

# Computational design of ACE2-based short peptide inhibitors of SARS-CoV-2

Yanxiao Han<sup>1</sup> and Petr Král<sup>1,2\*</sup>

<sup>1</sup>*Department of Chemistry, University of Illinois at Chicago, Chicago, Illinois 60607, United States and*

<sup>2</sup>*Departments of Physics, Biopharmaceutical Sciences, and Chemical Engineering,  
University of Illinois at Chicago, Chicago, Illinois 60607, United States.*

(Dated: April 1, 2020)

Peptide inhibitors against the SARS-CoV-2 coronavirus, currently causing a worldwide pandemic, are designed and simulated. The inhibitors are formed by two sequential self-supporting  $\alpha$ -helices (bundle) extracted from the protease domain (PD) of angiotensin-converting enzyme 2 (ACE2), which binds to the SARS-CoV-2 receptor binding domains. Molecular dynamics simulations revealed that the peptides maintain their secondary structure and provide a highly specific and stable binding (blocking) to SARS-CoV-2, determined by their sequences and conformations. The proposed peptide inhibitors could provide simple therapeutics against the COVID-19 disease.

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), previously known as 2019 novel coronavirus (2019-nCoV) [1], is causing a pandemic of coronavirus disease 2019 [2, 3]. SARS-CoV-2 shares about 80% of its genome identity with SARS-CoV, which emerged in 2002-3 [4]. SARS-CoV-2 is highly contagious in humans, which has rapidly caused an unprecedented pandemic, with a large number of fatalities worldwide.

The SARS-CoV-2 virion, 50–200 nm in diameter, contains four structural proteins, known as the S (spike), E (envelope), M (membrane), and N (nucleocapsid) proteins [2]. The S protein, imaged at the atomic level using Cryo-EM [5], is responsible for the host attachment and fusion of the viral and host-cell membranes [6, 7]. This process is triggered when the S1 subunit of S protein binds to a host-cell receptor. To engage a host-cell receptor, the receptor-binding domain (RBD) of S1 undergoes transient hinge-like conformational motions (receptor-accessible or receptor-inaccessible states) [8]. The angiotensin-converting enzyme 2 (ACE2) is the host cellular receptor with a higher affinity to SARS-CoV-2 than to SARS-CoV [5]. In the recognition of RBD, the protease domain (PD) of ACE2 mainly engages the  $\alpha_1$  helix with a minor contribution from the  $\alpha_2$  helix and the linker of the  $\beta_3$  and  $\beta_4$  sheets [8, 9].

Besides a hectic search of vaccines against COVID-19, there is a very fast ongoing search for therapeutics acting on SARS-CoV-2. Depending on the activity, the therapies can be divided into several main categories: (1) preventing the viral RNA synthesis and replication, (2) blocking the virus from binding to human cell receptors, (3) restoring the host’s innate immunity, and (4) blocking host’s specific receptors or enzymes [10]. Despite many experimental and computational studies currently exploring all these categories, to date, there is no confirmed effective treatment specifically available for COVID-19.

Computational approaches have been used to search potential therapeutics against SARS-CoV-2 protease (category 1) [11]. Analogous screening of potential drugs against the S protein of SARS-CoV-2 (category 2) pro-

vided small molecular compounds with a high binding affinity. Unfortunately, most of these compounds do not attach with the binding interface of the RBD-ACE2 complex. Hesperidin was predicted to lie on the surface of RBD, but it didn’t cover the whole interface [10]. In the early attempts of SARS-CoV blocking, short peptide inhibitors were studied and amino acids mutations were implemented to the S protein of SARS-CoV [12, 13]. However, the proposed peptide was too short (8 residues) to maintain secondary structure, so it was unable to block the whole SARS-CoV binding surface [12]. Broad-spectrum antiviral nanoparticles and cyclodextrins were designed, simulated, and implemented in blocking of other viruses [14–16]. They are category 2 or 3 inhibitors, but their applicability to SARS-CoV-2 is unknown. Proteins or rigid peptides with specific (multivalent) binding domains and conformations matching RBD could be promising therapeutics for COVID-19. Overall, protein therapies show a high specificity, small interference with biological processes, good tolerance to human organisms, and faster FDA approval times [17].

