

# Structural basis of receptor recognition by SARS-CoV-2

<https://doi.org/10.1038/s41586-020-2179-y>

Received: 16 February 2020

Accepted: 20 March 2020

Published online: 30 March 2020

 Check for updates

Jian Shang<sup>1,3</sup>, Gang Ye<sup>1,3</sup>, Ke Shi<sup>2,3</sup>, Yushun Wan<sup>1,3</sup>, Chuming Luo<sup>1</sup>, Hideki Aihara<sup>2</sup>, Qibin Geng<sup>1</sup>, Ashley Auerbach<sup>1</sup> & Fang Li<sup>1</sup>✉

A novel severe acute respiratory syndrome (SARS)-like coronavirus (SARS-CoV-2) recently emerged and is rapidly spreading in humans, causing COVID-19<sup>1,2</sup>. A key to tackling this pandemic is to understand the receptor recognition mechanism of the virus, which regulates its infectivity, pathogenesis and host range. SARS-CoV-2 and SARS-CoV recognize the same receptor—angiotensin-converting enzyme 2 (ACE2)—in humans<sup>3,4</sup>. Here we determined the crystal structure of the receptor-binding domain (RBD) of the spike protein of SARS-CoV-2 (engineered to facilitate crystallization) in complex with ACE2. In comparison with the SARS-CoV RBD, an ACE2-binding ridge in SARS-CoV-2 RBD has a more compact conformation; moreover, several residue changes in the SARS-CoV-2 RBD stabilize two virus-binding hotspots at the RBD–ACE2 interface. These structural features of SARS-CoV-2 RBD increase its ACE2-binding affinity. Additionally, we show that RaTG13, a bat coronavirus that is closely related to SARS-CoV-2, also uses human ACE2 as its receptor. The differences among SARS-CoV-2, SARS-CoV and RaTG13 in ACE2 recognition shed light on the potential animal-to-human transmission of SARS-CoV-2. This study provides guidance for intervention strategies that target receptor recognition by SARS-CoV-2.

The sudden emergence and rapid spread of SARS-CoV-2 is endangering global health and economy<sup>1,2</sup>. SARS-CoV-2 has caused many more infections, deaths and economic disruptions than SARS-CoV in 2002–2003<sup>5,6</sup>. The origin of SARS-CoV-2 remains unclear. Bats are considered the original source of SARS-CoV-2 because a closely related coronavirus, RaTG13, has been isolated from bats<sup>7</sup>. However, the molecular events that led to the possible bat-to-human transmission of SARS-CoV-2 are unknown. Clinically approved vaccines or drugs that specifically target SARS-CoV-2 are also lacking. Receptor recognition by coronaviruses is an important determinant of viral infectivity, pathogenesis and host range<sup>8,9</sup>. It presents a major target for vaccination and antiviral strategies<sup>10</sup>. Here we elucidate the structural and biochemical mechanisms of receptor recognition by SARS-CoV-2.

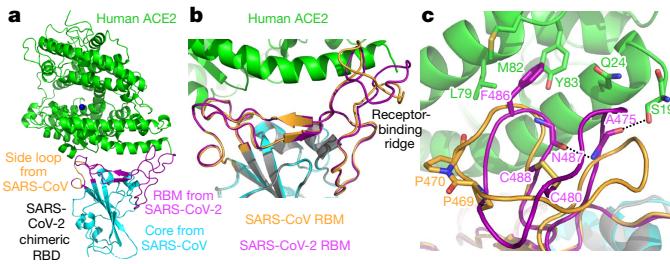
Receptor recognition by SARS-CoV has been extensively studied. A virus-surface spike protein mediates the entry of coronavirus into host cells. The spike protein of SARS-CoV contains a RBD that specifically recognizes ACE2 as its receptor<sup>3,4</sup>. A series of crystal structures of the SARS-CoV RBD from different strains in complex with ACE2 from different hosts has previously been determined<sup>3,11,12</sup>. These structures showed that SARS-CoV RBD contains a core and a receptor-binding motif (RBM); the RBM mediates contacts with ACE2. The surface of ACE2 contains two virus-binding hotspots that are essential for SARS-CoV binding. Several naturally selected mutations in the SARS-CoV RBM surround these hotspots and regulate the infectivity, pathogenesis, and cross-species and human-to-human transmissions of SARS-CoV<sup>3,11,12</sup>.

