

# Computational design and modeling of nanobodies toward SARS-CoV-2 receptor binding domain

Jingyi Yang<sup>1</sup> | Zhao Zhang<sup>1</sup> | Fengyuan Yang<sup>1,2</sup> | Haiwei Zhang<sup>3</sup> | Haibo Wu<sup>4</sup> |  
Feng Zhu<sup>1,2</sup>  | Weiwei Xue<sup>1</sup> 

<sup>1</sup>School of Pharmaceutical Sciences and Innovative Drug Research Centre, Chongqing Key Laboratory of Natural Product Synthesis and Drug Research, Chongqing University, Chongqing, China

<sup>2</sup>College of Pharmaceutical Sciences, Zhejiang University, Hangzhou, China

<sup>3</sup>Chongqing Key Laboratory of Translational Research for Cancer Metastasis and Individualized Treatment, Chongqing University Cancer Hospital, Chongqing, China

<sup>4</sup>School of Life Sciences, Chongqing University, Chongqing, China

## Correspondence

Feng Zhu and Weiwei Xue, School of Pharmaceutical Sciences and Innovative Drug Research Centre, Chongqing Key Laboratory of Natural Product Synthesis and Drug Research, Chongqing University, Chongqing, China.

Emails: zhufeng@zju.edu.cn; xuww@cqu.edu.cn

## Funding information

Chongqing Municipal Education Commission, Grant/Award Number: KYYG202002; Fundamental Research Funds for Central Universities, Grant/Award Number: 2019CDYGYB005; Technology Innovation and Application Demonstration Project of Chongqing, Grant/Award Number: cstc2018jscx-msybX0287

## Abstract

The ongoing pandemic of coronavirus disease 2019 (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has become a global health concern and pose a serious threat to humanity. There is an urgent need for developing therapeutic drugs and (or) biologics to prevent the spread of the virus. The life cycle of SARS-CoV-2 shows that the virus enters host cells by first binding to angiotensin-converting enzyme 2 (ACE2) through its spike protein receptor-binding domain (RBD). Therefore, blocking the binding between of ACE2 and SARS-CoV-2 RBD can inhibit the virus infection in the host cells. In this study, by grafting the complementarity-determining regions (CDRs) of developed SARS-CoV, MERS-CoVs specific neutralizing antibodies (nAbs) include monoclonal antibodies (mAbs) as well as SARS-CoV-2 mAbs onto a known stable nanobody (Nb) scaffold, and a total of 16 Nbs sequences were designed. Five Nbs, namely CS01, CS02, CS03, CS10, and CS16, were selected based on the free energy landscape of protein docking verified by the recently reported Nb-RBD cocrystal structures. CS01, CS02, and CS03 occupied the ACE2 binding site of RBD, while CS10 and CS16 were proposed to inhibit the interaction between RBD and ACE2 through an allosteric mechanism. Based on the structures of the five Nbs in complex with RBD, seven brand-new Nbs with enhanced binding affinities (CS02\_RD01, CS03\_RD01, CS03\_RD02, CS03\_RD03, CS03\_RD04, CS16\_RD01, and CS16\_RD02) were generated by redesign of residues on the interface of the five Nbs contact with SARS-CoV-2 RBD. In addition, the identified “hot spots” on the interface of each complex provide useful information to understand the binding mechanism of designed Nbs to SARS-CoV-2 RBD. In sum, the predicted stabilities and high binding affinities of the 11 (re)designed Nbs indicating the potential of the developed computational framework in this work to design effective agents to block the infection of SARS-CoV-2.

## KEYWORDS

computational protein design, COVID-19, neutralizing nanobody, protein, protein docking, SARS-CoV-2 RBD

## 1 | INTRODUCTION

The ongoing outbreak of novel coronavirus disease 2019 (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection has led to 136,291,755 confirmed cases and 2,941,128 deaths in 223 countries, areas or territories as of April 14th, 2021 (<https://www.who.int/emergencies/diseases/novel-coronavirus-2019>). The pandemic of the disease has become a global health concern and pose a serious threat to humanity (Wang et al., 2020). To meet this challenge, great efforts have been paid for the development of therapeutic approaches against SARS-CoV-2 (Li, Zhou, et al., 2020). Due to the rapid spread of COVID-19, there is an urgent need to develop highly potent and broad-spectrum and cost-effective anti-SARS-CoV-2 drugs and vaccines.

Like the other two CoVs, SARS-CoV and Middle East respiratory syndrome CoV (MERS-CoV), SARS-CoV-2 belongs to the beta-CoV genera in the family *Coronaviridae* (Zhou, Yang, et al., 2020). Life cycle of CoV shows that the spike protein plays an essential role in viral attachment, fusion, entry, and transmission (Hoffmann et al., 2020). During infection, CoV enters host cells by first binding to their respective cellular receptors angiotensin-converting enzyme 2 (ACE2) through the virus receptor-binding domain (RBD) of the spike protein (Shang, Wan, et al., 2020). Therefore, development of neutralizing antibodies (nAbs) blocking the interaction between RBD and ACE2 plays crucial roles in inhibiting the infection of pathogenic CoVs in the host cells (Jiang et al., 2020).

Nanobodies (Nbs) are single-domain antibodies (sdAbs) derived from camelids and sharks showing a large sequence identity with the human VH gene family III (Muyldermans, 2013). The small size (~15 kDa), thermostability, high binding specificity, and low immunogenicity of Nbs making them suitable for many biotechnology and medicine applications (Muyldermans, 2013). To fight viruses and prevent their spread, Nbs can interfere at different levels of the viral replication cycle (Steeland et al., 2016). Currently, nanobody maturation technology was deployed to develop several Nbs targeting SARS-CoV-2 spike protein (Zhou, Duyvesteyn, et al., 2020). Crystal structure showed that these molecules block the interaction between RBD and ACE2 (Zhou, Duyvesteyn, et al., 2020) and their neutralizing activity against SARS-CoV-2 suggested that Nbs may serve as useful therapeutics during CoVs outbreaks (Wrapp et al., 2020).

The present work focused on design of Nbs specifically binding to SARS-CoV-2 RBD by using an integrated computational approach. Based on the developed SARS-CoV, premier released SARS-CoV-2 and MERS-CoVs specific nAbs include monoclonal antibodies (mAbs; Chen et al., 2017; Hwang et al., 2006; Li et al., 2015; Pak et al., 2009; Prabakaran et al., 2006; Walls et al., 2019; Wang et al., 2015, 2018; Ying et al., 2015; Yu et al., 2015; Yuan et al., 2020; Zhang et al., 2018; Zhou et al., 2019), their functional

antigen-binding fragment (Fab), and the single-chain variable region fragment (scFv), and a series of novel Nbs were first designed by grafting the complementarity-determining regions (CDRs) of nAbs targeting the two CoVs heavy chain onto a known nanobody framework (Kang et al., 2019). Then, the structures of the designed Nbs were predicted by comparative modeling, and the possibility of Nbs-RBD complexes formation was assessed through protein-protein docking followed by molecular dynamics (MD) simulations. Considering the structural changes of SARS-CoV-2 RBD during infection, multiple conformations of RBDs binding to ACE2 were extracted from the crystal structure (Lan et al., 2020) or microsecond level MD simulation of a chimeric RBD in complex with ACE2 (Shang, Ye, et al., 2020). Computational redesign of protein-protein interaction affinity and specificity was further applied to generate brand-new Nbs. Finally, the results suggest that eleven of the (re)designed Nbs may bind the SARS-CoV-2 RBD and potentially neutralize the virus.

## 2 | MATERIAL AND METHODS

### 2.1 | Designing Nbs through CDRs grafting

There are 16 nAbs binding to SARS-CoV, MERS-CoV, or SARS-CoV-2 RBD which the atomic coordinates deposited in RCSB PDB database (Berman et al., 2000) at the time of analysis are collected in Table S1. To design Nbs targeting SARS-CoV-2 RBD, only the RBD of three CoVs in complex with the “single-domain antibodies (sdAb)” (referred to the heavy chain of the corresponding nAbs) were saved for all of the 16 structures. It is known that each sdAb (Nb) has a highly conserved framework with three CDRs of variable sequence composing the paratope (Wilton et al., 2018). Therefore, the 16 prepared “sdAb” sequences originated from the heavy chain of SARS-CoV, MERS-CoV, or SARS-CoV-2 nAbs were further divided into seven parts, including four fragment regions (FRs) and three CDRs parts (Figure S1a). Then, computational sequence design of 16 Nbs (CS01 to CS16) was performed in two steps. First, choosing the four FRs of a camelid Nbs (Figure S1b) as a scaffold (Hamers-Casterman et al., 1993) and grafting the three CDRs of each prepared “sdAb” onto the scaffold (Wagner et al., 2018; Table S2). Second, to make sure the correct folding of designed Nbs, several amino acids of the camelid Nbs, especially the residues located on the edge of the four FRs, were replaced by the sequence of prepared “sdAb”.

### 2.2 | Predicting the designed Nbs structures

The structures of designed Nbs were predicted by comparative modeling method. The BLAST algorithm (Schaffer

et al., 2001) in NCBI was used to search templates. For each Nb, three different structures were selected as templates (Table S3) for modeling using default parameter in Modeller (v.9.24; Sali & Blundell, 1993) and 10 models were predicted. The model for designed Nb was selected by picking the structure with the best DOPE assessment score considering the Lennard–Jones potential and GBSA implicit solvent interaction (Shen & Sali, 2006).

### 2.3 | Protein–protein docking

The docking of designed Nbs to RBDs was performed using the protein docking module of Rosetta (v.2019.31; Alford et al., 2017). All of the structures calculated by Rosetta were scored base on the Rosetta Energy Function2015 (REF2015; Alford et al., 2017; Chaudhury et al., 2011). The coordinates of Nb and RBD structures were first formed through the script of *clean\_pdb.py* and refined by running the Rosetta relax protocol (Nivon et al., 2013). To guide protein–protein docking, the PyMOL software (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC) was applied to construct the structures of Nbs contacted with RBDs based on the 16 prepared crystal structures of “sdAb” in complex with SARS-CoV, SARS-CoV-2, or MERS-CoV. The details of the constructed Nbs-RBD complexes are listed in Figure S2. The RBDs in different conformational states were extracted from the crystal structure (Lan et al., 2020) or microsecond level MD simulation of a chimeric RBD in complex with ACE2 (Shang, Ye, et al., 2020). To ensure low-energy starting side-chain conformations for docking, prepacking of the constructed Nbs-RBD complexes was conducted by Rosetta prepack protocol (Wang et al., 2005). For each complex, the prepacked low-energy conformation was used as a starting point for several rounds protein–protein docking to generate 1,000 or 10,000 decoys by running Rosetta docking protocol (Gray et al., 2003) with the Monte Carlo (MC) refinement method (Gray et al., 2003). For each docking trajectory, the RMSD was calculated from the heavy atoms of the interface residues using each pose of the top five scorers as a reference structure (Chaudhury et al., 2011) using Rosetta InterfaceAnalyzerMover (Fleishman et al., 2011). The docking funnel was then identified through plotting  $I_{\text{score}}$  against  $I_{\text{rmsd}}$ . The top scoring structure with the lowest  $I_{\text{rmsd}}$  was selected as the successful docking conformation of the Nb-RBD complex. Finally, five out of the 16 designed Nbs were found to be able to bind on RBD. To verify the reliability of selecting designed Nbs-RBD complex's near-native conformation from free energy landscape of protein docking, three recently reported Nbs (SR4 (Li, Cai, et al., 2020), H11-D4 (Huo et al., 2020), VHH-72 (Wrapp et al., 2020)) occupying different binding sites of RBD (Figure S3a) were used to redocking study with the same setup as mentioned above.