In this work, we design and simulate several peptide inhibitors against SARS-CoV-2, which included components from the virus-binding domains of ACE2; based on the recently released crystal structure (PDB code: 2AJF [9]). The inhibitors, which have relatively low molecular weights, are structurally stable, they conformationally match the S protein, and are highly specific to SARS-CoV-2. This study could provide a potential guidance in antigen recognition and structure-based designs of antibodies with high affinities. The proposed small peptides could be used as inhaled therapeutics for topical lung delivery, providing an efficient way to combat COVID-19 [18].

## Systems and results

In the crystal structure of ACE2 and RBD of SARS-CoV-2 (PDB: 2AJF [9]), we first analyzed the interacting amino acids at the ACE2 and RBD interface. In total, 15 residues from ACE2 interact with RBD: residues 24, 27, 30, 31, 34, 35, 37, 38, 41, and 42 are in  $\alpha_1$ , one residue (residue 82) comes from  $\alpha_2$ , residues 353, 354, 355 and

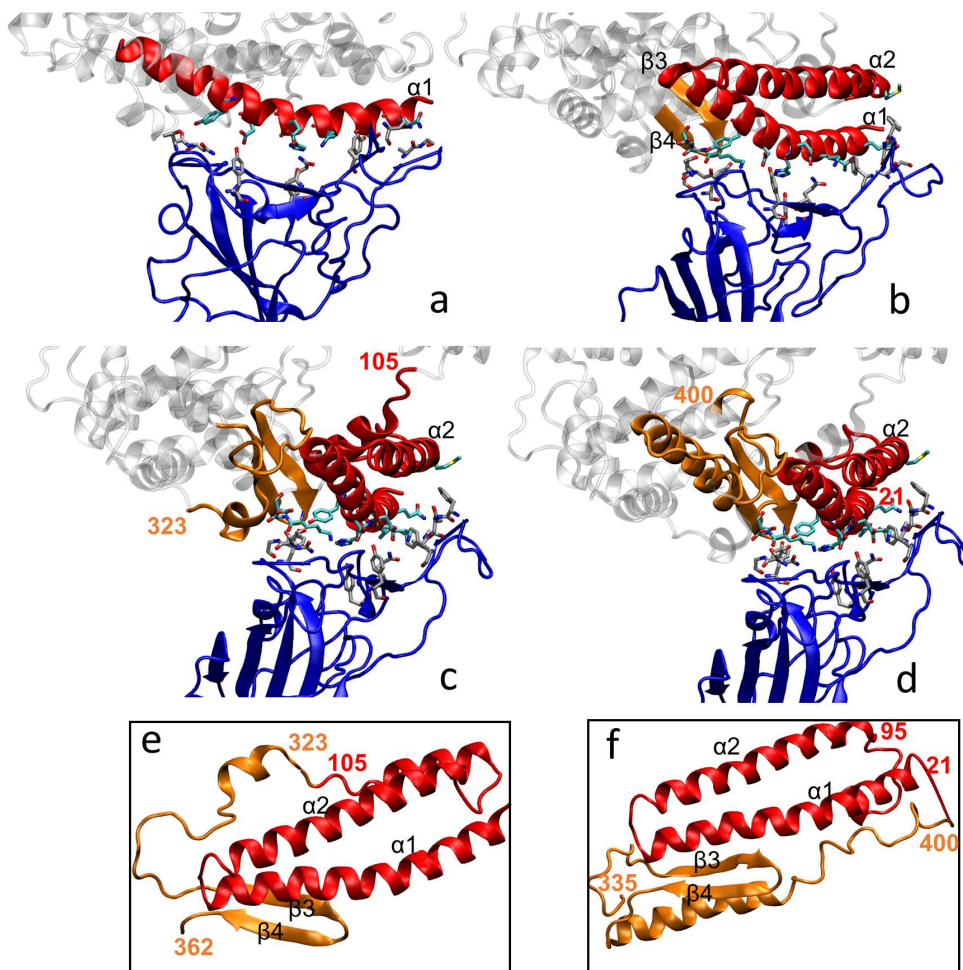


FIG. 1: Structural components of the inhibitors designed: (a) **inhibitor 1** is composed of  $\alpha_1$  (residues 21 to 55), (b) **inhibitor 2** is composed of  $\alpha_1$ ,  $\alpha_2$ , and loose chain between  $\beta_3$  and  $\beta_4$  connected by a C-C bond between residues 45 and 351 (residues 21 to 88 and 349 to 357); (c) **inhibitor 3** is composed of  $\alpha_1$ ,  $\alpha_2$ , and  $\beta_3$ ,  $\beta_4$  (residues 21 to 105 and 323 to 362) (d) **inhibitor 4** has the same composition as inhibitor 3 but different linkage (residues 21 to 95 and 335 to 400); (e) details of inhibitor 3 (c), reorganized with residue 323 connecting residue 105; (f) details of inhibitor 3 (c), reorganized with residue 21 connecting residue 400. In (e-f), the conformation of the  $\alpha$  helices and  $\beta$  sheets were maintained with the rest adapting to the connection. Coloring scheme: Red -  $\alpha$  helices; orange -  $\beta$  sheets or other linker components; blue - RBD of SARS-CoV-2; grey - other parts of ACE2; licorice - the initial contacting residues in the RBD-ACE2 interface.