Because of the sequence similarity between the spike proteins of SARS-CoV and SARS-CoV-2, it was recently predicted that SARS-CoV-2

also uses ACE2 as its receptor<sup>13</sup>, which has been validated by other studies<sup>7,14–16</sup>. Here we determined the structural basis of receptor recognition by SARS-CoV-2 and compared the ACE2-binding affinity among SARS-CoV-2, SARS-CoV and RaTG13. Our findings identify the molecular and structural features of the SARS-CoV-2 RBM that result in tight ACE2 binding. They provide insights into the animal origin of SARS-CoV-2, and can help to guide intervention strategies that target SARS-CoV-2–ACE2 interactions.

To understand the structural basis of ACE2 recognition by SARS-CoV-2, we aimed to crystallize the SARS-CoV-2 RBD–ACE2 complex. Our strategy was informed by previous crystallization of the SARS-CoV RBD–ACE2 complex<sup>3</sup>. In this crystal form, the core of the SARS-CoV RBD (along with the ACE2 surface) was mainly involved in crystal lattice contact; the essential ACE2-binding residues in the SARS-CoV RBM were buried at the RBD–ACE2 interface and did not affect crystallization. To facilitate crystallization, we designed a chimeric RBD that uses the core from the SARS-CoV RBD as the crystallization scaffold and the RBM from SARS-CoV-2 as the functionally relevant unit (Fig. 1a and Extended Data Fig. 1). To further enhance crystallization, we improved the ACE2-binding affinity of the chimeric RBD by keeping a short loop from the SARS-CoV RBM, which maintains a strong salt bridge between Arg426 of the RBD and Glu329 of ACE2 (Extended Data Fig. 2a). This loop sits on the side of the binding interface, away from the main binding interface. We expressed and purified the chimeric RBD and ACE2, and crystallized the complex under the same conditions and in the same crystal form as those used for the SARS-CoV RBD–ACE2 complex. On the basis of X-ray diffraction data,

<sup>1</sup>Department of Veterinary and Biomedical Sciences, University of Minnesota, Saint Paul, MN, USA. <sup>2</sup>Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota, Minneapolis, MN, USA. <sup>3</sup>These authors contributed equally: Jian Shang, Gang Ye, Ke Shi, Yushun Wan. ✉e-mail: lifang@umn.edu



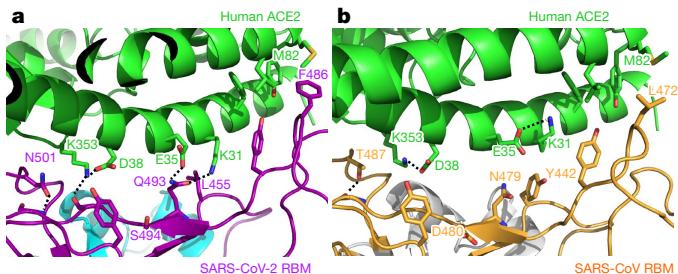
**Fig. 1 | Structure of the SARS-CoV-2 chimeric RBD complexed with ACE2.** **a**, Crystal structure of the SARS-CoV-2 chimeric RBD complexed with ACE2. ACE2 is shown in green. The RBD core is shown in cyan. The RBM is shown in magenta. A side loop in RBM is shown in orange. A zinc ion in ACE2 is shown in blue. **b**, Comparison of the conformations of the ridge in SARS-CoV-2 RBM (magenta) and SARS-CoV RBM (orange). **c**, Comparison of the conformations of the ridge from another viewing angle. In the SARS-CoV RBM, a proline-proline-alanine motif is shown. In the SARS-CoV-2 RBM, a newly formed hydrogen bond, Phe486, and some of the interactions of the ridge with the N-terminal helix of ACE2 are shown.