### 2.4 | Molecular dynamics simulations

The starting coordinate used to run MD simulation was near-native conformation of Nbs-RBD extracted from Rosetta protein–protein docking trajectory. Before the MD simulation, AMBER ff14SB (Maier et al., 2015) force field parameters were assigned to two proteins, and missing atoms were added by applying LEaP (Case et al., 2005) to each complex. Then, an appropriate number of counterions was added to maintain the electro-neutrality of the studied system, and each complex was immersed into a rectangular periodic box of TIP3P (Hornak et al., 2006) water molecular with an edge of 10.0 Å. All molecular dynamics simulations were carried out by running a GPU-accelerated simulation (Case et al., 2005; Gotz et al., 2012) in the PMEMD module of AMBER14 software. First, two steps of energy minimization were performed, the whole system was minimized by a harmonic restraint with a  $10.0 \text{ kcal mol}^{-1} \text{ Å}^{-2}$  force constant in the NVT simulation followed by second minimization with no restraint. Next, the system was gradually heated to 100.0K in 2,500 steps and then heated to 310.0 K in 5,000 steps also with the force constant  $10.0 \text{ kcal mol}^{-1} \text{ Å}^{-2}$  to the complex. Subsequently, the equilibrate simulation at constant pressure (Case et al., 2014) with isotropic scaling simulation at 310.0 K in 500 ps. Finally, a 200 ns production run carried out without any restraint on all studied systems at a temperature of 310.0 K controlled by the Langevin thermostat with a collision frequency (Case et al., 2014) of  $1 \text{ ps}^{-1}$  and a pressure of 1 atm under the control of Monte Carlo barostat (Case et al., 2014; barostat = 2). During the whole simulation, bond lengths concluding hydrogens were constrained using SHAKE (Case et al., 2014), a 2 fs time step was set and electrostatics and Lennard-Jones cutoff radius of 10.0 Å (Case et al., 2014) was considered. The trajectory analysis was done with the CPPTRAJ (Roe & Cheatham, 2013) module of AMBER14 (Case et al., 2014) and PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC) software packages were used to visualize the structure extracted from the MD trajectory.

### 2.5 | Calculating the binding free energies

The binding free energy ( $\Delta G_{\text{tol}}$ ) for each complex was calculated by the molecular mechanics/generalized Born surface area (MM/GBSA) method as below equation:

$$\Delta G_{\text{tol}} = \Delta E_{\text{vdW}} + \Delta E_{\text{ele}} + \Delta G_{\text{pol}} + \Delta G_{\text{nonpol}} \quad (1)$$

where  $\Delta E_{\text{vdW}}$  and  $\Delta E_{\text{ele}}$  represent van der Waals and electrostatic energy changes in the gas phase.  $\Delta G_{\text{pol}}$  and  $\Delta G_{\text{nonpol}}$  represent polar and non-polar solvent interaction changes.  $\Delta E_{\text{vdW}}$  and  $\Delta E_{\text{ele}}$  were calculated using AMBER ff14SB (Maier

et al., 2015) in the gas phase. The polar part  $\Delta G_{\text{pol}}$  is calculated by generalized Born (GB) model. Dielectric constants of 8.0 and 80.0 are used for solute and solvent, respectively. The non-polar solvation contribution is computed using the solvent access surface area (SASA) with LCPO method in AMBER molsurf module, according to:

$$\Delta G_{\text{nonpol}} = \gamma \text{SASA} + \beta \quad (2)$$

where  $\gamma$  and  $\beta$  represent the surface tension and constant were set to 0.005 kcal/(mol·Å<sup>2</sup>) and 0 (Wang et al., 2019). SASA was used to calculate the non-polar component with a probe radius of 1.4 Å (Yang et al., 2011). To identify important residues critical for Nb binding to RBD, energy decomposition analysis was performed by the MM/GBSA method at a per-residue basis with the SASA obtained from icosahedron ICOSA method (Weiser et al., 1999).

## 2.6 | Redesigning the protein–protein interface

The protein–protein interface redesign was carried out by RosettaDesign (Mandell et al., 2009; Wang et al., 2007) in two steps. First, Rosetta fast relax (Nivon et al., 2013) was performed on the PDB files of Nb bound RBD complex taken from the MD simulations. The chain ID of the complex was deleted before running the relax protocol. Then, a resfile which specifies the residues need to be redesigned was used to run the fixbb protocol (Huang et al., 2011) in the Rosetta. For each Nb-RBD complex, sequence alignment shows that the interface redesign may generate the same Nb sequences, and Nbs with different sequences among the top 5 redesigned structures were selected for further MD simulations and binding free energies calculation analysis using the same setup as mentioned above.

## 3 | RESULTS AND DISCUSSION

### 3.1 | Sequence and structural diversity of the designed Nbs

A total of 16 Nbs were designed by grafting the three CDRs as well as the FRs edge of “sdAbs” (Figure S1a) onto the camelid Nb scaffold (Figure S1b). The disulfide bonds of the camelid Nb scaffold were reserved during Nbs design. The Clustal Omega (v.1.2.4; Madeira et al., 2019) was used to create the multiple sequence alignment of the designed Nbs (Figure 1a). The length and type analysis of the amino acid composing the CDRs indicated the sequence diversity of the designed Nbs. Structural models of the 16 Nbs were predicted by comparative modeling method

in Modeller (v.9.24; Sali & Blundell, 1993) with suitable templates. Statistics on the models are summarized in Table S4. Structural alignment of the Nbs models is shown in Figure 1b. As expected, the four FRs are almost the same while the major difference of the 16 Nbs models is in the three CDRs, especially the CDR3 showing significant structure diversity, which is highly consistent with their sequences (Robert & Gouet, 2014).

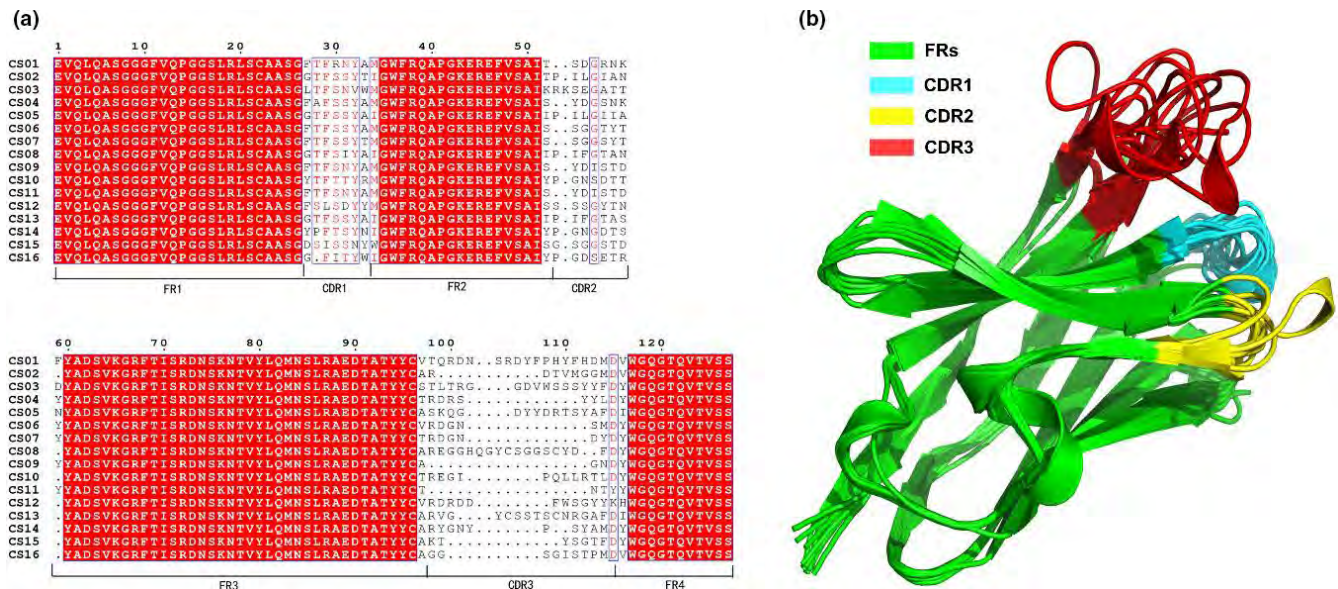
### 3.2 | Construction of the initial structures of Nbs-RBD complexes

The initial structures of Nbs in complex with RBD were constructed using the 16 prepared crystal structures of “sdAb” in complex with SARS-CoV, SARS-CoV-2, and MERS-CoV as references (Figure S2). Analysis of the constructed initial structures indicates Nbs contact with the different binding sites of RBD (Figure S4) which could be categorized into three groups by aligning 16 Nbs-RBD complex with ACE2-RBD complex. In group I, CS01, CS02, CS03, and CS04 occupy the ACE2 binding site of RBD (Figure S4a). In group II, CS05, CS06, CS07, CS08, CS10, CS12, CS13, and CS15 bind to the near-ACE2 binding site of RBD (Figure S4b). In group III, CS09, CS11, CS14, and CS16 bind to non-ACE2 binding site RBD (Figure S4c).

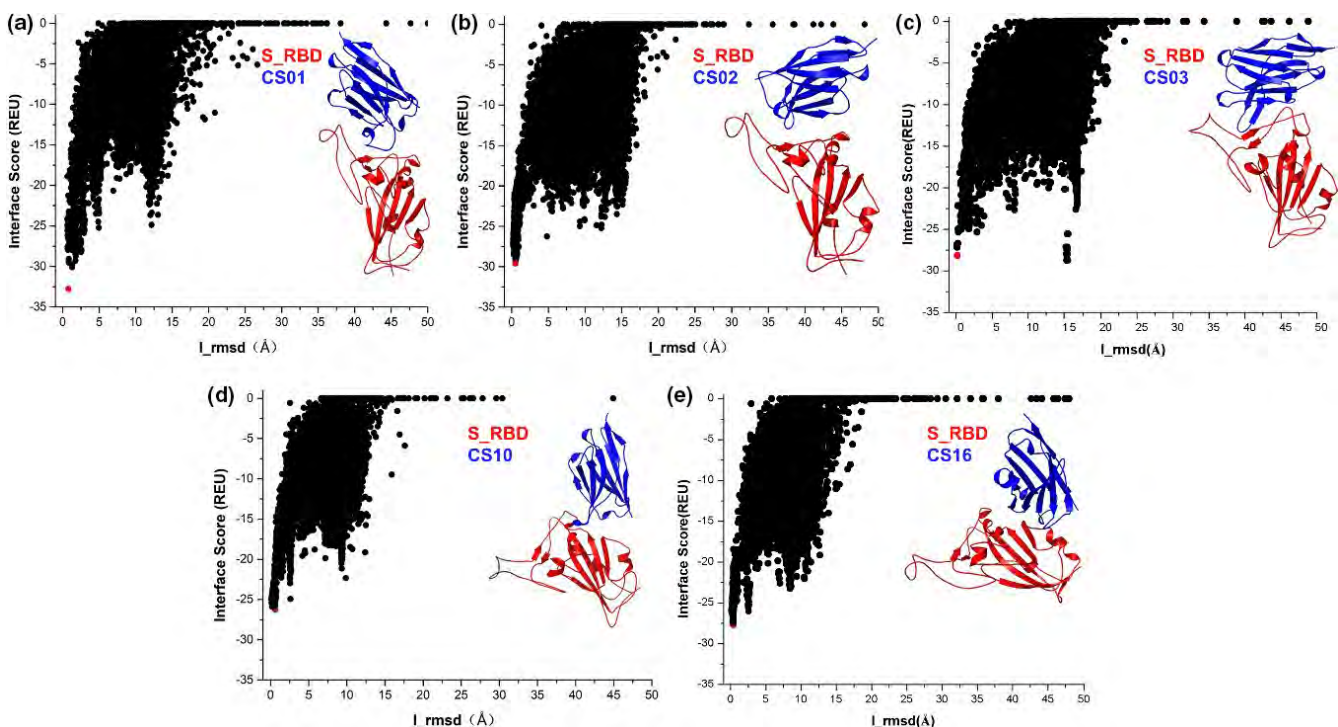
### 3.3 | The predicted near-native conformations of Nbs bind on RBD with different epitopes

Based on the constructed initial structures, Rosetta docking was used to search the near-native conformations of Nbs bound to SARS-CoV-2 RBD. Docking funnel of the trajectory describes the characteristics of the interface score (Rosetta energy units, REU) of each decoy and the interface root-mean-square deviation (I\_rmsd) with the initial structure to search for a near-native structure protein–protein binding complex (Lyskov et al., 2013). Therefore, docking funnel analysis was performed to evaluate the possibility of designed Nbs binding to RBD. As a result, the trajectories of CS01, CS02, CS03, CS10, and CS16 bind to RBD got the ideal docking funnels which contain lowest REU and I\_rmsd  $\leq 4$  Å (Chaudhury et al., 2011). For each complex, the structure with the lowest REU was extracted from the docking trajectory for further analysis (Figure 2). According to the classified information of constructed initial structures, CS01, CS02, and CS03 occupy the ACE2 binding site (Figure 2a–c), which may directly block the interaction between RBD and ACE2, and the interaction surface of RBD and nanobody is almost in loop regions of





