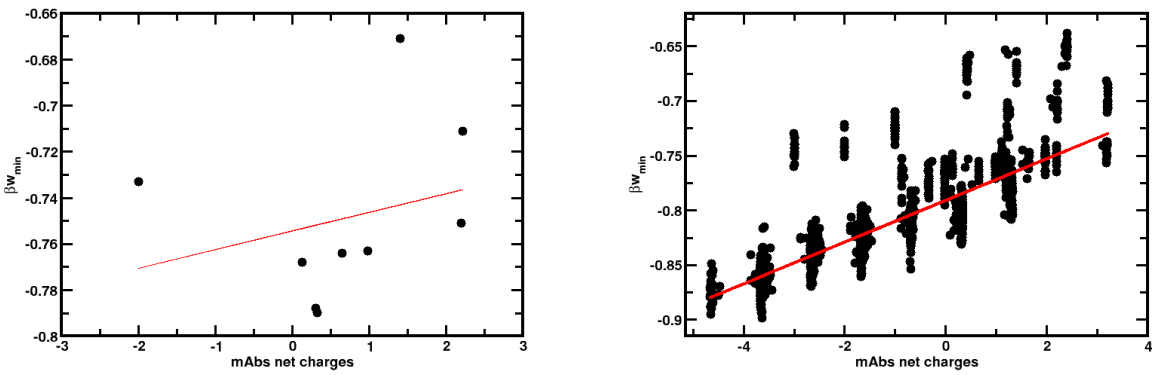

Figure S5. Correlation between the total net charge number of each fragment of mAb and the corresponding βw_{min} values for their complexations with SARS-CoV-2 RBD_{wt}. (a) *Left panel*: Data for each fragment of Rosetta-designed mAb (P01 to P10). (a) *Right panel*: Data for each mutated fragment of P01 and P06 was obtained during the theoretical ALA scanning procedure (correlation coefficient equals 0.8). The data for the total net charge numbers were obtained from titration simulations with a single protein in the absence of the RBD_{wt} by the FPTS at pH 7. All βw_{min} values were computed from RBD_{wt}-complexation studies using FORTE at pH 7 and 150mM of NaCl.

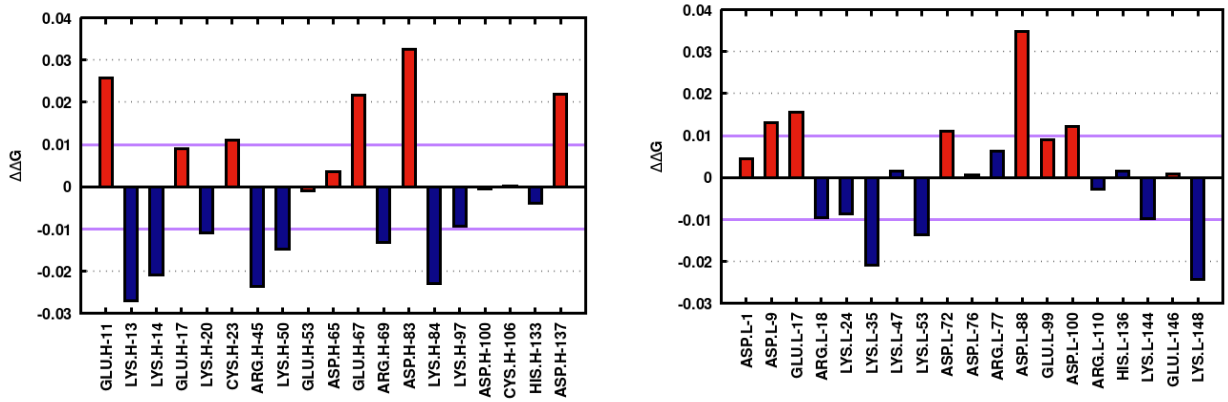

Figure S6. Individual residue electrostatic contributions to the complexation RBD_{wt}-P01 obtained with the theoretical ALA scanning procedure applied to P01. Acid and basic residues are represented by red and blue, respectively. Data for the amino acids belonging to the light (L) and heavy (H) chains are given in the left and right panels, respectively. Calculations were done with the FORTE approach. ΔΔG (in K_BT units) is defined as the difference between the minimum value measured in $\beta w(r)$ for the complexation RBD_{wt}-P01' [βw_{min}(P01')] and the corresponding quantity for the original P01 [βw_{min}(P01)] for each new mutation. See the main text for other details.

