

Is the rigidity of SARS-CoV-2 spike receptor-binding motif the hallmark for its enhanced infectivity? An answer from all-atoms simulations

Angelo Spinello,^{a,#} Andrea Saltalamacchia,^{b,#} Alessandra Magistrato,^{a,*}

a CNR-IOM c/o SISSA, via Bonomea 265, 34136 Trieste, Italy

b International School for Advanced Studies SISSA, via Bonomea 265, 34136 Trieste, Italy

Abstract

The severe acute respiratory syndrome coronavirus (SARS-CoV-2) pandemic is setting the global health crisis of our time, causing a devastating societal and economic burden. An idiosyncratic trait of coronaviruses is the presence of spike glycoproteins on the viral envelope, which mediate the virus binding to specific host receptor, enabling its entry into the human cells. In spite of the high sequence identity of SARS-CoV-2 with its closely related SARS-CoV emerged in 2002, the atomic-level determinants underlining the molecular recognition of SARS-CoV-2 to the angiotensin-converting enzyme 2 (ACE2) receptor and, thus, the rapid virus spread into human body, remain unresolved. Here, multi- μ s-long molecular dynamics simulations enabled us to unprecedentedly dissect the key molecular traits liable of the higher affinity/specificity of SARS-CoV-2 towards ACE2, as compared to SARS-CoV, supplying a minute per-residue topological map underlining its stunningly high infectivity. Harnessing this knowledge is pivotal for urgently developing effective medical countermeasures to face the ongoing global health crisis.

#Authors equally contributed to this work

*Correspondence to alessandra.magistrato@sissa.it

The latest outbreak of a severe viral pneumonia, commonly referred as coronavirus disease 19 (COVID-19), originated in December 2019 in the city of Wuhan, China, has soon thereafter rapidly spread worldwide, being officially declared on March 11 a pandemic by the World Health Organization.¹⁻² As yet of April 6th COVID-19 infected over 1.3 million patients and caused over 70,000 deaths worldwide. The pathogen responsible for this disease is a novel β -coronaviruses (β -CoV), named severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) after its closely related SARS-CoV,³ which, in 2002, caused 8,096 cases and 774 deaths worldwide.⁴ Additionally, a distinct coronavirus (Middle East respiratory syndrome coronavirus, MERS-CoV), in 2012, also spread in 27 different countries, causing 2,494 cases and 858 deaths.⁵ These numbers and the recurrence of this phenomenon underline that future outbreaks of new zoonotic threatening transmissions are not unlikely.

Similarly to other β -CoVs, the receptor-binding domain (RBD) of the homotrimeric viral spike (S) glycoprotein of SARS-CoVs mediates the molecular recognition to the host cellular receptor, acting as the trojan horse for the virus entry into host cells. Hence, the S protein is considered a key molecular target for the design and development of specific antibodies,⁶ currently being the object of burgeoning structural vaccinology studies. Stunningly, the phylogenetically similar S proteins of SARS-CoV and SARS-CoV-2 possess a sequence identity of about 77%,³ both hijacking angiotensin-converting enzyme 2 (ACE2), a zinc metallopeptidase entailed with cardiovascular and immune systems regulation,⁷⁻⁸ to enter and infect the human cells. The structural features of the interactions between the S protein and ACE2 are currently being addressed in many biophysical studies, aiming to clarify the reasons underlying the high human-to-human transmissibility of SARS-CoV-2 as compared to its closely related SARS-CoV variant. Cryo-EM and biophysical (surface plasmon resonance) studies provided insights on structure of the S glycoprotein's RBD in complex with ACE2, suggesting a higher binding affinity of SARS-CoV-2 towards ACE2 as compared to SARS-CoV.⁹ This was corroborated by a cytometry analysis and immunofluorescence staining based study.¹⁰ Conversely, two distinct studies, based on biolayer interferometry, proposed that the RBD of the two SARS variants share a similar binding affinity towards ACE2.¹¹⁻¹² In light of the ongoing global health emergency, these controversial evidences urgently call for a rapid defeasance. Aiming to extricate this puzzling scenario at atomic-level, we performed μ s-long all-atoms explicitly solvated molecular dynamics (MD) simulations of the RBD of SARS-CoV-2 and SARS-CoV/ACE2 adducts, starting from their recently released structures (pdb ID 6M0J and 6ACJ),¹³⁻¹⁴ respectively (see Supplementary Information (SI) for a detailed description). Most of the SARS-CoV(-2)'s residues binding to ACE2 belong to the RBD's receptor-binding motif (RBM, Figure 1A), which is made of 4 loops, divided by two small β -strands.