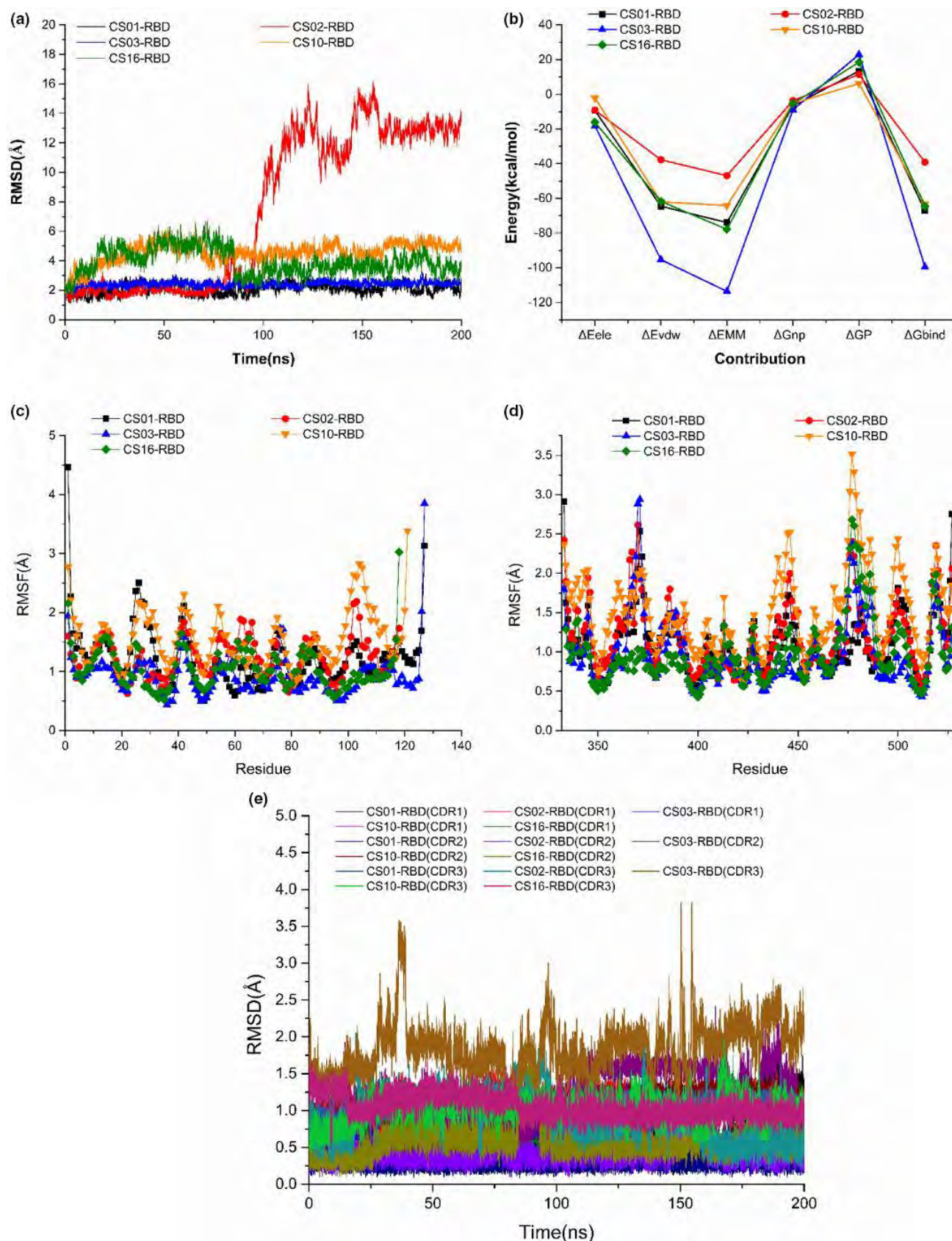
**FIGURE 1** Sequence and structural diversity of the 16 designed nanobodies (Nbs). (a) Multiple sequence alignment of the Nbs created by Clustal Omega (v.1.2.4) using default parameters. The red background residues are the scaffold of the Nbs, the residues in the blue box are the residues that share similarities, and the dot in the sequence means there is no residue at the corresponding position in the sequence. (b) Structure alignment of the Nbs predicted by comparative modeling method from Modeller (v.9.24). The protein structures are shown in cartoon representation. Green represents the four FRs regions; yellow, cyan, and red represent CDR1, CDR2, and CDR3 [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



**FIGURE 2** The Rosetta docking funnels of five Nbs (a–e): CS01, CS02, CS03, CS10, and CS16. The plots on red color have the lowest docking interface score with an I\_rmsd ≤ 4 Å. Inset: the lowest scoring conformation as near-native Nb-RBD complexes. There are complex\_05105 (CS01), complex\_09402 (CS02), complex\_09499 (CS03), complex\_02222 (CS10), and complex\_02898 (CS16) [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

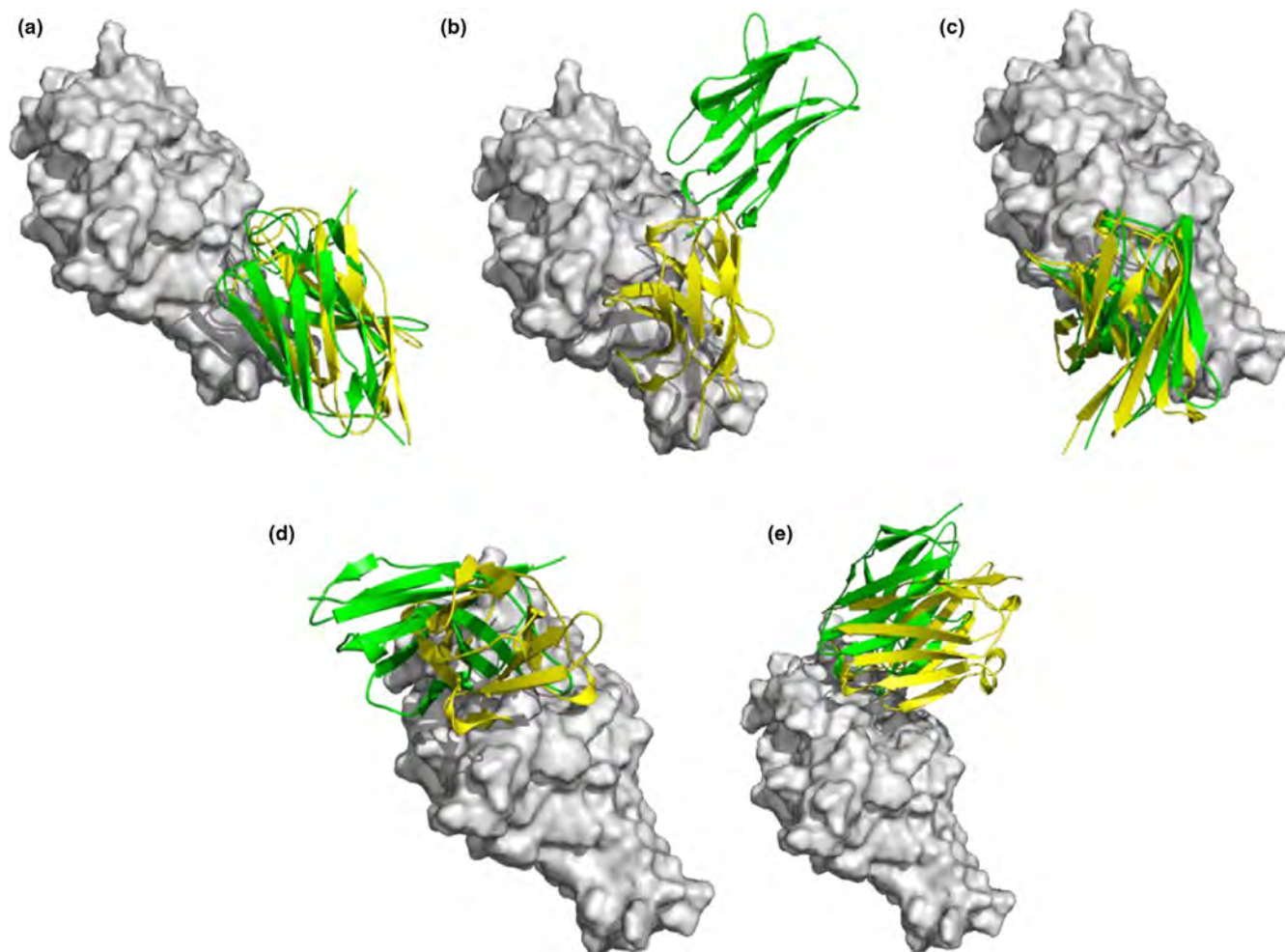
RBD, whereas CS10 and CS16 bind to the near-ACE2 and non-ACE2 binding sites which contain more helix and beta regions of RBD (Figure 2d–e), respectively, which may

inhibit the interaction between RBD and ACE2 through allosteric mechanism. The proposed inhibitory mechanism of design Nbs could be understood that MERS-CoV



**FIGURE 3** The stabilities and binding affinities for the formation of the five near-native Nb-RBD complexes. (a) Time evolution of the RMSD (Å) of the Nbs-RBD complex backbone Cα atoms for each snapshot relative to the initial structure during the MD simulations. (b) The binding free energies (kcal/mol) of the five complexes estimated from the snapshots of MD simulations. (c–d) The RMSF values of RBD and Nbs of five studied systems, respectively. (e) RMSD values of CDRs on five designed Nbs [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]





**FIGURE 4** Comparison of the Nbs-RBD binding site before (yellow) and after (green) MD simulation for designed Nbs, and white surface is SARS-CoV-2 RBD. (a–e) represent the complexes of CS01-RBD, CS02-RBD, CS03-RBD, CS10-RBD, and CS16-RBD, respectively [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

or SARS-CoV S protein adopted multiple conformation states and nAbs bind to the different epitopes of the antigen (Zavrtanik et al., 2018).

In addition, based on three recently reported cocrystal structures of Nbs-RBD complexes (PDB ID: 7C8V (Li, Cai, et al., 2020), 6YZ5 (Huo et al., 2020), and 6WAQ (Wrapp et al., 2020)), redocking was performed to verify the reliability of using free energy landscape to find the near-native conformation of protein docking in this study. From Figure S3a,b, the sequence of three Nbs SR4, HH-72, and VHH-72 is of great difference compared with designed stable Nbs (CS01, CS02, CS03, CS10, and CS16), and the structures of FRs are almost identical while the conformation of CDRs especially CDR3 is different. Interestingly, the three Nbs occupy the ACE2, near-ACE2, and non-ACE2 binding sites on RBD, similar to the five designed Nbs (Figures 2 and S3a). The results show that all the three Nbs bind to RBD got ideal docking funnels with a lowest REU and  $I_{\text{rmsd}} \leq 4$  Å and show similar trend with designed Nbs-RBD complexes (Figure S3b–d).

The selected near-native conformation of the three Nbs-RBD complexes protein docking is consistent well with their cocrystal structures, indicating that the selected near-native conformation of designed Nb-RBD complex from Rosetta docking is reliable and can be used for following analysis.

### 3.4 | The stabilities of the predicted near-native complexes

Starting from the near-native conformations of the five Nbs in complex with RBDs, MD simulations were performed to optimize their binding modes. The RMSD values of  $C\alpha$  atoms of each snapshot relative to the starting coordinates were calculated to monitor the stabilities of Nbs-RBD complexes. As shown in Figure 3a, all systems basically reached convergence after 160 ns simulation. The CS01-RBD, CS03-RBD, CS10-RBD, and CS16-RBD complexes have small fluctuations with RMSD values of 1.72, 1.17,

Complex	Contribution					
	$\Delta E_{\text{ele}}$	$\Delta E_{\text{vdW}}$	$\Delta E_{\text{MM}}$	$\Delta G_{\text{nonpol}}$	$\Delta G_{\text{pol}}$	$\Delta G_{\text{total}}$
CS01-RBD	-9.39	-64.55	-73.94	-6.58	13.44	-67.09
CS02-RBD	-9.17	-37.75	-46.92	-3.66	11.45	-39.13
CS03-RBD	-18.23	-95.39	-113.62	-8.89	22.93	-99.57
CS10-RBD	-2.12	-62.01	-64.14	-5.3	6.22	-63.22
CS16-RBD	-16.09	-61.75	-77.85	-5.33	18.52	-64.65

**TABLE 1** The calculated binding free energies (kcal/mol) of the five designed Nb-RBD complexes

1.93, and 3.11 Å, respectively, while CS02-RBD complex possesses a large fluctuation (from 1.85 to 8.94 Å) around 90 ns of the simulation. Compared with the initial conformation of CS02-RBD complex, it is found that CS02 associated from the initial site and bound to a new site on RBD after MD simulation (Figure 4b). For CDRs of the five Nbs, most of the calculated RMSD values range below 2.5 Å (Figure 3e). Among them, CDR3 region has larger fluctuation in complexes CS03-RBD and CS10-RBD, while CDR1 region has larger fluctuation in complexes CS01-RBD and CS02-RBD, and CDR2 region has larger fluctuation in complex CS16-RBD, indicating that those regions will accommodate to the conformation of RBD during the recognition process.