357 come from the linker between  $\beta_3$  and  $\beta_4$ . Therefore, the 15 amino acids can be labeled as critical amino acids and  $\alpha_1$ ,  $\alpha_2$ ,  $\beta_3$  and  $\beta_4$  as critical binding components.

Since most of the interacting residues are from  $\alpha_1$ , we picked as **inhibitor 1** the  $\alpha_1$  helix alone. In particular, the 21 - 55 residues, shown in Fig. 1 a, were selected. Realizing that  $\alpha_1$  (alone) might not be even stable, we next picked as **inhibitor 2** both  $\alpha_1$  and  $\alpha_2$  helices (residues 21 to 88), and the residues 349 to 357 (residues between  $\beta_3$  and  $\beta_4$  shown in orange in Fig. 1 b). This selection included all the 15 interacting residues from the crystal structure 2AJF [9]. Since the two  $\alpha$  helices are closely joined on one side (Fig. 1 b), they stabilize each other. To connect the two helices (red) with the  $\beta$  sheets with residues 349 to 357 (orange), as shown in Fig. 1 b, residues 45 (LEU) and 351 (LEU) were linked together

by a side chain with a carbon-carbon bond, as shown in Fig. 2 b.

We have also designed other inhibitors that are closer to the ACE2 protein, use purely peptide bonds, and contain all 15 residues that initial bind to RBD in the 2AJF crystal structure [9]. Figure 1 c (detail Fig. 1 e) shows **inhibitor 3**, where residues 323 to 362 (orange) include the two  $\beta$  sheets and a random coil (residues 323 to 348), while residues 21 to 105 (red) include the two  $\alpha$  helices with another random coil (residues 89 to 105). The two sequences are joined together by a peptide bond between residues 105 and 323, and the two pieces of random coils were moved close to each other. Finally, Fig. 1 d (detail Fig. 1 f) shows **inhibitor 4**, where two sequences including residues 21 to 95 (red) and residues 335 to 500 (orange) were selected. An extra peptide bonds was made