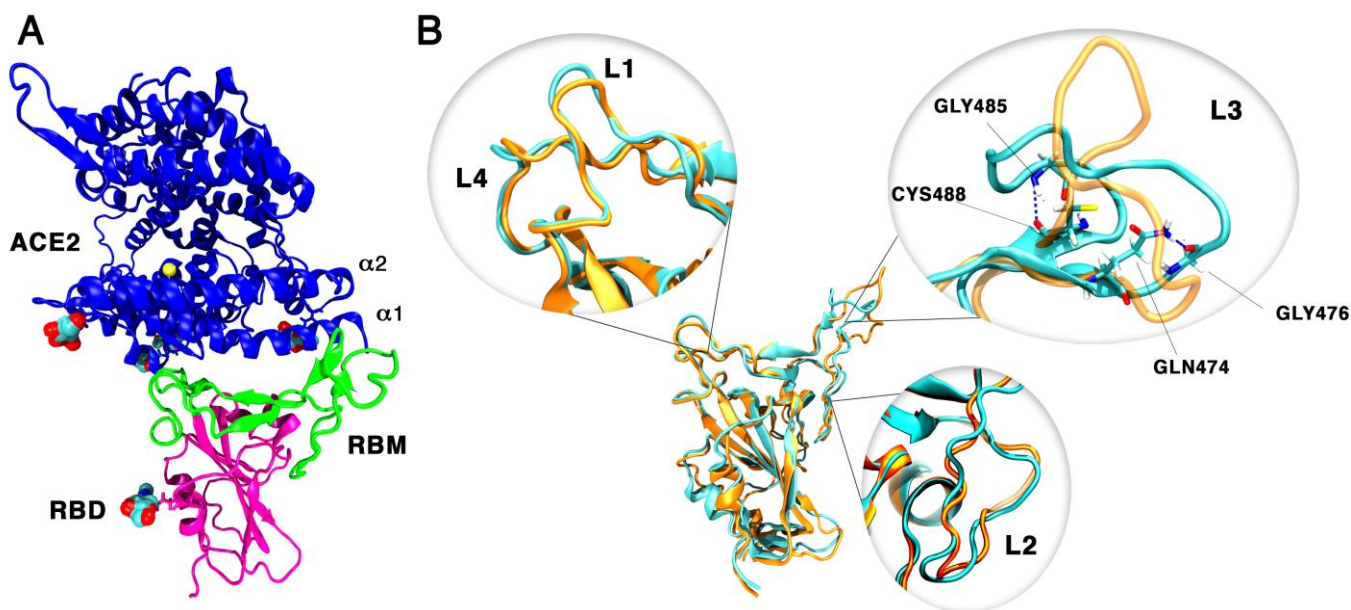
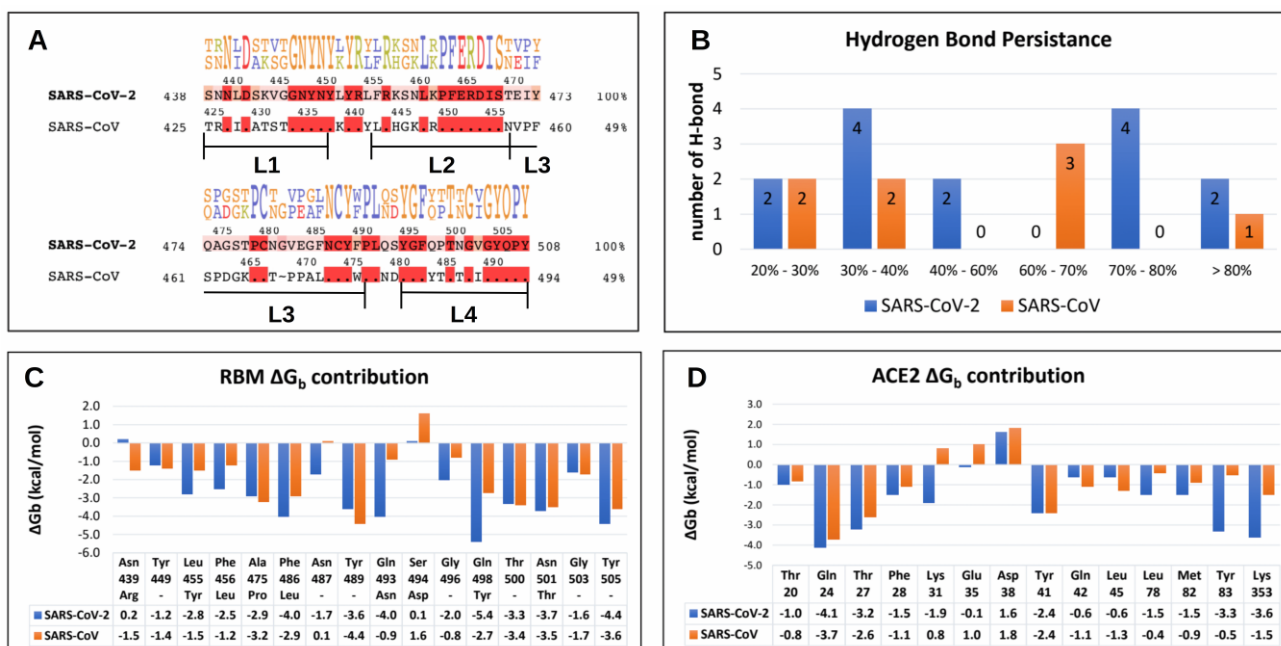