In addition, the averaged root-mean-square fluctuations (RMSFs) of each residue in Nbs and RBD protein during the last equilibrated 40 ns MD trajectories were calculated. In Figure 3c–d, all of the five complexes share similar trends of dynamics features. As expected, the fluctuation of  $\alpha$ -helix or  $\beta$ -sheet regions on RBD is smaller than loop region in Nbs-RBD complex. Interesting, for the loop region of RBD, the values of RMSF on the interface residues interacting with Nbs (CS01-RBD (472–502), CS02-RBD (400–425), CS03-RBD (442–454, 483–502), CS10-RBD (368–374, 436–441, 499–508), and CS16-RBD (368–382)) are almost smaller than 2 Å, suggesting that the loop region on interface of RBD is in stable state (Figure 3c). The peak RMSF values of RBD in five complexes occur at residues H519 (CS01), N370 (CS02), S371 (CS03), and S477 (CS10 and CS16) located at loop region, all of the residues are far away from the binding site of each complex. For Nbs, the fluctuation of RMSF values on CDR3 of CS01, CS03, and CS16 is more stable than other regions and all the RMSF values are around 1 Å (Figure 3d), while for CS02 the RMSF with lower values is concentrated in the regions near R38 on FR2 domain (Figure 3d), and residues focused on CDR2 have a relatively small RMSF value of CS10 (Figure 3d), indicating that CS01, CS03, and CS16 interact with RBD primarily through CDR3 domain. Apart from CDR3, the recognition of CS10 and RBD also depend on CDR2 in a degree. But for CS02, FR2 is the principal interacting domain which

can remind that CS02-RBD are not in a suitable binding state and compared to the other complexes, the relatively higher RMSF values of CS02 bound RBD implying the weak binding of the complex.

### 3.5 | The binding free energies of the Nbs-RBD complexes

We use the MM/GBSA method (Kollman et al., 2000), which has been widely employed to characterize the thermodynamics properties of different types of protein–protein interactions (PPIs; Du et al., 2020; Tu et al., 2018; Wang et al., 2019), to estimate the binding affinities between Nbs and RBD. Table 1 lists the total binding free energy ( $\Delta G_{\text{tol}}$ ) as well as their compositing terms including the electrostatic interaction energy ( $\Delta E_{\text{ele}}$ ), van der Waals interaction energy ( $\Delta E_{\text{vdW}}$ ), non-polar solvent energy ( $\Delta G_{\text{nonpol}}$ ), and polar solvent energy ( $\Delta G_{\text{pol}}$ ).

The calculated  $\Delta G_{\text{tol}}$  values of CS01, CS02, CS03, CS10, and CS16 binding to RBD are -67.09, -39.13, -99.57, -63.22, and -64.65 kcal/mol, respectively, which is probably due to the binding site of each Nb bound to the RBD is different as shown in Figure 2a–e. For the binding interface of Nbs-RBD, the number of key residues on the contact surface of Nbs as well as RBD recognized by MM/GBSA method can result in different calculated  $\Delta G_{\text{tol}}$  values. CS02-RBD complex contains four key residues on RBD interface and three key residues on Nb interface while CS03-RBD involves 10 key residues on RBD interface and 10 key residues on Nb interface, most likely the reason why the  $\Delta G_{\text{tol}}$  values between CS02-RBD and CS03-RBD make a huge difference. More information is shown in Table 2.

For all of the five complexes, the  $\Delta G_{\text{tol}}$  are contributed primary by the  $\Delta E_{\text{vdW}}$  and the  $\Delta E_{\text{ele}}$ . The  $\Delta G_{\text{nonpol}}$  contributed little to  $\Delta G_{\text{tol}}$ , while the  $\Delta G_{\text{polar}}$  seems to be unfavorable to the binding of the Nbs to RBD (Figure 3b). For CS03-RBD complex, the most favorable  $\Delta E_{\text{vdW}}$  contribution resulting in the highest binding affinities among the studied complexes. Although the CS02-RBD complex found a relative stable binding state (Figure 2b), the estimated value of  $\Delta G_{\text{tol}}$  further suggests that the binding interface between CS02 and RBD was not well designed.



**TABLE 2** The identified “hot spots” (energy contribution  $\leq -2$  kcal/mol) on the interface of (a) RBD and (b) the five designed Nbs

(a) RBD											
CS01-RBD		CS02-RBD		CS02_RD01-RBD		CS03-RBD		CS03_RD01-RBD		CS03_RD02-RBD	
Residue	Energy	Residue	Energy	Residue	Energy	Residue	Energy	Residue	Energy	Residue	Energy
N487	-3.86	T415	-2.67	F456	-4.61	Y449	-5.49	Y449	-5.94	Y449	-5.75
F486	-3.79	Q414	-2.62	Y489	-4.02	F490	-3.86	F490	-5.42	F490	-3.59
Y489	-3.08	Y380	-2.17	Y505	-2.89	N450	-3.1	Y489	-4.22	Y489	-2.86
F456	-2.48	R408	-2.15	K417	-2.8	Y505	-2.97	Q493	-3.57	Q498	-2.69
K417	-2.31			Q493	-2.28	E484	-2.88	Q498	-2.8	V445	-2.67
L455	-2.21			L455	-2.19	Q498	-2.84	L492	-2.66	L452	-2.64
A475	-2.04					Y489	-2.66	I468	-2.65	G446	-2.52
						T500	-2.07	V445	-2.17	E484	-2.38
						G446	-2.03	G446	-2.13	N450	-2.33
						V445	-2	E484	-2.05	F486	-2.17
										Y351	-2.05
CS03_RD03-RBD		CS03_RD04-RBD		CS10-RBD		CS16-RBD		CS16_RD01-RBD		CS16_RD02-RBD	
Residue	Energy	Residue	Energy	Residue	Energy	Residue	Energy	Residue	Energy	Residue	Energy
Y449	-5.71	449	-6.75	V503	-5.33	K378	-3.59	F377	-4.55	K378	-4.33
F490	-3.92	490	-3.41	N437	-3.49	F377	-3.53	T385	-4.37	F377	-3.83
F486	-2.86	452	-3.28	N440	-2.25	S383	-3.13	Y369	-3.99	Y380	-3.52
Q498	-2.86	486	-3.24	S373	-2.22	Y380	-3.05	K378	-2.44	R408	-2.93
Y489	-2.64	489	-2.8	N439	-2.2	T385	-2.52	F374	-2.39	S383	-2.52
V445	-2.31	498	-2.77			C379	-2.27			T376	-2.3
Q493	-2.3	487	-2.74			Y369	-2.18			C379	-2
G446	-2.22	445	-2.58								
C488	-2.15	446	-2.45								
Y505	-2.1	500	-2.34								
N487	-2.04	484	-2								
(b) Nbs											
CS01-RBD		CS02-RBD		CS02_RD01-RBD		CS03-RBD		CS03_RD01-RBD		CS03_RD02-RBD	
Residue	Energy	Residue	Energy	Residue	Energy	Residue	Energy	Residue	Energy	Residue	Energy
Y106	-4.75	S118	-3.8	L55	-4.07	E56	-4.03	T107	-5.66	Y110	-5.4
F107	-3.74	P41	-3.47	V101	-3.47	V107	-3.88	Q59	-4.27	Q56	-5.21
F59	-3.46	G42	-2.68	V54	-3.06	W108	-3.48	T110	-3.92	Y112	-4.49
N57	-2.61			D102	-2.86	D64	-3.22	V56	-3.9	Q47	-3.15
R56	-2.38			N59	-2.63	S110	-3.01	P64	-3.57	K111	-2.99
K65	-2.36			I57	-2.48	F47	-2.97	Y112	-3.35	P64	-2.97
Y60	-2.01			G103	-2.34	S55	-2.79	F47	-3.09	N55	-2.81
				F47	-2.26	Y112	-2.72	R108	-2.76	Q46	-2.64
						E46	-2.58	Q46	-2.62	T108	-2.42
						R103	-2.33	Y52	-2.51	N109	-2.41
								Y113	-2.35	P45	-2.01
								G57	-2.05		

(Continues)

TABLE 2 (Continued)

CS03_RD03-RBD		CS03_RD04-RBD		CS10-RBD		CS16-RBD		CS16_RD01-RBD		CS16_RD02-RBD	
Residue	Energy	Residue	Energy	Residue	Energy	Residue	Energy	Residue	Energy	Residue	Energy
R55	−4.31	M107	−4.45	I101	−4.41	I29	−4.35	T30	−5.16	I29	−4.43
V56	−3.98	Y110	−4.3	N55	−4.3	T30	−3.78	I29	−3.98	Y101	−4.18
Y110	−3.88	Q56	−3.89	Y52	−3.66	I101	−3.65	Y51	−3.44	Y51	−3.76
P64	−3.6	Q59	−3.53	P53	−3	S99	−3.16	Y31	−3.37	T30	−3.71
Q59	−2.98	D64	−3.26	G54	−2.37	T103	−2.44	S99	−3.15	S99	−2.57
F47	−2.75	K111	−3.1	T30	−2.29	G100	−2.02			E56	−2.1
Y112	−2.69	N109	−3.06								
T108	−2.66	Q65	−2.93								
Q46	−2.54	Q46	−2.87								
T107	−2.46	T108	−2.52								
Y113	−2.16	Y113	−2.36								
		E44	−2.19								
		F47	−2.13								

### 3.6 | The “Hot Spots” located at Nbs-RBD interface

To further understand the differentiated binding affinities of the designed Nbs binding to the different epitopes of the antigen, the important residues located at protein–protein interface were identified by decomposing the  $\Delta G_{\text{tot}}$  at per-residue basis. Herein the residues with a favorable energy contribution more than −2 kcal/mol were defined as “hot spots”. As shown in Table 2, various “hot spots” were discovered for the Nbs-RBDs complexes formation, and for more clearly, we identified “hot spots” on two parts, Table 2a,b represent the “hot spots” on the RBD interface and Nbs contact face of all studied systems, respectively. The number of “hot spots” on the RBD interface in CS01-RBD, CS02-RBD, CS03-RBD, CS10-RBD, CS16-RBD complexes are 7, 4, 10, 5, and 7, respectively, and the number of “hot spots” on the Nbs interface are 7, 3, 10, 6 and 6, consistent very well with the trends of their predicted binding affinities. Mapping the “hot spot” on the RBDs and Nbs interfaces is shown in Figure S5. Taking CS01-RBD complex as an example, N487 (−3.86 kcal/mol), F486 (−3.79 kcal/mol), Y489 (−3.08 kcal/mol), F456 (−2.48 kcal/mol), K417 (−2.31 kcal/mol), L455 (−2.21 kcal/mol) and A475 (−2.04 kcal/mol) form “hot spots” on the RBD interface, and Y106 (−4.75 kcal/mol), F107 (−3.74 kcal/mol), F59 (−3.46 kcal/mol), N57 (−2.61 kcal/mol), R56 (−2.38 kcal/mol), K65 (−2.36 kcal/mol), and Y60 (−2.01 kcal/mol) compose the “hot spots” of CS01.

All of the “hot spots” on RBDs interfaces of CS01, CS02, and CS10 bound complexes belong to loop region (Figure S5a–b,d), and more than half (4/7) “hot spots” on RBD

interface of CS16 (Figure S5e) bound complex located in loop region. The RBDs interfaces share the same “hot spot” Y489 in CS01 and CS03 bound complexes, demonstrating that Y489 of RBD will play an important role in interacting with neutralizing Nbs. Different from the CS01, CS02, and CS03 bound RBDs, the “hot spots” on the RBD of CS10 and CS16 bound RBDs interfaces represent the allosteric binding sites (Figure S5d–e).