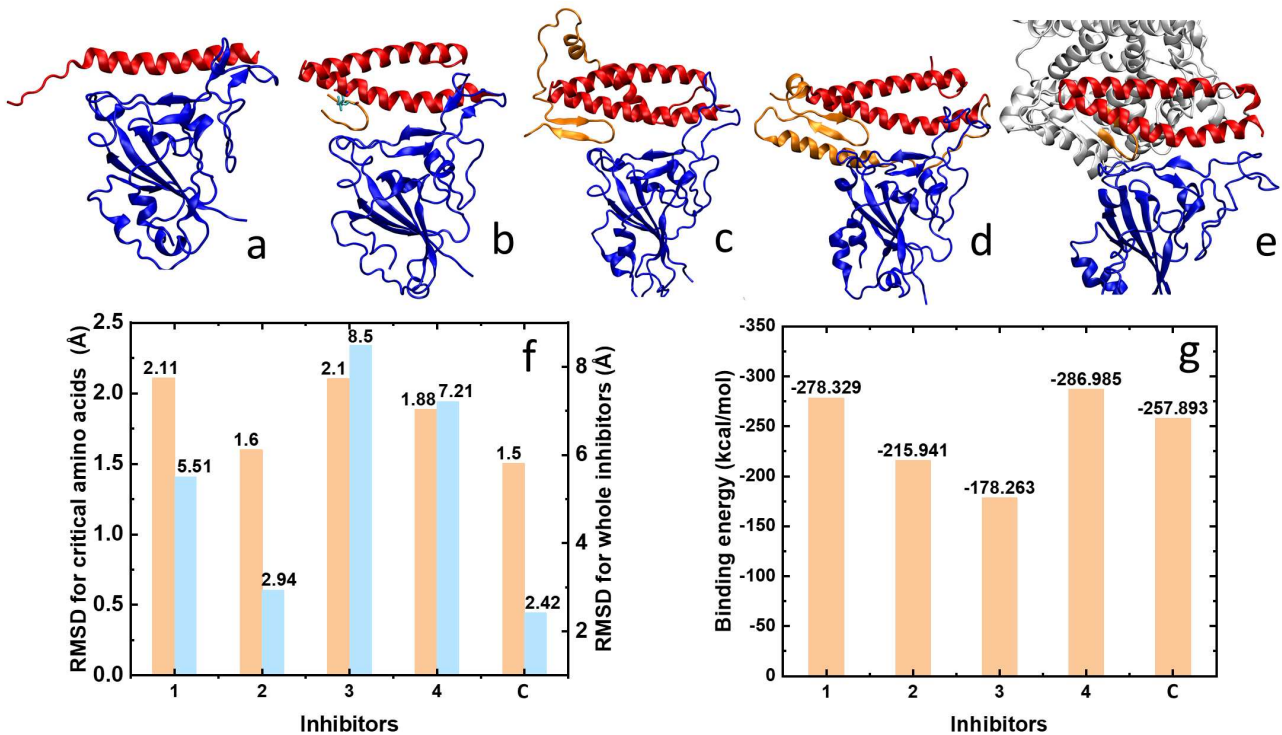


FIG. 2: (a-e) Final conformations of **inhibitors 1, 2, 3, 4** and **control**. (f) Averaged RMSD for the critical amino acids in each inhibitor and for the whole inhibitors when binding with the RBD of SARS-CoV-2. Numbering scale: 1-4 - inhibitors 1-4 with RBD; C - control system of PD from ACE2 and the RBD of SARS-CoV-2. (g) The average interaction energies between the contact residues of inhibitors 1-4 (or ACE2) and the RBD of SARS-CoV-2.

between residue 21 and residue 400 by adjusting the position of the corresponding sequences.

To examine how these potential inhibitors bind to RBD of SARS-CoV-2, we have prepared these systems in the initial position known from the crystal structure (PDB:2AJF), and simulated them in physiological solution (Methods), as shown in Figs. 2 a-d. As a control, the PD of ACE2 (residues 19 to 615) and RBD of SARS-CoV-2 were also simulated (Fig. 2 e).

#### Binding conformations

In Fig. 2 a, 170 ns long simulations showed that the helical structure of **inhibitor 1** deform from the left side - loose end unfolding, although it still binds to the RBD of SARS-CoV-2. In Fig. 2 b-d, 100 – 170 ns long simulations revealed that **inhibitors 2-4** bind in a stable way to the RBD of SARS-CoV-2, without  $\alpha_1$  losing its structure. Due to a different linkage among the critical binding components, the overall conformations of **inhibitors 2-4** vary. Specifically, the  $\alpha_1$  helix, which mostly contributes to the complementary sequence and conformational matching to RBD, is maintained in **inhibitors 2-4** with different degree of bending. The *beta* sheets in **inhibitors 3-4** structures are also preserved. Overall, the critical binding components in **inhibitors 2-4** bind to RBD in a very similar manner to the crystal structure.