Figure 1. (A) Adduct between angiotensin-converting enzyme 2 (ACE2, blue) and receptor-binding domain (RBD) of SARS-CoV-2 (pink) with the receptor binding motif (RBM) highlighted (green). Glycans and Zn²⁺ ions are depicted in van der Waals spheres and colored in cyan, red and yellow for carbon, oxygen and zinc atoms, respectively. (B) Alignment of the most representative structures of the RBD of SARS-CoV-2 (cyan) and SARS-CoV (orange), as obtained from a cluster analysis of the molecular dynamics trajectories. The insets compare the loops' (L1-2-3-4) structural organization.

During the MD simulations the two SARS-CoV(-2)/ACE2 adducts establish stable interfacial interactions (Figure S1), showing structurally similar binding features. The main differences are restricted to a loop of the RBM (composed by Thr470-Pro491 and Asn457-Pro477 for SARS-CoV-2 and SARS-CoV, respectively, Figure 1B), which engages persistent interactions with ACE2. This recognition loop (hereafter also referred as loop 3, L3) is markedly more rigid in the SARS-CoV-2/ACE2 adduct as compared to its older variant (Figures S2-S3). Indeed, L3@SARS-CoV-2 possesses a more defined secondary structure (composed by small β -sheets), that is preserved along the MD simulations (Figure S4). Ostensibly, L3's length is different in the two SARS variants, being characterized by the insertion of Gly482, in SARS-CoV-2. This makes L3 longer and more structured, enabling it to gain stabilizing interactions (namely, the mutated residues Gly485 with Cys488 and Gln474 with Gly476, Figure 1B) in the SARS-CoV-2/ACE2 adduct. This amino acid insertion, along with other amino acidic mutations (Figures 1B, 2A), stunningly convert an unessential part of the RBM into a strikingly effective recognition grasp for ACE2, allowing SARS-CoV-2 to stiffen, by establishing stronger interfacial interactions.



Indeed, the hydrogen (H)-bond analysis of the adducts pinpoints a larger number of more persistent H-bonds and salt bridges at SARS-CoV-2/ACE2's interface (Figure 2B and Table S1) as compared to its SARS predecessor. This translates in a striking difference of binding free energy (ΔG_b) of the two adducts, as obtained with a Molecular Mechanics Generalized Born Surface Area (MM-GBSA) method, which favors SARS-CoV-2/ACE2 by 20.8 kcal/mol (Table S2), with the highest observed difference of 29.5 kcal/mol in the last part of the trajectory, with both the electrostatic and van der Waals components contributing to the enhanced ΔG_b .¹⁵⁻¹⁶ This remarkable difference in the $\Delta\Delta G_b$ of SARS-CoV(-2)/ACE2 adducts is overestimated by the employed methodology, being nevertheless qualitatively consistent with the experimental K_d values, which pinpoint a significantly higher ACE2 binding affinity toward the novel coronavirus (15–44 nM and 185–326 nM for SARS-CoV-2 and SARS-CoV, respectively).^{9,17}