The identified “hot spots” of the five designed Nbs are nearly located in the CDRs interfere with RBDs except for CS02 (Figure S5a–e). The most favorable residues of the five Nbs are Y106 (CS01, −4.75 kcal/mol, CDR3), S118 (CS02, −3.8 kcal/mol, FR4), E56 (CS03, −4.03 kcal/mol, CDR2), I101 (CS10, −4.41 kcal/mol, CDR3), and I29 (CS16, −4.35 kcal/mol, CDR1). The “hot spots” occupancies in CDR2 and CDR3 of CS01, CS03, CS10, and CS16 are 4/7, 7/10, 4/6, and 4/6, respectively. For CS02, all key residues are in FRs, which are consistent with the lower binding affinity (−39.13 kcal/mol) of CS02-RBD compared to the other complexes. Therefore, the residues on CDR2 and CDR3 regions especially CDR3 are of great importance for nanobody bound to RBD, enhancing interaction of CDR2 and CDR3 to RBD can do a great help to design Nbs.

In addition, “warm spots” contributing the binding free energies from −0.50 to −2.00 kcal/mol located at Nbs-RBD interfaces were also identified. As shown in Table S5, residue Q493 of RBD was identified as “warm spots” both in CS01 and CS03 bound complexes. It is interesting that residues L455 and F486 of RBD were identified as “hot spots” in CS01 bound complexes, while they were “warm spots” for CS03 binding. The result further verified the

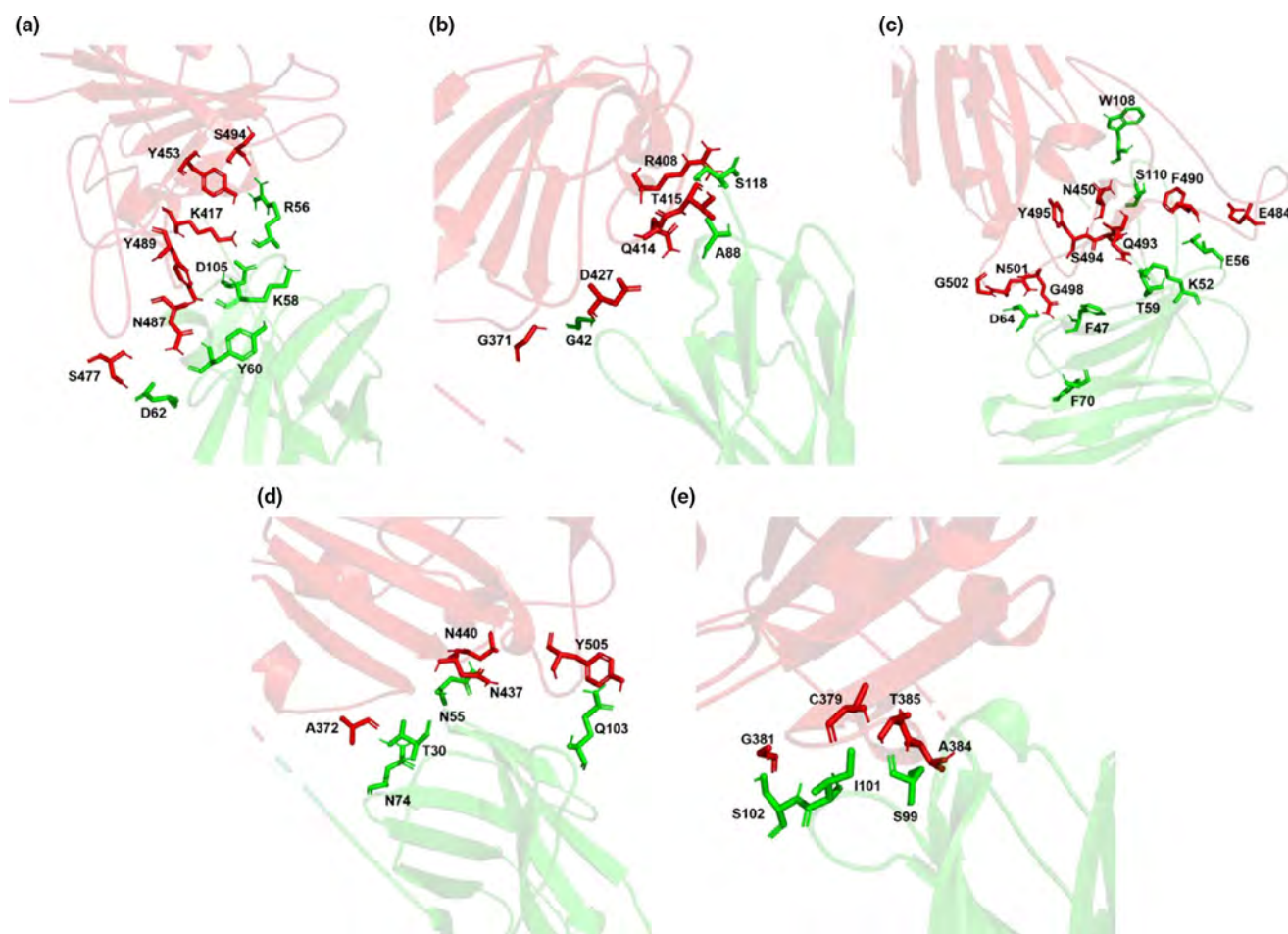
mechanism that CS01 and CS03 occupied the ACE2 binding site on RBD.

Figure 5 shows the interfaces of the representative structures extracted from the last equilibrated 40 ns MD trajectories of the five Nbs-RBD complexes. For clearly view, only “interacting residues” were displayed. It should be noted that residues involved in protein–protein interactions are nearly all “hot spots” and “warm spots” (Figure S5), proving that per-residue energy decomposition calculation is suitable for recognizing interactions between the Nbs and RBD interface.

Those information located at the Nbs-RBD interfaces not only imply that targeting loop region of RBD with different mechanisms could be an essential way for SARS-CoV-2 neutralizing Nbs design, but also provide useful information for better understanding the results that CS01 and CS03, CS10, and CS16 still bind at the ACE2, near-ACE2, and non-ACE2 binding sites of RBDs (Figure 4a,c–e), while CS02 found another binding site on RBD (Figure 4b) after MD simulations.

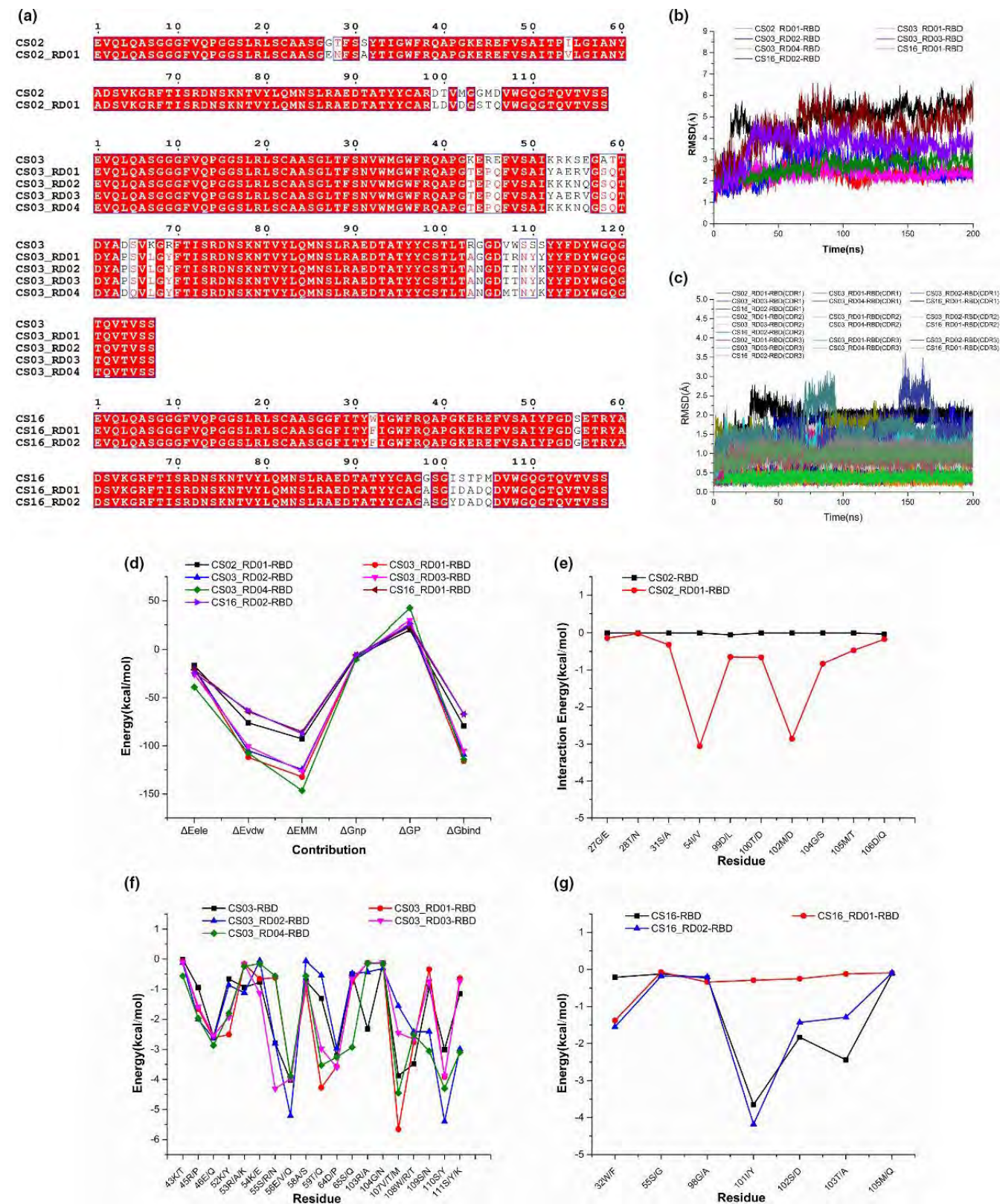
### 3.7 | The improved binding profiles of redesigned Nbs on RBD

Through grafting the CDRs of “sdAbs” onto a selected camelid Nb scaffold, five designed Nbs theoretically binding to SARS-CoV-2 RBD have been obtained by Rosetta docking and MD simulation evaluation. Here, RosettaDesign (Mandell et al., 2009; Wang et al., 2007) of the Nb interface residues was performed to enhance the Nbs-RBD interaction specificity while maintaining or improving the binding affinity (Kortemme et al., 2004; Potapov et al., 2008). Based on the representative conformations of CS01, CS02, CS03, CS10, and CS16 bound complexes, a number of brand-new sequences were generated. Different sequences of top five redesigned structures (Figure S6) were selected for further MD simulations and binding free energies calculations. Here we listed redesigned sequences bind to the RBD with more stability and improved binding affinities after MD simulations, one redesigned sequence of CS02 (CS02\_RD01), four redesigned sequences of CS03 (CS03\_RD01, CS03\_RD02, CS03\_RD03,