#### RMSD and interaction energies

To further quantify the binding of these inhibitors to RBD, we calculated RMSD for the 15 critical amino acids in each inhibitor and for the whole inhibitors. Figure 2 f shows the average RMSD at the end of our simulations (see also Fig. S1). **Inhibitor 1** has the largest RMSD for the critical amino acids and the largest fluctuations both for the critical amino acids and the overall RMSD (Fig. S1 a and b). This can be attributed to unfolding of  $\alpha_1$ , shown in Fig. 2 a. A highly promising **inhibitor 2** has RMSD of the critical amino acids and the overall RMSD similar to those in the control (lowest). **Inhibitors 3-4** have higher RMSD of the critical amino acids and the overall RMSD. However, Fig. S1 b shows that **inhibitors 3-4** have a relatively smooth overall RMSD at latter times. This may be due to a poor adaption of their added connections at early times.

The interaction energies have van der Waals (vdW) and electrostatic components, calculated by the NAMD energy plugin. The total energies are shown in Fig. 2 g and Fig. S2 (detail). The residues which contribute to the interaction energies between inhibitors and SARS-CoV-2 are selected with a cutoff of 3 Å. The selections are updated in every frame. **Inhibitors 1, 4** show similar interaction energies as the control, while **inhibitors 2, 3** show a slightly lower interaction energies than the control. The larger interaction energy in **inhibitor 1**

might be due to non-specific interactions caused by the deformed helix.

In summary, using classical MD simulations, we have shown that **inhibitors 2-4** designed in this work provide highly promising trails for experimental testing of SARS-CoV-2 blocking. Suitable inhibitors need to have a stable binding, lower RMSD for the critical amino acids and relatively high binding energy. The incorporation of the two helices can help to maintain the conformational stability. The two helices in **inhibitors 2-4** keep in the bent shape (Fig. 2 b-d), which provide a conformational matching to the RBD of SARS-CoV-2 and a full cover of the RBD surface. Precise conformational matching between the designed peptides and the virus provides room for improving the binding affinity, which should be considered in future inhibitor design protocols. With further experimental evaluations, the inhibitors designed could be used as inhaled protein therapeutics in preventing the virus attachment in lungs.

### Methods

The inhibitors and RBD of the virus were simulated by NAMD [19] and the CHARMM36 protein force field [20]. The PME method was used for the evaluation of long-range Coulombic interactions. The time step was set to 2 fs. The simulations were performed in the NpT ensemble ( $p = 1$  bar and  $T = 298$  K), using the Langevin dynamics with a damping constant of  $1 \text{ ps}^{-1}$ . After 2,000 steps of minimization, ions and water molecules were equilibrated for 2 ns around proteins, which were restrained using harmonic forces with a spring constant of  $2 \text{ kcal}/(\text{mol } \text{Å}^2)$ . The last frames of restrained equilibration were used to start simulations of free inhibitors and partially constrained PD of ACE2 (two residues on the bottom).