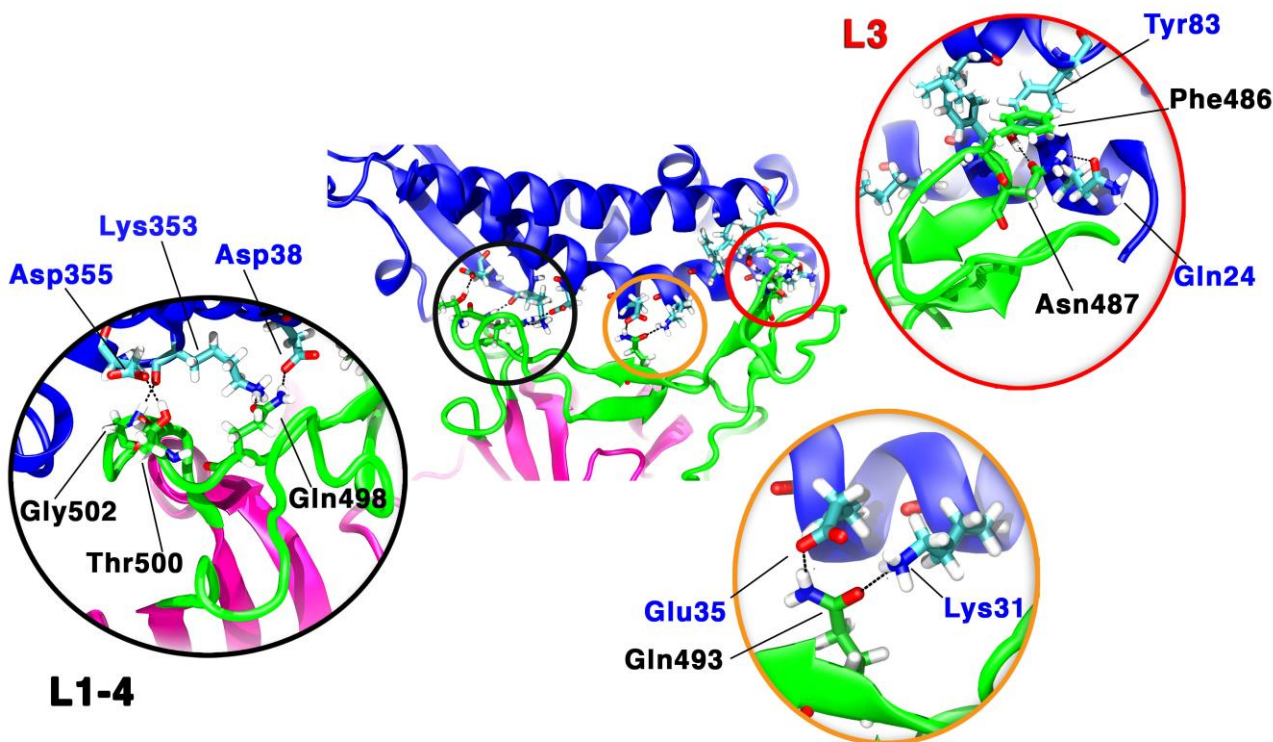


Figure 3. Key residues engaging hydrogen bonds and π -stacking interactions at the SARS-CoV-2/ACE2 interface (blue and green ribbons, respectively), as evidenced by MD simulations trajectory. Hydrogen bonds are highlighted as black dashed lines.

This detailed per-residue topological signature of the binding interface also discloses that Asp480@SARS-CoV has a negative impact on ΔG_b . This residue forms only a low-persistent salt-bridge with Lys439@SARS-CoV, which due to RBM's flexibility breaks for most of the MD trajectory. Thus, since Asp480 destabilizes the SARS-CoV/ACE2 adduct owing to the electrostatic repulsion with two nearby negatively charged residues (Glu35 and Asp38@ACE2- α 1, Figures 2C-D), the mutation to Ser494@SARS-CoV-2 resolves this unfavorable contribution.

Finally, a comparison of the electrostatic potential, at the ACE2's interface with the two SARS variants did not reveal marked differences. In particular, SARS-CoV-2 shows only a slightly wider negatively charged surface (Figure S5).