we determined the structure of the chimeric RBD–ACE2 complex by molecular replacement using the structure of the SARS-CoV RBD–ACE2 complex as the search template. We refined the structure to 2.68 Å (Extended Data Table 1 and Extended Data Fig. 3). The structure of this chimeric RBD–ACE2 complex, particularly in the RBM region, is highly similar to another recently determined structure of the SARS-CoV-2 wild-type RBD–ACE2 complex<sup>17</sup>, confirming that the chimeric RBD is a successful design.

The overall structure of the chimeric RBD–ACE2 complex is similar to that of the SARS-CoV RBD–ACE2 complex (Fig. 1a). Similar to the SARS-CoV RBM, SARS-CoV-2 RBM forms a gently concave surface with a ridge on one side; it binds to the exposed outer surface of the claw-like structure of ACE2 (Fig. 1a). The strong salt bridge between SARS-CoV RBD and ACE2 became a weaker (as judged by the longer distance of the interaction), but still energetically favourable, N–O bridge between Arg439 of the chimeric RBD and Glu329 of ACE2<sup>18</sup> (Extended Data Fig. 2b). In comparison to the SARS-CoV RBM, the SARS-CoV-2 RBM forms a larger binding interface and more contacts with ACE2 (Extended Data Fig. 4a, b). Our structural model also contained glycans attached to four ACE2 sites and one RBD site (Extended Data Fig. 5a). The glycan attached to Asn90 of ACE2 forms a hydrogen bond with Arg408 of the RBD core (Extended Data Fig. 5b); this glycan-interacting arginine is conserved between SARS-CoV-2 and SARS-CoV (Extended Data Fig. 1). The overall structural similarity in ACE2 binding by SARS-CoV-2 and SARS-CoV supports a close evolutionary relationship between the two viruses.

We measured the binding affinities between each of the three RBDs (SARS-CoV-2, chimeric and SARS-CoV) and ACE2 using surface plasmon resonance (SPR) (Extended Data Figs. 4c, 6). We found that the chimeric RBD has a higher ACE2-binding affinity than the SARS-CoV-2 RBD, consistent with the introduced N–O bridge between the chimeric RBD and ACE2. Both the chimeric and SARS-CoV-2 RBDs have significantly higher ACE2-binding affinities than the SARS-CoV RBD. These dissociation constant  $K_d$  values are consistent with other SPR studies<sup>12,19</sup>, although the exact  $K_d$  values vary depending on the specific approaches of each SPR experiment (Extended Data Table 2). Here we investigate the structural differences between the RBMs of SARS-CoV-2 and SARS-CoV that account for their different ACE2-binding affinities.

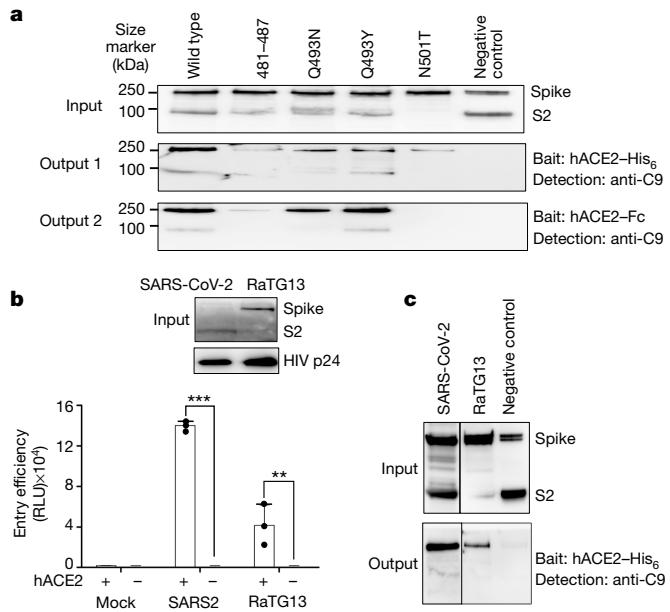
A marked structural difference between the RBMs of SARS-CoV-2 and SARS-CoV is the conformation of the loops in the ACE2-binding ridge (Fig. 1b, c). In both RBMs, one of the ridge loops contains an essential disulfide bond and the region between the disulfide-bond-forming cysteines is variable (Fig. 1c and Extended Data Fig. 1). Specifically, human and civet SARS-CoV strains and bat coronavirus RS3367 all contain a three-residue motif proline-proline-alanine in this loop;