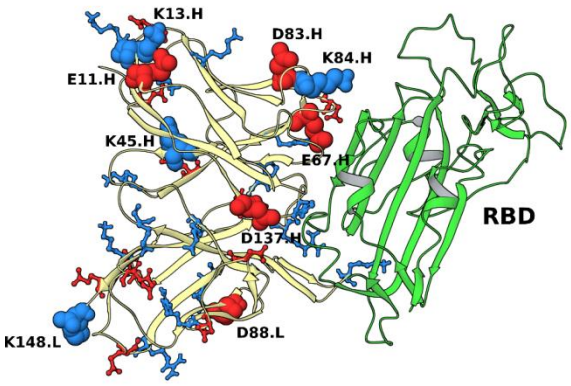

Figure S7. Molecular structure of the RBD_{wt} complexed with P01. The RBD_{wt} is represented in green. All ionizable residues from the binder P01 whose single ALA mutation resulted in representative ΔΔGs (values larger than the estimated errors) are highlighted by blue (basic residues) and red (acid residues). The most important mutations (see Figure S5) are labeled and represented with vdW spheres. The letters L and H are used to refer to the light and heavy chain, respectively.

S4) Supplementary discussions on the precision of the FORTE data

One of the main advantages of CG methods such as FORTE is the smoother energy landscape that allows for reasonable precision in the results. This can be easily seen by comparing different replicate runs. Some examples are given in Figure S8 and Table S2 above. Figure S8 displays the free energy of interactions for the complexation of the $RBD_{(wt)}$-P01 system obtained by different replicate runs. In table S2, we compiled βw_{min} values for some selected simulated cases for runs with a different number of MC steps. The other studied systems behave in the same manner. As larger the sampling, **i)** better populated are all the histogram bins used for computing the g(r) during the sampling, **ii)** and reduced the statistical noises in the calculated βw(r).[17] A shorter number of MC steps decreases the precision of the βw_{min} values as expected. Due to the relatively lower CPU cost of the FORTE runs than other approaches, longer simulation runs are possible increasing the precision of the computed free energy of interactions for the studied protein-protein systems by FORTE.

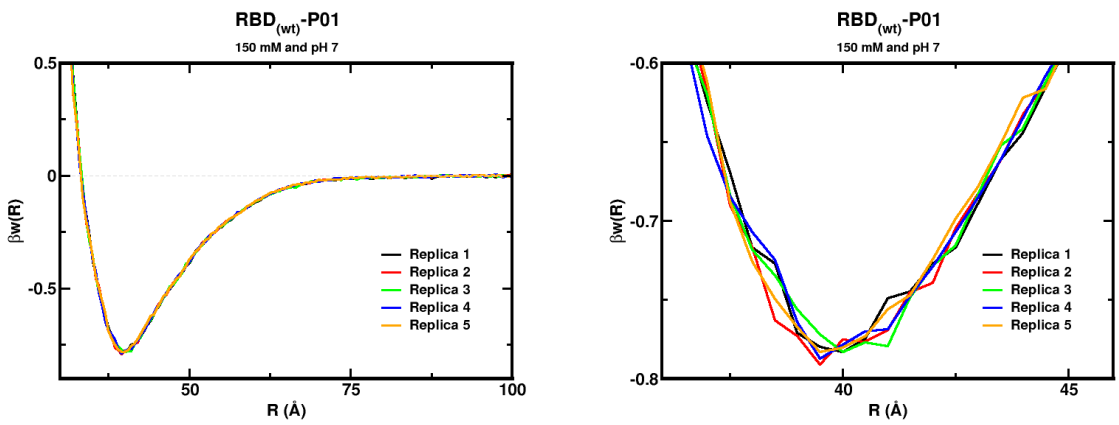

Figure S8. Free energy profiles for the interaction of $RBD_{(wt)}$ with P01. The simulated free energy of interactions as a function of the separation distances [βw(r)] between the centers of the RBD and the binder is given at pH 7.0 and 150 mM of NaCl for 5 different independent runs. (a) *Left panel:* Data for a full range of separation distances. (b) *Right panel:* Data for the short-range separation distance highlighting the well-depths of each replicate run.