To better decrypt the relative importance of the distinct components of RBM, we even computed the Pearson-based cross-correlation matrix (CCM) of both SARS-CoV(-2)/ACE2 adducts, capturing the dynamically coupled motions and pinpointing dynamical differences between the two β -CoV variants.¹⁸ Although by comparing the two matrices no difference has been detected on the ACE2's dynamics (Figure S6), by focusing on the interfacial residues some diversities appear. Strikingly, all the RBM's loops, besides L2, which is the farthest from ACE2, exhibit an increased correlation in SARS-CoV-2/ACE2 (Figure 4) as compared to SARS-CoV/ACE2 adduct. This is particularly true for L3, whose increased content of secondary structure (Figure S4) affects the strength of the interactions, and in turn the positive correlations, with the residues of the two interfacial ACE2's α -helices (residues 19-45@ACE2- α 1 and 79-83@ACE2- α 2, respectively). As well, L1 and L4 increase the grip to ACE2, gaining correlations with its nearby residues (324-330, 351-357, 386-393).

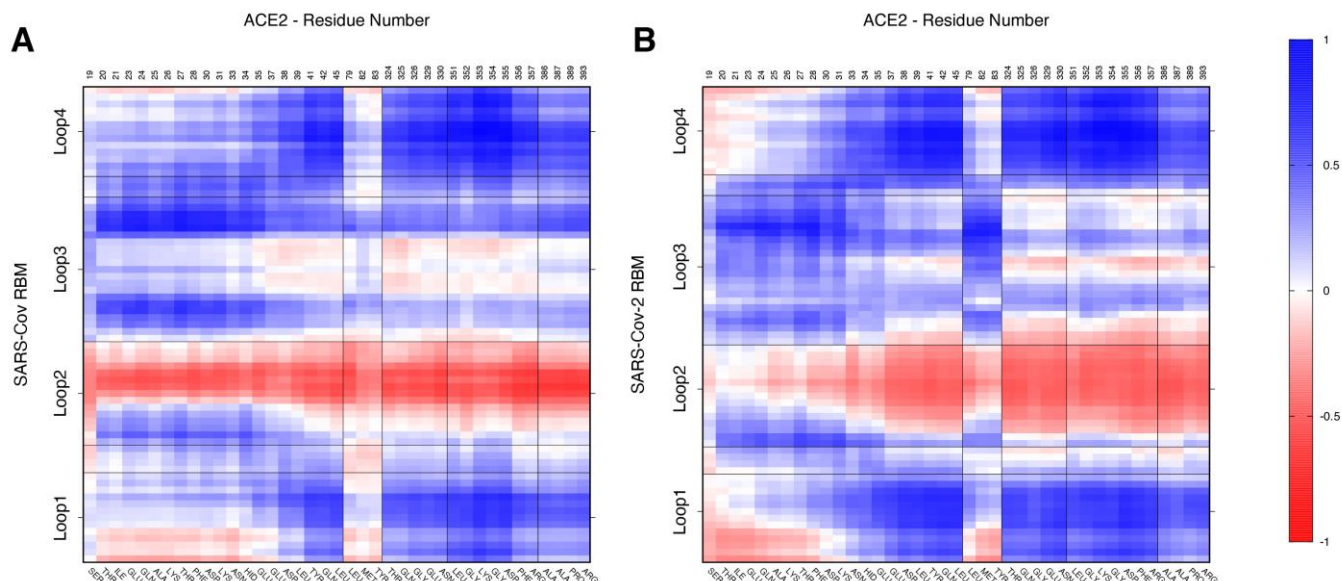


Figure 4. Cross-correlation matrix of the residues at the interface between ACE2 and (A) SARS-CoV and (B) SARS-CoV-2, based on per-residue Pearson's correlation coefficients (CCs) as derived from the mass-weighted covariance matrix calculated over 800 ns of molecular dynamics simulation trajectory. CCs values range from -1 (red, anti-correlated motions) to +1 (blue, correlated motions). RBM is separated into loops. ACE2 residues are divided according to their secondary structure element. ACE2's residue names and numbers are listed on the bottom and top of the matrix, respectively.

By summing the per-residue correlations of these matrices (i.e. measuring the per-residue correlation significance) we pinpoint the cardinal RBM's residues for ACE2 recognition (Table S3). Among these, the atomic fluctuations of Asn501@SARS-CoV-2 have a striking impact on the surrounding interactions, being at the center of an intricate H-bonding network. Because of its intermolecular H-bond with Gln498@SARS-CoV-2, Asn501 enables it to firmly interact with Asp38@ACE2- α 1 and Lys353@ACE2. Remarkably, this analysis pinpoints the relevance of another mutated residue, i.e. Phe486@SARS-CoV-2, located on L3, which performs π -stacking interaction with Tyr83@ACE2- α 2 (Figure 3), contributing to L3 stabilization and, thus, to ACE2 recognition.