**Fig. 2 | Structural details at the interface between the SARS-CoV-2 RBM and ACE2.** **a**, The interface between the SARS-CoV-2 RBM and ACE2. **b**, The interface between SARS-CoV RBM and ACE2.

the tandem prolines allow the loop to take a sharp turn. By contrast, SARS-CoV-2 and bat coronavirus RaTG13 both contain a four-residue motif glycine-valine/glutamine-glutamate/threonine-glycine; the two relatively bulky residues and two flexible glycines enable the loop to take a different conformation (Fig. 1c and Extended Data Fig. 1). Because of these structural differences, an additional main-chain hydrogen bond forms between Asn487 and Ala475 in the SARS-CoV-2 RBM, causing the ridge to take a more compact conformation and the loop containing Ala475 to move closer to ACE2 (Fig. 1c). As a consequence, the ridge in the SARS-CoV-2 RBM forms more contacts with the N-terminal helix of ACE2 (Extended Data Fig. 4b). For example, the N-terminal residue Ser19 of ACE2 forms a new hydrogen bond with the main chain of Ala475 of the SARS-CoV-2 RBM, and Gln24 in the N-terminal helix of ACE2 also forms a new contact with the SARS-CoV-2 RBM (Fig. 1c and Extended Data Fig. 4b). Moreover, compared with the corresponding Leu472 of the SARS-CoV RBM, Phe486 of the SARS-CoV-2 RBM points in a different direction and inserts into a hydrophobic pocket involving Met82, Leu79 and Tyr83 of ACE2 (Figs. 1c, 2a, b). In comparison to the SARS-CoV RBM, these structural changes in the SARS-CoV-2 RBM are more favourable for ACE2 binding.

In comparison to the SARS-CoV RBM–ACE2 interface, subtle yet functionally important structural changes take place near the two virus-binding hotspots at the SARS-CoV-2 RBM–ACE2 interface (Fig. 2a, b). At the SARS-CoV–ACE2 interface, two virus-binding hotspots were previously identified<sup>11,12</sup>: hotspot Lys31 (that is, hotspot 31) consists of a salt bridge between Lys31 and Glu35, and hotspot Lys353 (that is, hotspot 353) consists of a salt bridge between Lys353 and Asp38. Both salt bridges are weak, as judged by the relatively long distance of these interactions. Burial of these weak salt bridges in hydrophobic environments on virus binding would enhance their energy, owing to a reduction in the dielectric constant. This process is facilitated by interactions between the hotspots and nearby RBD residues. First, at the SARS-CoV RBM–ACE2 interface, hotspot 31 requires support from Tyr442 of the SARS-CoV RBM (Fig. 2b). In comparison, at the SARS-CoV-2 RBM–ACE2 interface, Leu455 of the SARS-CoV-2 RBM (corresponding to Tyr442 of the SARS-CoV RBM) has a less bulky side chain, providing less support to Lys31 of ACE2. As a result, the structure of hotspot 31 has rearranged: the salt bridge between Lys31 and Glu35 breaks apart, and each of the residues forms a hydrogen bond with Gln493 of the SARS-CoV-2 RBM (Fig. 2a). Second, at the SARS-CoV RBM–ACE2 interface, hotspot 353 requires support from the side-chain methyl group of Thr487 of the SARS-CoV RBM, whereas the side-chain hydroxyl group of Thr487 forms a hydrogen bond with the RBM main chain (which fixes the conformation of the Thr487 side chain) (Fig. 2b). In comparison, at the SARS-CoV-2 RBM–ACE2 interface, Asn501 of the SARS-CoV-2 RBM also has its conformation fixed through a hydrogen bond between its side chain and the RBM main chain; correspondingly, its side chain provides less support to hotspot 353 than the corresponding Thr487 of the SARS-CoV RBM does (Fig. 2a). Consequently, Lys353