Table S2. Mean βw_{min} values for some selected simulated cases as a function of the number of MC steps of the production phase. (*) The number of MC steps used for production in this work. The estimated errors are given between parentheses and were based on 5 replicates.

Studied complex	Mean value for βw_{min}	
	$3x10^9$ MC steps(*)	**$3x10^8$ MC steps**
RBD(wt)-P01	-0.785(3)	-0.81(2)
RBD(Delta)-P01	-0.796(5)	-0.83(2)
RBD(Omicron)-P01	-0.832(6)	-0.87(2)
RBD(wt)-P01'''	-0.883(6)	-0.91(3)
RBD(Delta)-P01'''	-0.951(5)	-0.98(3)
RBD(Omicron)-P01'''	-0.985(6)	-1.03(4)
RBD(wt)-P06	-0.790(7)	-0.80(2)
RBD(Delta)-P06	-0.798(5)	-0.82(2)
RBD(Omicron)-P06	-0.83(1)	-0.85(2)
RBD(wt)-P06'''	-0.843(5)	-0.85(2)
RBD(Delta)-P06'''	-0.880(7)	-0.89(1)
RBD(Omicron)-P06'''	-0.93(1)	-0.95(2)

References

(1) Barroso da Silva, F. L.; Carloni, P.; Cheung, D.; Cottone, G.; Donnini, S.; Foegeding, E. A.; Gulzar, M.; Jacquier, J. C.; Lobaskin, V.; MacKernan, D.; Mohammad Hosseini Naveh, Z.; Radhakrishnan, R.; Santiso, E. E. Understanding and Controlling Food Protein Structure and Function in Foods: Perspectives from Experiments and Computer Simulations. *Annu. Rev. Food Sci. Technol.* **2020**, *11* (1), 365–387. https://doi.org/10.1146/annurev-food-032519-051640.

(2) Dascalu, A. I.; Ardeleanu, R.; Neamtu, A.; Maier, S. S.; Uritu, C. M.; Nicolescu, A.; Silion, M.; Peptanariu, D.; Calin, M.; Pinteala, M. Transfection-Capable Polycationic Nanovectors Which Include PEGylated-Cyclodextrin Structural Units: A New Synthesis Pathway. *J. Mater. Chem. B* **2017**, *5* (34), 7164–7174. https://doi.org/10.1039/C7TB01722G.

(3) Engelbrecht, L. de V.; Mocci, F.; Wang, Y.; Perepelytsya, S.; Vasiliu, T.; Laaksonen, A. Molecular Perspective on Solutions and Liquid Mixtures from Modelling and Experiment. In *Soft Matter Systems for Biomedical Applications*; Bulavin, L., Lebovka, N., Eds.; Springer Proceedings in Physics; Springer International Publishing: Cham, 2022; Vol. 266, pp 53–84. https://doi.org/10.1007/978-3-030-80924-9_3.

(4) Gunsteren, W. F. van; Luque, F. J.; Timms, D.; Torda, A. E. Molecular Mechanics in Biology: From Structure to Function, Taking Account of Solvation. *Annu Rev Biophys Biomol Struct.* **1994**, *23*, 847–863.

(5) Lyubartsev, A.; Laaksonen, A. Inverse Problems and Hierarchical Multiscale Modelling of Biological Matter. In *New Trends in Macromolecular and Supramolecular Chemistry for Biological Applications*; J.M. Abadie, M., Pinteala, M., Rotaru, A., Eds.; Springer International Publishing: Cham, 2021; pp 213–237. https://doi.org/10.1007/978-3-030-57456-7_11.