Our results support and expand the first attempts to capture the key interactions of SARS-CoV-2/ACE2 complex performed at structural and bioinformatics level,¹⁹⁻²⁰ predicting the importance of critical residues (i.e. Gln498 and Phe486 mutations in SARS-CoV-2). The so-far available experimental structures of the SARS-CoV-2(RBD)/ACE2^{14,21} supplied precious information on their interfacial H-bond network. Nevertheless, these structural studies can only provide a static picture, being unable to exhaustively unravel the relative importance of these key interfacial interactions. Finite temperature all-atoms simulations complement these pieces of information, uncovering the relative strengths and persistence of the principal interactions engaged by key mutations occurring in SARS-CoV-2 (Table S4), pinpointing, as a result, the critical chemical determinants for the higher stability of the SARS-CoV-2/ACE2 adduct. Additionally, our MD simulations disclose an induced fit of ACE2- α 1 helix, which along the MD trajectory better adapts to the RBD@SARS-CoV-2 by increasing its tilt (Figure S7).

Owing to its extremely high contagiousness, COVID-19 pandemic is dramatically hitting one by one every country, while taking an increasingly deadly toll of lives. Since, no therapeutic options are available for COVID-19, identifying effective medical countermeasures is a current global clinical urgency. Despite the high sequence identity between SARS-CoV and SARS-CoV-2, potent neutralizing antibodies, specific for targeting ACE2's binding site of SARS-CoV/ACE2 (Figure S8),²²⁻²³ failed to bind SARS-CoV-2 spike protein.^{9,12} Remarkably, the crystal structure of CR3022 antibody, isolated from a convalescent SARS patient,²⁴ disclosed a cryptic epitope on the S protein of SARS-CoV-2 (Figure S8),²⁵ for which no competition with ACE2 is required, suggesting that this virus has a complex interactome to unravel and exploit for vaccine and drug design studies. In this context, the structural and dynamical differences recorded at the interfaces of the SARS-CoV(-2)/ACE2 adducts and the higher ΔG_b of SARS-CoV-2 toward ACE2 provide a rationale to this lack of cross-reactivity, fostering the identification of novel epitopes as exclusive and specific hallmarks of SARS-CoV-2.

In conclusion, our outcomes unprecedentedly disclose the key molecular traits underlying the higher affinity of SARS-CoV-2 toward the human host ACE2 receptor, supplying a meticulous atomic-level topological map of the critical residues and the pivotal interactions possibly triggering SARS-CoV-2's entrance into the host cells. The very first step of SARS viral infection is indeed the ACE2 receptor recognition by the S protein. Hence, the enhanced binding affinity of SARS-CoV-2 towards ACE2, as compared to that of SARS-CoV, disclosed by us and by recent experimental studies,⁹⁻¹⁰ is most likely correlated with its remarkable human-to-human transmissibility.^{11,26} Besides rationalizing the high infectivity of SARS-CoV-2, our study provides novel fundamental advances for boosting the urgent development of effective therapeutic strategies to face the ongoing COVID-19 pandemic and possibly to fight future coronavirus outbreaks.

Acknowledgement

AS was supported by a FIRC-AIRC "Mario e Valeria Rindi" fellowship for Italy.

Supporting information. Computational Details, Figure S1 to S8 and Table S1 to S4. This material is available free of charge at via the Internet at <http://pubs.acs.org>.