**Fig. 3 | Biochemical data showing the interactions between SARS-CoV-2 or bat RaTG13 spike and ACE2.** **a**, Protein pull-down assay using ACE2 as the bait and cell-associated SARS-CoV-2 spike molecules (wild type and mutants) as the targets. Top, cell-expressed SARS-CoV-2 spike. Middle, pull-down results using His<sub>6</sub>-tagged ACE2. Bottom, pull-down results using Fc-tagged ACE2. MERS-CoV spike was used as a negative control. **b**, Entry of SARS-CoV-2 and bat RaTG13 pseudoviruses into ACE2-expressing cells. Top, packaged SARS-CoV-2 and bat RaTG13 pseudoviruses. HIV p24 was detected as an internal control. Bottom, pseudovirus entry efficiency. Mock, no pseudoviruses. Data are mean ± s.d. A comparison (two-tailed Student's *t*-test) between SARS-CoV-2 with ACE2 (*n* = 3 independent samples) and SARS-CoV-2 without ACE2 (*n* = 4 independent samples) showed a significant difference ( $P < 1.16 \times 10^{-8}$ ). A comparison (two-tailed Student's *t*-test) between RaTG13 with ACE2 (*n* = 3 independent samples) and RaTG13 without ACE2 (*n* = 4 independent samples) showed a significant difference,  $P = 0.0097$ . Individual data points are shown as black dots. \*\*\* $P < 0.001$ ; \*\* $P < 0.01$ . **c**, Protein pull-down assay using ACE2 as the bait and cell-associated RaTG13 spike as the target. All experiments were repeated independently three times with similar results.

of ACE2 takes a slightly different conformation, forming a hydrogen bond with the main chain of the SARS-CoV-2 RBM while maintaining the salt bridge with Asp38 of ACE2 (Fig. 2a). Thus, both hotspots have adjusted to the reduced support from nearby RBD residues, yet still become well-stabilized at the SARS-CoV-2 RBM–ACE2 interface.

To corroborate the structural observations, we characterized ACE2-binding affinities of the SARS-CoV-2 spike that contains mutations in critical ACE2-interacting residues. To this end, protein pull-down assays were performed, with purified recombinant ACE2 as the bait and cell-associated SARS-CoV-2 spike as the target (Fig. 3a). For cross-validation, we used ACE2 with two different tags, His<sub>6</sub> and Fc. The SARS-CoV-2 spike contained one of the following RBM changes: 481–487 (481-NGVEGFN-487 in SARS-CoV-2 were mutated to TPPALN as in SARS-CoV), Q493N (Gln493 in SARS-CoV-2 was mutated to an asparagine as in human SARS-CoV), Q493Y (Gln493 in SARS-CoV-2 was mutated to a tyrosine as in bat RaTG13) and N501T (Asn501 in SARS-CoV-2 was mutated to a threonine as in human SARS-CoV). The results showed that all of these introduced mutations reduced the ACE2-binding affinity of the SARS-CoV-2 spike. They confirm that the structural features of the SARS-CoV-2 RBM, including the ACE2-binding ridge and the hotspots-stabilizing residues, all contribute to the high ACE2-binding affinity of SARS-CoV-2.