(6) Minea, B.; Marangoci, N.; Peptanariu, D.; Rosca, I.; Nastasa, V.; Corciova, A.; Varganici, C.; Nicolescu, A.; Fifere, A.; Neamtu, A.; Mares, M.; Barboiu, M.; Pinteala, M. Inclusion Complexes of Propiconazole Nitrate with Substituted β-Cyclodextrins: The Synthesis and in Silico and in Vitro Assessment of Their Antifungal Properties. *New J. Chem.* **2016**, *40* (2), 1765–1776. https://doi.org/10.1039/C5NJ01811K.

(7) Steinhauser, M.; Hiermaier, S. A Review of Computational Methods in Materials Science: Examples from Shock-Wave and Polymer Physics. *Int. J. Mol. Sci.* **2009**, *10* (12), 5135–5216. https://doi.org/10.3390/ijms10125135.

(8) van Gunsteren, W. F.; Dolenc, J. Thirty-Five Years of Biomolecular Simulation: Development of Methodology, Force Fields and Software. *Mol. Simul.* **2012**, *38* (14–15), 1271–1281. https://doi.org/10.1080/08927022.2012.701744.

(9) Andreatta, M.; Nielsen, M. Bioinformatics Tools for the Prediction of T-Cell Epitopes. In *Epitope Mapping Protocols*; Rockberg, J., Nilvebrant, J., Eds.; Methods in Molecular Biology; Springer New York: New York, NY, 2018; pp 269–281. https://doi.org/10.1007/978-1-4939-7841-0_18.

(10) Ibrahim, B.; McMahon, D. P.; Hufsky, F.; Beer, M.; Deng, L.; Mercier, P. L.; Palmarini, M.; Thiel, V.; Marz, M. A New Era of Virus Bioinformatics. *Virus Res.* **2018**, *251*, 86–90. https://doi.org/10.1016/j.virusres.2018.05.009.

(11) Pappas, N.; Roux, S.; Hölzer, M.; Lamkiewicz, K.; Mock, F.; Marz, M.; Dutilh, B. E. Virus Bioinformatics. In *Encyclopedia of Virology*; Elsevier, 2021; pp 124–132. https://doi.org/10.1016/B978-0-12-814515-9.00034-5.

(12) Sironi, M.; Kaderali, L. Bioinformatics Algorithms and Predictive Models: The Grand Challenge in Computational Virology. *Front. Virol.* **2021**, *1*, 684608. https://doi.org/10.3389/fviro.2021.684608.

(13) Barnes, C. O.; Jette, C. A.; Abernathy, M. E.; Dam, K.-M. A.; Esswein, S. R.; Gristick, H. B.; Malyutin, A. G.; Sharaf, N. G.; Huey-Tubman, K. E.; Lee, Y. E.; Robbiani, D. F.; Nussenzweig, M. C.; West, A. P.; Bjorkman, P. J. SARS-CoV-2 Neutralizing Antibody Structures Inform Therapeutic Strategies. *Nature* **2020**, *588* (7839), 682–687. https://doi.org/10.1038/s41586-020-2852-1.

(14) Tian, X.; Li, C.; Huang, A.; Xia, S.; Lu, S.; Shi, Z.; Lu, L.; Jiang, S.; Yang, Z.; Wu, Y.; Ying, T. Potent Binding of 2019 Novel Coronavirus Spike Protein by a SARS Coronavirus-Specific Human