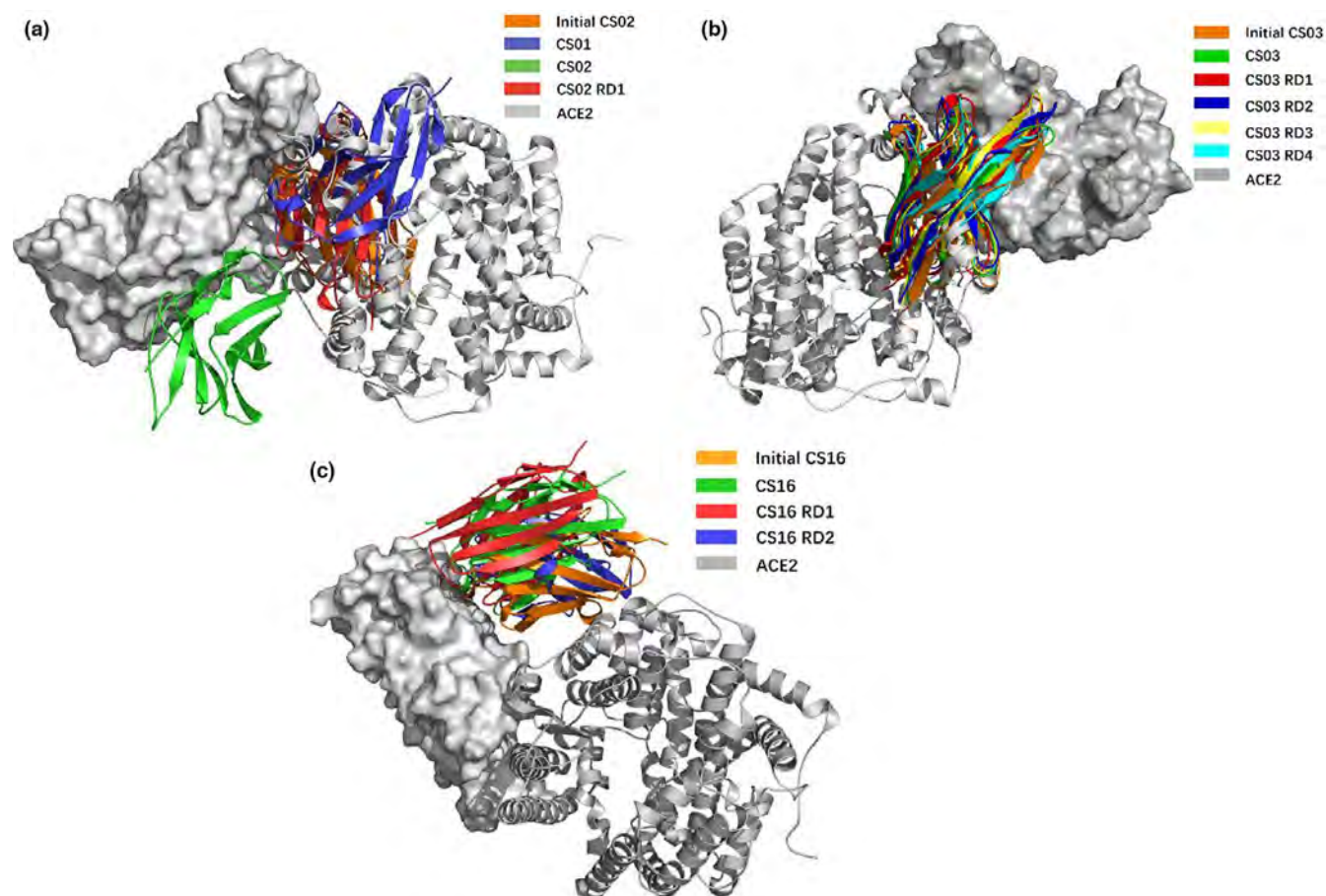


**FIGURE 5** View of Nbs (green) and RBD (red) complex interfaces with selected interacting residues in stick representation. The CS01-RBD, CS02-RBD, CS03-RBD, CS10-RBD, and CS16-RBD complexes are shown in (a–e), respectively [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]





**FIGURE 6** Comparison of binding profiles of Nbs on RBD before and after redesign. (a) Sequence alignment of wild-type and redesigned Nbs CS02, CS03, and CS16. The residues with red background are conservative residues, and the residues with empty background or red color are the parts that are mutated. (b) Time evolution of RMSD of the backbone Cα atoms for the redesigned Nbs-RBD complexes during MD simulations. (c) Time evolution of RMSD of the CDRs' backbone Cα atoms for the wild-type and redesigned Nbs-RBD complexes. (d) The calculated binding free energy changes of the redesigned Nbs-RBD complexes. (e–g) The per-residue energy contributions of the Nbs mutant positions before and after redesign of CS02, CS03, and CS16, respectively [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



**FIGURE 7** Comparison of binding sites between the initial structures and (re)designed Nbs of (a) CS02\_Nbs-RBD, (b) CS03\_Nbs-RBD, and (c) CS16\_Nbs-RBD complexes, respectively. CS01 shown in (a) demonstrates its similar binding site with CS02\_RD01. The RBD protein of SARS-CoV-2 is represented in white surface [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

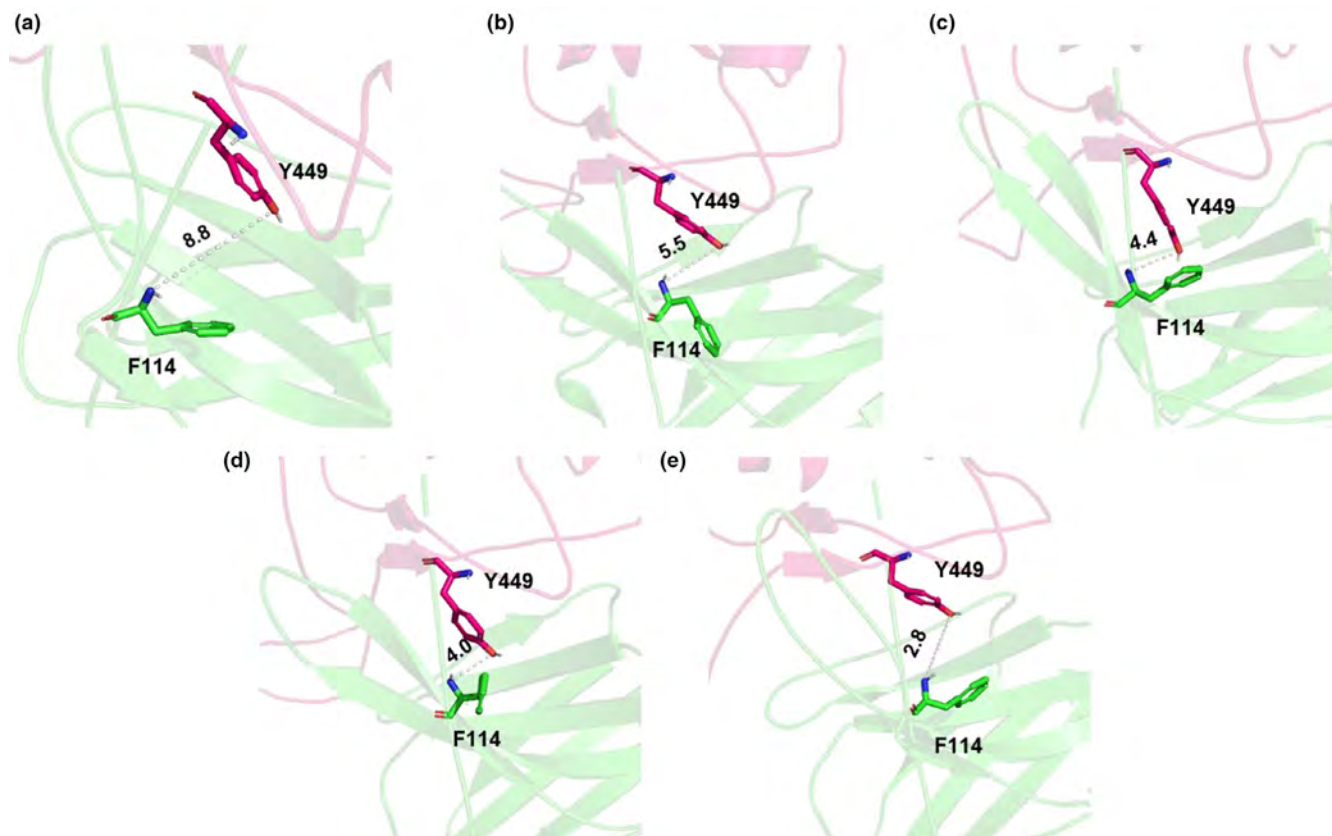
and CS03\_RD04), two redesigned sequence of CS16 (CS16\_RD01, CS16\_RD02; Figure 6a–c). Details of the calculated binding free energies are shown in Table S6. The  $\Delta E_{vdW}$  is still the main driving force for redesigned Nbs binding to RBDs followed by  $\Delta E_{ele}$ . Sequence alignments of the redesigned Nbs and their templates are displayed in Figure 6a with the mutations highlighted in white background. Figure 6b shows the  $\alpha$  RMSD of the seven redesigned Nbs-RBD complexes. CS02\_RD01-RBD and CS16\_RDs-RBD become more stable after redesign, and CS03\_RDs-RBD still maintain slight RMSD values fluctuation throughout the MD simulations.

Figure 6d indicates that the binding affinity of CS02\_RD01-RBD (−79.13 kcal/mol) is almost two times higher than that of CS02-RBD (−39.13 kcal/mol). Compared with the per-residue energy decomposition result of CS02\_RD01 and CS02 after MD simulation (Table 2), it can be assumed that the binding site of CS02\_RD01 on RBD has changed after redesign. This was further supported by several “hot spots” or “warm spots” on RBD interface shared by CS02\_RD01 and CS01, such as Y449, Y505, Y489, and Q493. Structural alignments show that the binding site of CS02\_RD01 on RBD is close to that of CS01 (Figure 7),

which can directly block the interaction between RBD and ACE2.

After redesign, the total number of “hot spots” on the Nb-RBD interfaces enlarged (Table 2) except for CS16\_RD01 bound complex. Table 2 also shows that some of the “hot spots” on the RBD (Table 2a) or Nb (Table 2b) interfaces contributing more energies for protein–protein interactions for 12 studied complexes. Compared with CS02-RBD and CS02\_RD01-RBD, the energy contribution of “hot spot” changed from −2.67 kcal/mol (Q415) to −4.61 kcal/mol (F456) on the interface of RBD, the binding site of CS02\_RD01 on RBD has changed a lot from CS02 (Figure 7a) closing to that of CS01. The residue Y449 of RBD is important for the binding of CS03. In all CS03\_RDs-RBDs, the energy contribution of Y449 was improved and with CS03\_RD04 increased by −1.26 kcal/mol, by aligning CS03-RBD and CS03\_RDs-RBD on the basis of  $\alpha$  backbone of RBD. Figure 8 shows the orientation of benzene ring on Y449 of RBD getting closer to Nb through monitoring the atom between nitrogen atom on residue F114 of Nb and oxygen atom on residue Y449 of RBD. In particular, the change in distance from 8.8 Å in CS03-RBD complex to 2.8 Å in CS03\_RD04-RBD





**FIGURE 8** Comparison of distance between oxygen atom on residue Y449 of RBD (red) and nitrogen atom on residue F114 of CS03\_Nbs (green). (a–e) represent the complex of CS03-RBD, CS03\_RD01-RBD, CS03\_RD02-RBD, CS03\_RD03-RBD, and CS03\_RD04-RBD, respectively [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

complex suggests that redesign method can make tight contact and improve the binding affinity of Nb-RBD complex. The “hot spot” with largest energy contribution on RBD was changed from Y378 (−3.59 kcal/mol) to C377 (−4.55 kcal/mol) and Y378 (−4.33 kcal/mol) in CS16, CS16\_RD01, and CS16\_RD02, respectively, and this little difference may be due to the Nb conformation readjusted after CS16 redesign (Figure 7c).

Similar to RBD, “hot spots” on the interface of CS02\_RD01 vary greatly. Figure 6e shows that the mutations I54V and M102D lead to the energy contributions from 0 and 0 kcal/mol (CS02) to −3.06 and −2.86 kcal/mol (CS02\_RD01), and these two major changes happened to CS02 reveal the information of the binding epitope of CS02 to RBD has reorganized after redesign and residues at position 54 and 102 deserve more attention. After the redesign, residue Y51 on CS16\_RDs emerged and made a difference in the binding of CS16\_RDs and RBD than CS16, and the energy of Y51 both is greater than −3.0 kcal/mol indicating the importance of Y51 means to CS16\_RDs. Figure 6g demonstrates mutation energy change of CS16 Nbs, and mutations W32F (CS16\_RD01, CS16\_RD02) and I101Y (CS16\_RD02) are beneficial for CS16 binding.

In all four CS03\_RDs, the increased energy contributions of mutations R45P, T59Q, and S110Y (Figure 6f) suggested that the mutant residues may be more suitable to recognize RBD. Apart from that, mutations E46Q, E56V(Q), D64P, and W108T(R) still have great energy values higher than −2 kcal/mol, the residues on the position of 46, 56, 64, and 108 are critical to CS03 Nbs, and those mutations are of benefit to CS03 binding. Mutations K52Y (CS03\_RD01, CS03\_RD03), S55N (CS03\_RD02), S55R (CS03\_RD03), S65Q (CS03\_RD04), V107M (CS03\_RD04), and S109N (CS03\_RD02, CS03\_RD04) bring improvement in energy contributions. It should be noted that the energy contributions of the same mutation (S55N, S109N, or S119K) in different redesigned complexes of CS03\_RDs show opposite results, such as the energy of mutation S55R gets lower for CS03\_RD03 but higher for CS03\_RD01 after redesign. Interestingly, in the series of CS03\_RDs, the values of the residue's largest energy were changed, but the number of “hot spots” and “warm hot spots” was hardly to be changed. This could be explained by the more conserved binding mode between CS03\_Nbs and RBD before and after redesign (Figure 7b), formed by residue from CDR2 and CDR3 of CS03\_Nbs and residues V113-L120, E152-F158, and F165-T168 from RBD.