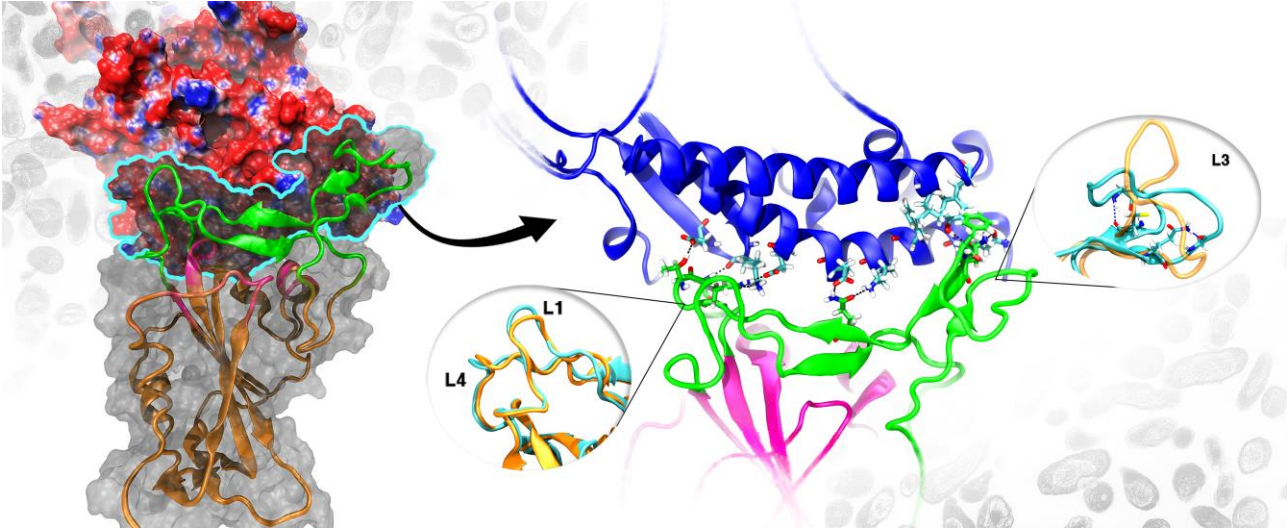
References

1. Huang, C.; Wang, Y.; Li, X.; Ren, L.; Zhao, J.; Hu, Y.; Zhang, L.; Fan, G.; Xu, J.; Gu, X.; Cheng, Z.; Yu, T.; Xia, J.; Wei, Y.; Wu, W.; Xie, X.; Yin, W.; Li, H.; Liu, M.; Xiao, Y.; Gao, H.; Guo, L.; Xie, J.; Wang, G.; Jiang, R.; Gao, Z.; Jin, Q.; Wang, J.; Cao, B., Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **2020**, 395 (10223), 497-506.
2. Chan, J. F.; Yuan, S.; Kok, K. H.; To, K. K.; Chu, H.; Yang, J.; Xing, F.; Liu, J.; Yip, C. C.; Poon, R. W.; Tsoi, H. W.; Lo, S. K.; Chan, K. H.; Poon, V. K.; Chan, W. M.; Ip, J. D.; Cai, J. P.; Cheng, V. C.; Chen, H.; Hui, C. K.; Yuen, K. Y., A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet* **2020**, 395 (10223), 514-523.
3. 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.; Zheng, X. S.; Zhao, K.; Chen, Q. J.; Deng, F.; Liu, L. L.; Yan, B.; Zhan, F. X.; Wang,

- Y. Y.; Xiao, G. F.; Shi, Z. L., A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **2020**, 579 (7798), 270-273.
4. de Wit, E.; van Doremalen, N.; Falzarano, D.; Munster, V. J., SARS and MERS: recent insights into emerging coronaviruses. *Nat. Rev. Microbiol.* **2016**, 14 (8), 523-34.
5. Zaki, A. M.; van Boheemen, S.; Bestebroer, T. M.; Osterhaus, A. D.; Fouchier, R. A., Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N. Engl. J. Med.* **2012**, 367 (19), 1814-20.
6. Du, L.; He, Y.; Zhou, Y.; Liu, S.; Zheng, B. J.; Jiang, S., The spike protein of SARS-CoV--a target for vaccine and therapeutic development. *Nat. Rev. Microbiol.* **2009**, 7 (3), 226-36.
7. Turner, A. J.; Hiscox, J. A.; Hooper, N. M., ACE2: from vasopeptidase to SARS virus receptor. *Trends Pharmacol. Sci.* **2004**, 25 (6), 291-4.
8. Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Kruger, N.; Herrler, T.; Erichsen, S.; Schiergens, T. S.; Herrler, G.; Wu, N. H.; Nitsche, A.; Muller, M. A.; Drosten, C.; Pohlmann, S., SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **2020**.
9. Wrapp, D.; Wang, N.; Corbett, K. S.; Goldsmith, J. A.; Hsieh, C. L.; Abiona, O.; Graham, B. S.; McLellan, J. S., Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **2020**, 367 (6483), 1260-1263.
10. Tai, W.; He, L.; Zhang, X.; Pu, J.; Voronin, D.; Jiang, S.; Zhou, Y.; Du, L., Characterization of the receptor-binding domain (RBD) of 2019 novel coronavirus: implication for development of RBD protein as a viral attachment inhibitor and vaccine. *Cell Mol. Immunol.* **2020**.
11. Walls, A. C.; Park, Y. J.; Tortorici, M. A.; Wall, A.; McGuire, A. T.; Veasley, D., Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **2020**.
12. 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.
13. Song, W.; Gui, M.; Wang, X.; Xiang, Y., Cryo-EM structure of the SARS coronavirus spike glycoprotein in complex with its host cell receptor ACE2. *PLoS Pathog.* **2018**, 14 (8), e1007236.
14. Lan, A.; Ge, J.; Yu, J.; Shan, S.; Zhou, H.; Fan, S.; Zhang, Q.; Shi, X.; Wang, L.; Zhang, L.; Wang, X., Crystal structure of the 2019-nCoV spike receptor-binding domain bound with the ACE2 receptor. *bioRxiv* **2020**.
15. Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E., 3rd, Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.* **2000**, 33 (12), 889-97.
16. Chen, F.; Liu, H.; Sun, H.; Pan, P.; Li, Y.; Li, D.; Hou, T., Assessing the performance of the MM/PBSA and MM/GBSA methods. 6. Capability to predict protein-protein binding free energies and re-rank binding poses generated by protein-protein docking. *Phys Chem Chem Phys* **2016**, 18 (32), 22129-39.
17. Shang, J.; Ye, G.; Shi, K.; Wan, Y.; Luo, C.; Aihara, H.; Geng, Q.; Auerbach, A.; Li, F., Structural basis of receptor recognition by SARS-CoV-2. *Nature* **2020**.
18. Casalino, L.; Palermo, G.; Spinello, A.; Rothlisberger, U.; Magistrato, A., All-atom simulations disentangle the functional dynamics underlying gene maturation in the intron lariet spliceosome. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, 115 (26), 6584-6589.
19. 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 (7).
20. Chen, Y.; Guo, Y.; Pan, Y.; Zhao, Z. J., Structure analysis of the receptor binding of 2019-nCoV. *Biochem. Biophys. Res. Commun.* **2020**.