**Acknowledgments** Y.H. acknowledges the support from the Dean's Scholar Fellowship (UIC).

\* The addresses for correspondence: pkral@uic.edu

- 
- [1] Gorbalenya, A.E.; Baker, S.C.; Baric, R.S. et al. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat. Microbiol.* **2020**, *5*, 536–544.
  - [2] Chen, N.; Zhou, M.; Dong, X.; Qu, J.; Gong, F.; Han, Y.; et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *The Lancet.* **2020**, *395*, 507–513.
  - [3] Huang, C.; Wang, Y.; Li, X.; Ren, L.; Zhao, J.; Hu, Y.; Zhang, L. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet* **2020**, *395*, 497–506.
  - [4] Zhou, P.; Yang, X.-L.; Wang, X.-G.; Hu, B.; Zhang, L. et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **2020**, *579*, 270–273.
  - [5] Wrapp, D.; Wang, N.; Corbett, K. S.; Goldsmith, J. A.; Hsieh, C.-L.; et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **2020**, *367*, 1260–1263. DOI: 10.1126/science.abb2507
  - [6] Li, F. Structure, Function, and Evolution of Coronavirus Spike Proteins. *Annu. Rev. Virol.* **2016**, *3*, 237–61.
  - [7] Bosch, B. J.; Zee, R.; Haan, C. A. M.; Rottier, P. J. M. The Coronavirus Spike Protein Is a Class I Virus Fusion Protein: Structural and Functional Characterization of the Fusion Core Complex. *J. Virol.* **2003**, *77*, 8801–8811.
  - [8] Wan, Y.; Shang, J.; Graham, R.; Baric, R. S.; Li, F. Receptor Recognition by the Novel Coronavirus from Wuhan: an Analysis Based on Decade-Long Structural Studies of SARS Coronavirus. *J. Virol.* **2020**, *94*, 00127–20.
  - [9] Yan, R.; Zhang, Y.; Xia, L.; Guo, Y. Structural basis for the recognition of the SARS-CoV-2 by full-length human ACE2. *Science* 10.1126/science.abb2762 (2020).
  - [10] Wu, C.; Liu, Y.; Yang, Y.; Zhang, P.; Zhong, W.; Wang, Y.; Wang, Q.; Xu, Y.; Li, M.; Li, X.; Zheng, M.; Chen, L.; Li, H. Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. *Acta Pharm. Sin. B.* **2020** In Press.
  - [11] Wang, J. Fast Identification of Possible Drug Treatment of Coronavirus Disease -19 (COVID-19). Through Computational Drug Repurposing Study. ChemRxiv. Preprint. <https://doi.org/10.26434/chemrxiv.11875446.v1>
  - [12] Du, Q.; Wang, S.; Wei, D.; Sirois, S.; Chou, K.-C. Molecular modeling and chemical modification for finding peptide inhibitor against severe acute respiratory syndrome coronavirus main proteinase. *Anal. Biochem.* **2005**, *337*, 262–270.
  - [13] Zhang, Y.; Zheng, N.; Nan, P.; Cao, Y.; Hasegawa, M. et al. Computational simulation of interactions between SARS coronavirus spike mutants and host species-specific receptors. *Comput. Biol. Chem.* **2007**, *31*, 134–137.
  - [14] Cagno, V.; Andreozzi, P.; D'Alicarnasso, M.; Silva, P. J.; Mueller, M. et al. Broad-spectrum non-toxic antiviral nanoparticles with a virucidal inhibition mechanism. *Nature Mater.* **2018**, *17*, 195–203.
  - [15] Jone, S.; Cagno, V.; Janecsek, M.; Ortiz, D.; Gasilova, N.; et al. Modified cyclodextrins as broad-spectrum antivirals. *Sci. Adv.* **2020**, *6* eaax9318.
  - [16] Sen, S.; Han, Y.; Rehak, P.; Vuković, L.; Král, P. Computational Studies of Micellar and Nanoparticle Nanomedicines. *Chem. Soc. Rev.* **2018**, *47*, 3849–3860.
  - [17] Leader, B.; Baca, Q.; Golan, D.; Protein therapeutics: a summary and pharmacological classification. *Nat. Rev. Drug Discov.* **2008**, *7*, 21–39.
  - [18] Bodier-Montagutelli, E.; Mayor, A.; Vecellio, L.; Respaud, R.; Heuzé-Vourc'h, N. Designing inhaled protein therapeutics for topical lung delivery: what are the next steps? *Expert Opin. Drug Deliv.* **2018**, *15*, 729–736
  - [19] Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802.
  - [20] MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D. et al. All-atom empirical potential for molecular modeling and dynamics Studies of proteins. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.