## 4 | CONCLUSION

In this work, 16 Nbs (CS01-CS16) were designed by grafting CDRs sequence of CoVs' nAbs or mAbs into a known stable nanobody scaffold. And five out of the 16 Nbs (CS01, CS02, CS03, CS10, and CS16) were successfully docked onto SARS-CoV-2 RBD with different binding mechanisms. Among them, CS01, CS02, and CS03 occupy the ACE2 binding site, while CS10 and CS16 bind to the near-ACE2 and non-ACE2 binding sites, respectively. Further MD simulation and binding free energy analysis indicated that, except for CS02, the other four designed Nbs stably bind to the RBD of SARS-CoV-2. Based on the structures of the five Nbs in complex with RBD and seven brand-new Nbs with improved stabilities and binding affinities were generated by redesign of residues on the interface of the five Nbs contact with RBD. Especially for redesigned CS02-RBD complex, the calculated binding free energy increased from  $-39.13$  kcal/mol (CS02-RBD) to  $-79.13$  kcal/mol (CS02\_RD01-RBD). As a result, 11 (re)designed Nbs (CS01, CS02\_RD01, CS03, CS03\_RD01, CS03\_RD02, CS03\_RD03, CS03\_RD04, CS10, CS16, CS16\_RD01, CS16\_RD02) show stabilities and high binding affinities to SARS-CoV-2 RBD. Moreover, per-residue binding free energy decomposition analysis was performed to identify the "hot spots" between the Nb and RBD interface. For example, residue Y449 of RBD is important for CS03\_Nbs-RBD bound complex, and mutations R45P, T59Q, and S110Y of CS03 can do benefit for the binding of CS03\_Nbs and RBD. The identified "hot spots" provided useful information for us to understand the binding mechanism of designed Nbs to SARS-CoV-2 RBD.

## ACKNOWLEDGMENTS

This work was supported by COVID-19 Research Grant of Chongqing Municipal Education Commission (KYYG202002), Fundamental Research Funds for Central Universities (2019CDYGYB005), and Technology Innovation and Application Demonstration Project of Chongqing (cstc2018jscx-msybX0287).

## CONFLICT OF INTEREST

The authors declare no competing financial interests.

## AUTHOR CONTRIBUTIONS

W.X. and F.Z. designed the research. J.Y. and Z.Z. performed the research. J.Y., Z.Z., F.Y., H.Z., H.W., and W.X. analyzed the data. J.Y., Z.Z., F.Z., and W.X. wrote the manuscript. All authors reviewed the manuscript.

## DATA AVAILABILITY STATEMENT

The data that supports the findings of this study are available in the supplementary material of this article

## ORCID

Feng Zhu <http://orcid.org/0000-0001-8069-0053>

Weiwei Xue  <https://orcid.org/0000-0002-3285-0574>

## REFERENCES

- Alford, R. F., Leaver-Fay, A., Jeliazkov, J. R., O'Meara, M. J., DiMaio, F. P., Park, H., Shapovalov, M. V., Renfrew, P. D., Mulligan, V. K., Kappel, K., Labonte, J. W., Pacella, M. S., Bonneau, R., Bradley, P., Dunbrack, R. L., Das, R., Baker, D., Kuhlman, B., Kortemme, T., & Gray, J. J. (2017). The Rosetta all-atom energy function for macromolecular modeling and design. *Journal of Chemical Theory and Computation*, 13(6), 3031–3048. <https://doi.org/10.1021/acs.jctc.7b00125>
- Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., & Bourne, P. E. (2000). The protein data bank. *Nucleic Acids Research*, 28(1), 235–242. <https://doi.org/10.1093/nar/28.1.235>
- Case, D. A., V. B., Berryman, J. T., Betz, R. M., Cai, Q., Cerutti, D. S., Cheatham, T. E. III, Darden, T. A., Duke, R. E., H. G., Goetz, A. W., Gusarov, S., Homeyer, N., Janowski, P., Kaus, J., Kolossváry, I., Kovalenko, A., Lee, T. S. S. L., Luchko, T., ... Kollman, P. A. (2014). *AMBER 14*. University of California, San Francisco.
- Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M., Onufriev, A., Simmerling, C., Wang, B., & Woods, R. J. (2005). The Amber biomolecular simulation programs. *Journal of Computational Chemistry*, 26(16), 1668–1688. <https://doi.org/10.1002/jcc.20290>
- Chaudhury, S., Berrondo, M., Weitzner, B. D., Muthu, P., Bergman, H., & Gray, J. J. (2011). Benchmarking and analysis of protein docking performance in Rosetta v3.2. *PLoS One*, 6(8), e22477. <https://doi.org/10.1371/journal.pone.0022477>
- Chen, Z., Bao, L., Chen, C., Zou, T., Xue, Y., Li, F., Lv, Q. I., Gu, S., Gao, X., Cui, S., Wang, J., Qin, C., & Jin, Q. I. (2017). Human neutralizing monoclonal antibody inhibition of middle east respiratory syndrome coronavirus replication in the common marmoset. *Journal of Infectious Diseases*, 215(12), 1807–1815. <https://doi.org/10.1093/infdis/jix209>
- Du, Q., Qian, Y., & Xue, W. (2020). Molecular simulation of oncostatin M and receptor (OSM-OSMR) interaction as a potential therapeutic target for inflammatory bowel disease. *Frontiers in Molecular Biosciences*, 7, 29. <https://doi.org/10.3389/fmolb.2020.00029>
- Fleishman, S. J., Leaver-Fay, A., Corn, J. E., Strauch, E.-M., Khare, S. D., Koga, N., Ashworth, J., Murphy, P., Richter, F., Lemmon, G., Meiler, J., & Baker, D. (2011). RosettaScripts: A scripting language interface to the Rosetta macromolecular modeling suite. *PLoS One*, 6(6), e20161. <https://doi.org/10.1371/journal.pone.0020161>
- Gotz, A. W., Williamson, M. J., Xu, D., Poole, D., Le Grand, S., & Walker, R. C. (2012). Routine microsecond molecular dynamics simulations with AMBER on GPUs. 1. Generalized born. *Journal of Chemical Theory and Computation*, 8(5), 1542–1555. <https://doi.org/10.1021/ct200909j>
- Gray, J. J., Moughon, S., Wang, C., Schueler-Furman, O., Kuhlman, B., Rohl, C. A., & Baker, D. (2003). Protein–protein docking with simultaneous optimization of rigid-body displacement and side-chain conformations. *Journal of Molecular Biology*, 331(1), 281–299. [https://doi.org/10.1016/s0022-2836\(03\)00670-3](https://doi.org/10.1016/s0022-2836(03)00670-3)
- Hamers-Casterman, C., Atarhouch, T., Muyldermans, S., Robinson, G., Hammers, C., Songa, E. B., Bendahman, N., & Hammers, R. (1993).

- Naturally occurring antibodies devoid of light chains. *Nature*, 363(6428), 446–448. <https://doi.org/10.1038/363446a0>
- Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., Schiergens, T. S., Herrler, G., Wu, N.-H., Nitsche, A., Müller, M. A., Drosten, C., & Pöhlmann, S. (2020). SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell*, 181(2), 271–280 e278. <https://doi.org/10.1016/j.cell.2020.02.052>
- Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., & Simmerling, C. (2006). Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins*, 65(3), 712–725. <https://doi.org/10.1002/prot.21123>
- Huang, P. S., Ban, Y. E., Richter, F., Andre, I., Vernon, R., Schief, W. R., & Baker, D. (2011). RosettaRemodel: A generalized framework for flexible backbone protein design. *PLoS One*, 6(8), e24109. <https://doi.org/10.1371/journal.pone.0024109>
- Huo, J., Le Bas, A., Ruza, R. R., Duyvesteyn, H. M. E., Mikolajek, H., Malinauskas, T., Tan, T. K., Rijal, P., Dumoux, M., Ward, P. N., Ren, J., Zhou, D., Harrison, P. J., Weckener, M., Clare, D. K., Vogirala, V. K., Radecke, J., Moynié, L., Zhao, Y., ... Naismith, J. H. (2020). Neutralizing nanobodies bind SARS-CoV-2 spike RBD and block interaction with ACE2. *Nature Structural & Molecular Biology*, 27(9), 846–854. <https://doi.org/10.1038/s41594-020-0469-6>
- Hwang, W. C., Lin, Y., Santelli, E., Sui, J., Jaroszewski, L., Stec, B., Farzan, M., Marasco, W. A., & Liddington, R. C. (2006). Structural basis of neutralization by a human anti-severe acute respiratory syndrome spike protein antibody, 80R. *Journal of Biological Chemistry*, 281(45), 34610–34616. <https://doi.org/10.1074/jbc.M603275200>
- Jiang, S., Hillyer, C., & Du, L. (2020). Neutralizing ANTIBODIES against SARS-CoV-2 and other human coronaviruses. *Trends in Immunology*, 41(5), 355–359. <https://doi.org/10.1016/j.it.2020.03.007>
- Kang, S., Davidsen, K., Gomez-Castillo, L., Jiang, H., Fu, X., Li, Z., Liang, Y. U., Jahn, M., Moussa, M., DiMaio, F., & Gu, L. (2019). COMBINES-CID: An efficient method for de novo engineering of highly specific chemically induced protein dimerization systems. *Journal of the American Chemical Society*, 141(28), 10948–10952. <https://doi.org/10.1021/jacs.9b03522>
- Kollman, P. A., Massova, I., Reyes, C., Kuhn, B., Huo, S., Chong, L., & Cheatham, T. E. III. (2000). Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Accounts of Chemical Research*, 33(12), 889–897. <https://doi.org/10.1021/ar000033j>
- Kortemme, T., Joachimiak, L. A., Bullock, A. N., Schuler, A. D., Stoddard, B. L., & Baker, D. (2004). Computational redesign of protein-protein interaction specificity. *Nature Structural & Molecular Biology*, 11(4), 371–379. <https://doi.org/10.1038/nsmb749>
- Lan, J., Ge, J., Yu, J., Shan, S., Zhou, H., Fan, S., Zhang, Q. I., Shi, X., Wang, Q., Zhang, L., & Wang, X. (2020). Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature*, 581(7807), 215–220. <https://doi.org/10.1038/s41586-020-2180-5>
- Li, H., Zhou, Y., Zhang, M., Wang, H., Zhao, Q., & Liu, J. (2020). Updated Approaches against SARS-CoV-2. *Antimicrobial Agents and Chemotherapy*, 64(6), e00483–20. <https://doi.org/10.1128/AAC.00483-20>
- Li, T., Cai, H., Yao, H., Zhou, B., Zhang, N., Gong, Y., Zhao, Y., Shen, Q., Qin, W., Hutter, C. A. J., Lai, Y., Kuo, S. M., Bao, J., Lan, J., Seeger, M. A., Wong, G., Bi, Y., Lavillette, D., & Li, D. (2020). A potent synthetic nanobody targets RBD and protects mice from SARS-CoV-2 infection. *BioRxiv*, 143438–v2. <https://www.biorxiv.org/content/10.1101/2020.06.09.143438v2>
- Li, Y., Wan, Y., Liu, P., Zhao, J., Lu, G., Qi, J., Wang, Q., Lu, X., Wu, Y., Liu, W., Zhang, B., Yuen, K.-Y., Perlman, S., Gao, G. F., & Yan, J. (2015). A humanized neutralizing antibody against MERS-CoV targeting the receptor-binding domain of the spike protein. *Cell Research*, 25(11), 1237–1249. <https://doi.org/10.1038/cr.2015.113>
- Lyskov, S., Chou, F.-C., Conchúir, S. Ó., Der, B. S., Drew, K., Kuroda, D., Xu, J., Weitzner, B. D., Renfrew, P. D., Sripakdeevong, P., Borgo, B., Havranek, J. J., Kuhlman, B., Kortemme, T., Bonneau, R., Gray, J. J., & Das, R. (2013). Serverification of molecular modeling applications: The Rosetta Online Server that Includes Everyone (ROSIE). *PLoS One*, 8(5), e63906. <https://doi.org/10.1371/journal.pone.0063906>
- Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A. R. N., Potter, S. C., Finn, R. D., & Lopez, R. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Research*, 47(W1), W636–W641. <https://doi.org/10.1093/nar/gkz268>
- Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., & Simmerling, C. (2015). ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *Journal of Chemical Theory and Computation*, 11(8), 3696–3713. <https://doi.org/10.1021/acs.jctc.5b00255>
- Mandell, D. J., Coutsiar, E. A., & Kortemme, T. (2009). Sub-angstrom accuracy in protein loop reconstruction by robotics-inspired conformational sampling. *Nature Methods*, 6(8), 551–552. <https://doi.org/10.1038/nmeth0809-551>
- Muyldermans, S. (2013). Nanobodies: Natural single-domain antibodies. *Annual Review of Biochemistry*, 82, 775–797. <https://doi.org/10.1146/annurev-biochem-063011-092449>
- Nivon, L. G., Moretti, R., & Baker, D. (2013). A Pareto-optimal refinement method for protein design scaffolds. *PLoS One*, 8(4), e59004. <https://doi.org/10.1371/journal.pone.0059004>
- Pak, J. E., Sharon, C., Satkunarajah, M., Auperin, T. C., Cameron, C. M., Kelvin, D. J., Seetharaman, J., Cochrane, A., Plummer, F. A., Berry, J. D., & Rini, J. M. (2009). Structural insights into immune recognition of the severe acute respiratory syndrome coronavirus S protein receptor binding domain. *Journal of Molecular Biology*, 388(4), 815–823. <https://doi.org/10.1016/j.jmb.2009.03.042>
- Potapov, V., Reichmann, D., Abramovich, R., Filchtinski, D., Zohar, N., Ben Halevy, D., Edelman, M., Sobolev, V., & Schreiber, G. (2008). Computational redesign of a protein-protein interface for high affinity and binding specificity using modular architecture and naturally occurring template fragments. *Journal of Molecular Biology*, 384(1), 109–119. <https://doi.org/10.1016/j.jmb.2008.08.078>
- Prabakaran, P., Gan, J., Feng, Y., Zhu, Z., Choudhry, V., Xiao, X., Ji, X., & Dimitrov, D. S. (2006). Structure of severe acute respiratory syndrome coronavirus receptor-binding domain complexed with neutralizing antibody. *Journal of Biological Chemistry*, 281(23), 15829–15836. <https://doi.org/10.1074/jbc.M600697200>
- Robert, X., & Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research*, 42(W1), W320–W324. <https://doi.org/10.1093/nar/gku316>
- Roe, D. R., & Cheatham, T. E. III. (2013). PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics

- trajectory data. *Journal of Chemical Theory and Computation*, 9(7), 3084–3095. <https://doi.org/10.1021/ct400341p>
- Sali, A., & Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *Journal of Molecular Biology*, 234(3), 779–815. <https://doi.org/10.1006/jmbi.1993.1626>
- Schaffer, A. A., Aravind, L., Madden, T. L., Shavirin, S., Spouge, J. L., Wolf, Y. I., & Altschul, S. F. (2001). Improving the accuracy of PSI-BLAST protein database searches with composition-based statistics and other refinements. *Nucleic Acids Research*, 29(14), 2994–3005. <https://doi.org/10.1093/nar/29.14.2994>
- Shang, J., Wan, Y., Luo, C., Ye, G., Geng, Q., Auerbach, A., & Li, F. (2020). Cell entry mechanisms of SARS-CoV-2. *Proceedings of the National Academy of Sciences USA*, 117(21), 11727–11734. <https://doi.org/10.1073/pnas.2003138117>
- Shang, J., Ye, G., Shi, K. E., Wan, Y., Luo, C., Aihara, H., Geng, Q., Auerbach, A., & Li, F. (2020). Structural basis of receptor recognition by SARS-CoV-2. *Nature*, 581(7807), 221–224. <https://doi.org/10.1038/s41586-020-2179-y>
- Shen, M. Y., & Sali, A. (2006). Statistical potential for assessment and prediction of protein structures. *Protein Science*, 15(11), 2507–2524. <https://doi.org/10.1110/ps.062416606>
- Steeland, S., Vandenbroucke, R. E., & Libert, C. (2016). Nanobodies as therapeutics: Big opportunities for small antibodies. *Drug Discov Today*, 21(7), 1076–1113. <https://doi.org/10.1016/j.drudis.2016.04.003>
- Tu, G., Fu, T., Yang, F., Yao, L., Xue, W., & Zhu, F. (2018). Prediction of GluN2B-CT1290-1310/DAPK1 Interaction by protein(-)peptide docking and molecular dynamics simulation. *Molecules*, 23(11), 1290–1310. <https://doi.org/10.3390/molecules23113018>
- Wagner, H. J., Wehrle, S., Weiss, E., Cavallari, M., & Weber, W. (2018). A two-step approach for the design and generation of nanobodies. *International Journal of Molecular Sciences*, 19(11), 3444. <https://doi.org/10.3390/ijms19113444>
- Walls, A. C., Xiong, X., Park, Y.-J., Tortorici, M. A., Snijder, J., Quispe, J., Cameroni, E., Gopal, R., Dai, M., Lanzavecchia, A., Zambon, M., Rey, F. A., Corti, D., & Velesler, D. (2019). Unexpected receptor functional mimicry elucidates activation of coronavirus fusion. *Cell*, 176(5), 1026–1039 e1015. <https://doi.org/10.1016/j.cell.2018.12.028>
- Wang, C., Bradley, P., & Baker, D. (2007). Protein-protein docking with backbone flexibility. *Journal of Molecular Biology*, 373(2), 503–519. <https://doi.org/10.1016/j.jmb.2007.07.050>
- Wang, C., Horby, P. W., Hayden, F. G., & Gao, G. F. (2020). A novel coronavirus outbreak of global health concern. *The Lancet*, 395(10223), 470–473. [https://doi.org/10.1016/S0140-6736\(20\)30185-9](https://doi.org/10.1016/S0140-6736(20)30185-9)
- Wang, C., Schueler-Furman, O., & Baker, D. (2005). Improved side-chain modeling for protein-protein docking. *Protein Science*, 14(5), 1328–1339. <https://doi.org/10.1110/ps.041222905>
- Wang, E., Sun, H., Wang, J., Wang, Z., Liu, H., Zhang, J. Z. H., & Hou, T. (2019). End-point binding free energy calculation with MM/PBSA and MM/GBSA: Strategies and applications in drug design. *Chemical Reviews*, 119(16), 9478–9508. <https://doi.org/10.1021/acs.chemrev.9b00055>
- Wang, L., Shi, W., Chappell, J. D., Joyce, M. G., Zhang, Y. I., Kanekiyo, M., Becker, M. M., van Doremalen, N., Fischer, R., Wang, N., Corbett, K. S., Choe, M., Mason, R. D., Van Galen, J. G., Zhou, T., Saunders, K. O., Tatti, K. M., Haynes, L. M., Kwong, P. D., ... Graham, B. S. (2018). Importance of neutralizing monoclonal antibodies targeting multiple antigenic sites on the middle east respiratory syndrome coronavirus spike glycoprotein to avoid neutralization escape. *Journal of Virology*, 92(10), e02002–17. <https://doi.org/10.1128/JVI.02002-17>
- Wang, L., Shi, W., Joyce, M. G., Modjarrad, K., Zhang, Y. I., Leung, K., Lees, C. R., Zhou, T., Yassine, H. M., Kanekiyo, M., Yang, Z.-Y., Chen, X., Becker, M. M., Freeman, M., Vogel, L., Johnson, J. C., Olinger, G., Todd, J. P., Bagci, U., ... Graham, B. S. (2015). Evaluation of candidate vaccine approaches for MERS-CoV. *Nature Communications*, 6, 7712. <https://doi.org/10.1038/ncomms8712>
- Weiser, J., Shenkin, P. S., & Still, W. C. (1999). Approximate atomic surfaces from linear combinations of pairwise overlaps (LCPO). *Journal of Computational Chemistry*, 20(2), 217–230. [https://doi.org/10.1002/\(sici\)1096-987x\(19990130\)20:2<217:aid-jcc4>3.0.co;2-a](https://doi.org/10.1002/(sici)1096-987x(19990130)20:2<217:aid-jcc4>3.0.co;2-a)
- Wilton, E. E., Opyr, M. P., Kailasam, S., Kothe, R. F., & Wieden, H. J. (2018). sdAb-DB: The single domain antibody database. *ACS Synthetic Biology*, 7(11), 2480–2484. <https://doi.org/10.1021/acssynbio.8b00407>
- Wrapp, D., De Vlieger, D., Corbett, K. S., Torres, G. M., Wang, N., Van Breedam, W., Roose, K., van Schie, L., Hoffmann, M., Pöhlmann, S., Graham, B. S., Callewaert, N., Schepens, B., Saelens, X., & McLellan, J. S. (2020). Structural basis for potent neutralization of betacoronaviruses by single-domain Camelid antibodies. *Cell*, 181(5), 1004–1015 e1015. <https://doi.org/10.1016/j.cell.2020.04.031>
- Yang, Y., Shen, Y., Liu, H., & Yao, X. (2011). Molecular dynamics simulation and free energy calculation studies of the binding mechanism of allosteric inhibitors with p38alpha MAP kinase. *Journal of Chemical Information and Modeling*, 51(12), 3235–3246. <https://doi.org/10.1021/ci200159g>
- Ying, T., Prabakaran, P., Du, L., Shi, W., Feng, Y., Wang, Y., Wang, L., Li, W., Jiang, S., Dimitrov, D. S., & Zhou, T. (2015). Junctional and allele-specific residues are critical for MERS-CoV neutralization by an exceptionally potent germline-like antibody. *Nature Communications*, 6, 8223. <https://doi.org/10.1038/ncomms9223>
- Yu, X., Zhang, S., Jiang, L., Cui, Y. E., Li, D., Wang, D., Wang, N., Fu, L., Shi, X., Li, Z., Zhang, L., & Wang, X. (2015). Structural basis for the neutralization of MERS-CoV by a human monoclonal antibody MERS-27. *Scientific Reports*, 5, 13133. <https://doi.org/10.1038/srep13133>
- Yuan, M., Wu, N. C., Zhu, X., Lee, C.-C., So, R. T. Y., Lv, H., Mok, C. K. P., & Wilson, I. A. (2020). A highly conserved cryptic epitope in the receptor binding domains of SARS-CoV-2 and SARS-CoV. *Science*, 368(6491), 630–633. <https://doi.org/10.1126/science.abb7269>
- Zavrtanik, U., Lukan, J., Loris, R., Lah, J., & Hadzi, S. (2018). Structural basis of epitope recognition by heavy-chain camelid antibodies. *Journal of Molecular Biology*, 430(21), 4369–4386. <https://doi.org/10.1016/j.jmb.2018.09.002>
- Zhang, S., Zhou, P., Wang, P., Li, Y., Jiang, L., Jia, W., Wang, H., Fan, A., Wang, D., Shi, X., Fang, X., Hammel, M., Wang, S., Wang, X., & Zhang, L. (2018). Structural definition of a unique neutralization epitope on the receptor-binding domain of MERS-CoV spike glycoprotein. *Cell Reports*, 24(2), 441–452. <https://doi.org/10.1016/j.celrep.2018.06.041>
- Zhou, D., Duyvesteyn, H. M. E., Chen, C.-P., Huang, C.-G., Chen, T.-H., Shih, S.-R., Lin, Y.-C., Cheng, C.-Y., Cheng, S.-H., Huang, Y.-C., Lin, T.-Y., Ma, C., Huo, J., Carrique, L., Malinauskas, T., Ruza, R. R., Shah, P. N. M., Tan, T. K., Rijal, P., ... Huang, K.-Y. (2020). Structural basis for the neutralization of SARS-CoV-2 by an antibody



- from a convalescent patient. *Nature Structural & Molecular Biology*, 27(10), 950–958. <https://doi.org/10.1038/s41594-020-0480-y>
- Zhou, H., Chen, Y., Zhang, S., Niu, P., Qin, K., Jia, W., Huang, B., Zhang, S., Lan, J., Zhang, L., Tan, W., & Wang, X. (2019). Structural definition of a neutralization epitope on the N-terminal domain of MERS-CoV spike glycoprotein. *Nature Communications*, 10(1), 3068. <https://doi.org/10.1038/s41467-019-10897-4>
- Zhou, P., Yang, X.-L., Wang, X.-G., Hu, B., Zhang, L., Zhang, W., Si, H.-R., Zhu, Y., Li, B., Huang, C.-L., Chen, H.-D., Chen, J., Luo, Y., Guo, H., Jiang, R.-D., Liu, M.-Q., Chen, Y., Shen, X.-R., Wang, X. I., ... Shi, Z.-L. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, 579(7798), 270–273. <https://doi.org/10.1038/s41586-020-2012-7>

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Yang J, Zhang Z, Yang F, et al. Computational design and modeling of nanobodies toward SARS-CoV-2 receptor binding domain. *Chem Biol Drug Des*. 2021;98:1–18. <https://doi.org/10.1111/cbdd.13847>