Monoclonal Antibody. *Emerg. Microbes Infect.* **2020**, *9* (1), 382–385. https://doi.org/10.1080/22221751.2020.1729069.
(15) Wu, N. C.; Yuan, M.; Bangaru, S.; Huang, D.; Zhu, X.; Lee, C.-C. D.; Turner, H. L.; Peng, L.; Yang, L.; Burton, D. R.; Nemazee, D.; Ward, A. B.; Wilson, I. A. A Natural Mutation between SARS-CoV-2 and SARS-CoV Determines Neutralization by a Cross-Reactive Antibody. *PLOS Pathog.* **2020**, *16* (12), e1009089. https://doi.org/10.1371/journal.ppat.1009089.
(16) Yuan, M.; Wu, N. C.; Zhu, X.; Lee, C.-C. D.; So, R. T. Y.; Lv, H.; Mok, C. K. P.; Wilson, I. A. A Highly Conserved Cryptic Epitope in the Receptor-Binding Domains of SARS-CoV-2 and SARS-CoV. *Science* **2020**, eabb7269. https://doi.org/10.1126/science.abb7269.
(17) Corrêa Giron, C.; Laaksonen, A.; Barroso da Silva, F. L. On the Interactions of the Receptor-Binding Domain of SARS-CoV-1 and SARS-CoV-2 Spike Proteins with Monoclonal Antibodies and the Receptor ACE2. *Virus Res.* **2020**, *285*, 198021. https://doi.org/10.1016/j.virusres.2020.198021.
(18) Ding, H.; Yin, Y.; Ni, S.; Sheng, Y.; Ma, Y. Accurate Evaluation on the Interactions of SARS-CoV-2 with Its Receptor ACE2 and Antibodies CR3022/CB6*. *Chin. Phys. Lett.* **2021**, *38* (1), 018701. https://doi.org/10.1088/0256-307X/38/1/018701.
(19) Nguyen, H.; Lan, P. D.; Nissley, D. A.; O'Brien, E. P.; Li, M. S. Electrostatic Interactions Explain the Higher Binding Affinity of the CR3022 Antibody for SARS-CoV-2 than the 4A8 Antibody. *J. Phys. Chem. B* **2021**, *125* (27), 7368–7379. https://doi.org/10.1021/acs.jpcb.1c03639.
(20) Shariatifar, H.; Farasat, A. Affinity Enhancement of CR3022 Binding to RBD; in Silico Site Directed Mutagenesis Using Molecular Dynamics Simulation Approaches. *J. Biomol. Struct. Dyn.* **2021**, *0* (0), 1–10. https://doi.org/10.1080/07391102.2021.2004230.
(21) Lagoumintzis, G.; Chasapis, C. T.; Alexandris, N.; Kouretas, D.; Tzartos, S.; Eliopoulos, E.; Farsalinos, K.; Poulas, K. Nicotinic Cholinergic System and COVID-19: In Silico Identification of Interactions between A7 Nicotinic Acetylcholine Receptor and the Cryptic Epitopes of SARS-Co-V and SARS-CoV-2 Spike Glycoproteins. *Food Chem. Toxicol.* **2021**, *149*, 112009. https://doi.org/10.1016/j.fct.2021.112009.
(22) van Zundert, G. C. P.; Rodrigues; Trellet, M.; Schmitz, C.; Kastritis, P. L.; Karaca, E.; Melquiond, A. S. J.; van Dijk, M.; de Vries, S. J.; Bonvin, A. M. J. J. The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J. Mol. Biol.* **2016**, *428* (4), 720–725. https://doi.org/10.1016/j.jmb.2015.09.014.
(23) Verkhivker, G. M.; Di Paola, L. Integrated Biophysical Modeling of the SARS-CoV-2 Spike Protein Binding and Allosteric Interactions with Antibodies. *J. Phys. Chem. B* **2021**, *125* (18), 4596–4619. https://doi.org/10.1021/acs.jpcb.1c00395.
(24) Riahi, S.; Lee, J. H.; Wei, S.; Cost, R.; Masiero, A.; Prades, C.; Olfati-Saber, R.; Wendt, M.; Park, A.; Qiu, Y.; Zhou, Y. Application of an Integrated Computational Antibody Engineering Platform to Design SARS-CoV-2 Neutralizers. *Antib. Ther.* **2021**, tbab011. https://doi.org/10.1093/abt/tbab011.
(25) Burley, S.; Berman, H.; Kleywegt, G.; Markley, J.; Nakamura, H.; Velankar, S. Protein Data Bank (PDB): The Single Global Macromolecular Structure Archive. In *Protein Crystallography*; Wlodawer, A., Dauter, Z., Jaskolski, M., Eds.; Methods in Molecular Biology; Springer New York, 2017; Vol. 1607, pp 627–641. https://doi.org/10.1007/978-1-4939-7000-1_26.
(26) Vilar, S.; Cozza, G.; Moro, S. Medicinal Chemistry and the Molecular Operating Environment (MOE): Application of QSAR and Molecular Docking to Drug Discovery. *Curr. Top. Med. Chem.* **2008**, *8* (18), 1555–1572. https://doi.org/10.2174/156802608786786624.
(27) Ó Conchúir, S.; Barlow, K. A.; Pache, R. A.; Ollikainen, N.; Kundert, K.; O'Meara, M. J.; Smith, C. A.; Kortemme, T. A Web Resource for Standardized Benchmark Datasets, Metrics, and Rosetta Protocols for Macromolecular Modeling and Design. *PLOS ONE* **2015**, *10* (9), e0130433. https://doi.org/10.1371/journal.pone.0130433.