21. Yan, R.; Zhang, Y.; Li, Y.; Xia, L.; Guo, Y.; Zhou, Q., Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* **2020**, *367* (6485), 1444-1448.
22. Zhu, Z.; Chakraborti, S.; He, Y.; Roberts, A.; Sheahan, T.; Xiao, X.; Hensley, L. E.; Prabakaran, P.; Rockx, B.; Sidorov, I. A.; Corti, D.; Vogel, L.; Feng, Y.; Kim, J. O.; Wang, L. F.; Baric, R.; Lanzavecchia, A.; Curtis, K. M.; Nabel, G. J.; Subbarao, K.; Jiang, S.; Dimitrov, D. S., Potent cross-reactive neutralization of SARS coronavirus isolates by human monoclonal antibodies. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (29), 12123-8.
23. Sui, J.; Li, W.; Roberts, A.; Matthews, L. J.; Murakami, A.; Vogel, L.; Wong, S. K.; Subbarao, K.; Farzan, M.; Marasco, W. A., Evaluation of human monoclonal antibody 80R for immunoprophylaxis of severe acute respiratory syndrome by an animal study, epitope mapping, and analysis of spike variants. *J. Virol.* **2005**, *79* (10), 5900-6.
24. ter Meulen, J.; van den Brink, E. N.; Poon, L. L.; Marissen, W. E.; Leung, C. S.; Cox, F.; Cheung, C. Y.; Bakker, A. Q.; Bogaards, J. A.; van Deventer, E.; Preiser, W.; Doerr, H. W.; Chow, V. T.; de Kruif, J.; Peiris, J. S.; Goudsmit, J., Human monoclonal antibody combination against SARS coronavirus: synergy and coverage of escape mutants. *PLoS Med.* **2006**, *3* (7), e237.
25. Yuan, M.; Wu, N. C.; Zhu, X.; Lee, 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**.
26. Wu, K.; Peng, G.; Wilken, M.; Geraghty, R. J.; Li, F., Mechanisms of host receptor adaptation by severe acute respiratory syndrome coronavirus. *J. Biol. Chem.* **2012**, *287* (12), 8904-11.

Table of contents



The latest outbreak of a new pathogenic coronavirus (SARS-CoV-2) is provoking a global health, economic and societal crisis. All-atom simulations enabled us to uncover the key molecular traits underlying the high affinity of SARS-CoV-2 spike glycoprotein towards its human receptor, providing a rationale to its high infectivity. Harnessing this knowledge can boost developing effective medical countermeasures to fight the current global pandemic (see picture).