(28) Weitzner, B. D.; Jeliazkov, J. R.; Lyskov, S.; Marze, N.; Kuroda, D.; Frick, R.; Adolf-Bryfogle, J.; Biswas, N.; Dunbrack, R. L.; Gray, J. J. Modeling and Docking of Antibody Structures with Rosetta. *Nat. Protoc.* **2017**, *12* (2), 401–416. https://doi.org/10.1038/nprot.2016.180.
(29) Maguire, J. B.; Haddox, H. K.; Strickland, D.; Halabiya, S. F.; Coventry, B.; Griffin, J. R.; Pulavarti, S. V. S. R. K.; Cummins, M.; Thieker, D. F.; Klavins, E.; Szyperski, T.; DiMaio, F.; Baker, D.; Kuhlman, B. Perturbing the Energy Landscape for Improved Packing during Computational Protein Design. *Proteins Struct. Funct. Bioinforma.* **2021**, *89* (4), 436–449. https://doi.org/10.1002/prot.26030.
(30) Martí, D.; Alsina, M.; Alemán, C.; Bertran, O.; Turon, P.; Torras, J. Unravelling the Molecular Interactions between the SARS-CoV-2 RBD Spike Protein and Various Specific Monoclonal Antibodies. *Biochimie* **2021**, S0300-9084(21)00249-2. https://doi.org/10.1016/j.biochi.2021.10.013.
(31) Abraham, M. J.; van der Spoel, D.; Lindahl, E.; Hess, B.; and the GROMACS development team. GROMACS User Manual Version 2019, 2019.
(32) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25. https://doi.org/10.1016/j.softx.2015.06.001.
(33) Bai, C.; Warshel, A. Critical Differences between the Binding Features of the Spike Proteins of SARS-CoV-2 and SARS-CoV. *J. Phys. Chem. B* **2020**, *124* (28), 5907–5912. https://doi.org/10.1021/acs.jpcb.0c04317.
(34) Nguyen, H. L.; Lan, P. D.; Thai, N. Q.; Nissley, D. A.; O'Brien, E. P.; Li, M. S. Does SARS-CoV-2 Bind to Human ACE2 More Strongly Than Does SARS-CoV? *J. Phys. Chem. B* **2020**, *124* (34), 7336–7347. https://doi.org/10.1021/acs.jpcb.0c04511.
(35) Poveda-Cuevas, S. A.; Etchebest, C.; Barroso da Silva, F. L. Identification of Electrostatic Epitopes in Flavivirus by Computer Simulations: The PROCEEDpKa Method. *J. Chem. Inf. Model.* **2020**, *60* (2), 944–963. https://doi.org/10.1021/acs.jcim.9b00895.
(36) Poveda-Cuevas, S. A.; Etchebest, C.; Barroso da Silva, F. L. Insights into the ZIKV NS1 Virology from Different Strains through a Fine Analysis of Physicochemical Properties. *ACS Omega* **2018**, *3* (11), 16212–16229. https://doi.org/10.1021/acsomega.8b02081.
(37) Xie, Y.; Guo, W.; Lopez-Hernadez, A.; Teng, S.; Li, L. The PH Effects on SARS-CoV and SARS-CoV-2 Spike Proteins in the Process of Binding to HACE2. *Res. Sq.* **2021**, rs.3.rs-871118. https://doi.org/10.21203/rs.3.rs-871118/v1.
(38) Poveda-Cuevas, S. A.; Etchebest, C.; Silva, F. L. B. da. Self-Association Features of NS1 Proteins from Different Flaviviruses. *Virus Res.* **2022**, 198838. https://doi.org/10.1016/j.virusres.2022.198838.
(39) Giron, C. C.; Laaksonen, A.; Silva, F. L. B. da. Differences between Omicron SARS-CoV-2 RBD and Other Variants in Their Ability to Interact with Cell Receptors and Monoclonal Antibodies. *J. Biomol. Struct. Dyn.* **2022**, *0* (0), 1–21. https://doi.org/10.1080/07391102.2022.2095305.
(40) Barroso da Silva, F. L.; Pasquali, S.; Derreumaux, P.; Dias, L. G. Electrostatics Analysis of the Mutational and PH Effects of the N-Terminal Domain Self-Association of the Major Ampullate Spidroin. *Soft Matter* **2016**, *12*, 5600–5612.
(41) Delboni, L.; Barroso da Silva, F. L. On the Complexation of Whey Proteins. *Food Hydrocoll.* **2016**, *55*, 89–99.
(42) Kurut, A.; Persson, B. A.; \AAkesson, T.; Forsman, J.; Lund, M. Anisotropic Interactions in Protein Mixtures: Self Assembly and Phase Behavior in Aqueous Solution. *J Phys Chem Lett* **2012**, *3* (6), 731–734.
(43) Lund, M.; Vrbka, L.; Jungwirth, P. Specific Ion Binding to Nonpolar Surface Patches of Proteins. *J. Am. Chem. Soc.* **2008**, *130* (35), 11582–11583. https://doi.org/10.1021/ja803274p.
(44) Mendonça, D. C.; Macedo, J. N.; Guimarães, S. L.; Barroso da Silva, F. L.; Cassago, A.; Garratt, R. C.; Portugal, R. V.; Araujo, A. P. U. A Revised Order of Subunits in Mammalian Septin Complexes. *Cytoskeleton* **2019**, *76* (9–10), 457–466. https://doi.org/10.1002/cm.21569.

(45) Persson, B.; Lund, M.; Forsman, J.; Chatterton, D. E. W.; Torbjörn \AAkesson. Molecular Evidence of Stereo-Specific Lactoferrin Dimers in Solution. *Biophys Chem* **2010**, *3* (3), 187–189.
(46) Marcus, R. A. Calculation of Thermodynamic Properties of Polyelectrolytes. *J Chem Phys* **1955**, *23*, 1057.
(47) Barroso da Silva, F. L.; Dias, L. G. Development of Constant-PH Simulation Methods in Implicit Solvent and Applications in Biomolecular Systems. *Biophys. Rev.* **2017**, *9* (5), 699–728. https://doi.org/10.1007/s12551-017-0311-5.
(48) Jönsson, B.; Svensson, B. Monte Carlo Simulation of Ion–Protein Binding. In *Computer Simulation of Biomolecular Systems*; Gunsteren, W. F. van, Weiner, P. K., Wilkinson, A., Eds.; ESCOM: Leiden, 1993; Vol. 2, pp 464–482.
(49) Svensson, B.; Jönsson, B.; Woodward, C. E. Electrostatic Contributions of the Binding of $Ca^{2+\$}$ in Calbindin Mutants. A Monte Carlo Study. *Biophys Chem* **1990**, *38*, 179–183.
(50) Wade, R. C.; Gabdoulline, R. G.; Lüdemann, S. K.; Lounnas, V. Electrostatic Steering and Ionic Tethering in Enzyme-Ligand Binding: Insights from Simulations. *Proc Natl Acad Sci USA* **1998**, *95*, 5942–5949.
(51) Parums, D. V. Editorial: Revised World Health Organization (WHO) Terminology for Variants of Concern and Variants of Interest of SARS-CoV-2. *Med. Sci. Monit. Int. Med. J. Exp. Clin. Res.* **2021**, *27*, e933622-1-e933622-2. https://doi.org/10.12659/MSM.933622.
(52) Annavajhala, M. K.; Mohri, H.; Wang, P.; Nair, M.; Zucker, J. E.; Sheng, Z.; Gomez-Simmonds, A.; Kelley, A. L.; Tagliavia, M.; Huang, Y.; Bedford, T.; Ho, D. D.; Uhlemann, A.-C. Emergence and Expansion of the SARS-CoV-2 Variant B.1.526 Identified in New York. *medRxiv* **2021**, 2021.02.23.21252259. https://doi.org/10.1101/2021.02.23.21252259.
(53) Callaway, E. Multitude of Coronavirus Variants Found in the US — but the Threat Is Unclear. *Nature* **2021**, *591* (7849), 190–190. https://doi.org/10.1038/d41586-021-00564-4.
(54) Deng, X.; Garcia-Knight, M. A.; Khalid, M. M.; Servellita, V.; Wang, C.; Morris, M. K.; Sotomayor-González, A.; Glasner, D. R.; Reyes, K. R.; Gliwa, A. S.; Reddy, N. P.; Sanchez San Martin, C.; Federman, S.; Cheng, J.; Balcerek, J.; Taylor, J.; Streithorst, J. A.; Miller, S.; Sreekumar, B.; Chen, P.-Y.; Schulze-Gahmen, U.; Taha, T. Y.; Hayashi, J. M.; Simoneau, C. R.; Kumar, G. R.; McMahon, S.; Lidsky, P. V.; Xiao, Y.; Hemarajata, P.; Green, N. M.; Espinosa, A.; Kath, C.; Haw, M.; Bell, J.; Hacker, J. K.; Hanson, C.; Wadford, D. A.; Anaya, C.; Ferguson, D.; Frankino, P. A.; Shivram, H.; Lareau, L. F.; Wyman, S. K.; Ott, M.; Andino, R.; Chiu, C. Y. Transmission, Infectivity, and Neutralization of a Spike L452R SARS-CoV-2 Variant. *Cell* **2021**, *184* (13), 3426-3437.e8. https://doi.org/10.1016/j.cell.2021.04.025.
(55) Ford, C. T.; Machado, D. J.; Janies, D. A. *Predictions of the SARS-CoV-2 Omicron Variant (B.1.1.529) Spike Protein Receptor-Binding Domain Structure and Neutralizing Antibody Interactions*; 2021; p 2021.12.03.471024. https://doi.org/10.1101/2021.12.03.471024.
(56) Khateeb, J.; Li, Y.; Zhang, H. Emerging SARS-CoV-2 Variants of Concern and Potential Intervention Approaches. *Crit. Care* **2021**, *25* (1), 244. https://doi.org/10.1186/s13054-021-03662-x.
(57) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Greenblatt, G. S.; Meng, E. C.; Ferrin, T. E. UCSF Chimera: A Visualization System for Exploratory Research and Analysis. *J Comp Chem* **2004**, *25*, 1605–1612.
(58) Barroso da Silva, F. L. B. da; MacKernan, D. Benchmarking a Fast Proton Titration Scheme in Implicit Solvent for Biomolecular Simulations. *J Chem Theory Comput* **2017**, *13* (6), 2915–2929.
(59) Kumar, S.; Nussinov, R. Close-Range Electrostatic Interactions in Proteins. *ChemBioChem* **2002**, *3* (7), 604–617. https://doi.org/10.1002/1439-7633(20020703)3:7%3C604::aid-cbic604%3E3.0.co;2-x.
(60) Wodak, S. J.; Janin, J. Structural Basis of Macromolecular Recognition. In *Advances in Protein Chemistry*; Protein Modules and Protein-Protein Interaction; Academic Press, 2002; Vol. 61, pp 9–73. https://doi.org/10.1016/S0065-3233(02)61001-0.

(61) Barroso da Silva, F. L.; Giron, C. C.; Laaksonen, A. Electrostatic Features for the Receptor Binding Domain of SARS-COV-2 Wildtype and Its Variants. Compass to the Severity of the Future Variants with the Charge-Rule. *J. Phys. Chem. B* **2022**, acs.jpcb.2c04225. https://doi.org/10.1021/acs.jpcb.2c04225.