Having compared ACE2 recognition by SARS-CoV-2 and SARS-CoV, we further investigated human ACE2 binding by bat RaTG13. To this end,

we performed a pseudovirus entry assay in which retroviruses pseudotyped with RaTG13 spike (that is, RaTG13 pseudoviruses) were used to enter ACE2-expressing human cells (Fig. 3b). The results showed that RaTG13 pseudovirus entry into the cells depends on ACE2. Additionally, RaTG13 spike was not cleaved on the pseudovirus surface. SARS-CoV-2 pseudovirus entry also depends on ACE2, but its spike was cleaved to S2 on the pseudovirus surface (probably because of a furin site insertion<sup>16</sup>) (Fig. 3b). Moreover, we performed a protein pull-down assay using ACE2 as the bait and cell-associated RaTG13 spike as the target (Fig. 3c). We found that the RaTG13 spike was pulled down by ACE2. Therefore, similar to SARS-CoV-2, bat RaTG13 binds to human ACE2 and can use human ACE2 as its entry receptor.

The current SARS-CoV-2 outbreak has become a global pandemic. Previous structural studies on SARS-CoV have established receptor recognition as an important determinant of SARS-CoV infectivity, pathogenesis and host range<sup>9</sup>. On the basis of the structural information presented here, along with biochemical data, we discuss the receptor recognition and evolution of SARS-CoV-2.

We will first discuss how well SARS-CoV-2 recognizes ACE2 in comparison to SARS-CoV. We show that, compared with SARS-CoV, the SARS-CoV-2 RBM contains structural changes in the ACE2-binding ridge, largely caused by a four-residue motif (residues 482–485: Gly-Val-Glu-Gly). This structural change allows the ridge to become more compact and form better contacts with the N-terminal helix of ACE2 (Fig. 1b, c). In addition, Phe486 of the SARS-CoV-2 RBM inserts into a hydrophobic pocket (Fig. 1c). The corresponding residue in the SARS-CoV RBM is a leucine, which probably forms a weaker contact with ACE2 owing to its smaller side chain. Finally, both virus-binding hotspots are more stabilized at the RBM–ACE2 interface through interactions with the SARS-CoV-2 RBM. As previous studies have shown<sup>11,12</sup>, these hotspots on ACE2 are important for coronavirus binding, because they involve two lysine residues that need to be accommodated properly in hydrophobic environments. Neutralizing the charges of the lysines is key to the binding of coronavirus RBDs to ACE2. The SARS-CoV-2 RBM has evolved strategies to stabilize the two hotspots: Gln493 and Leu455 stabilize hotspot 31, whereas Asn501 stabilizes hotspot 353 (Fig. 2a). Our biochemical data confirm that the SARS-CoV-2 RBD has a significantly higher ACE2-binding affinity than the SARS-CoV RBD and that the above structural features of the SARS-CoV-2 RBM contribute to the high ACE2-binding affinity of SARS-CoV-2 RBD (Fig. 3a). Thus, both structural and biochemical data reveal that the SARS-CoV-2 RBD recognizes ACE2 better than SARS-CoV RBD does.

Next, we investigated how SARS-CoV-2 may have been transmitted from bats to humans. First, we found that bat RaTG13 uses human ACE2 as its receptor (Fig. 3b, c), suggesting that RaTG13 may infect humans. Second, as with SARS-CoV-2, bat RaTG13 RBM contains a similar four-residue motif in the ACE2-binding ridge, supporting the notion that SARS-CoV-2 may have evolved from RaTG13 or a RaTG13-related bat coronavirus (Extended Data Table 3 and Extended Data Fig. 7). Third, the L486F, Y493Q and D501N residue changes from RaTG13 to SARS-CoV-2 enhance ACE2 recognition and may have facilitated the bat-to-human transmission of SARS-CoV-2 (Extended Data Table 3 and Extended Data Fig. 7). A lysine-to-asparagine mutation at the 479 position in the SARS-CoV RBD (corresponding to the 493 position in the SARS-CoV-2 RBD) enabled SARS-CoV to infect humans<sup>3</sup>. Fourth, Leu455 contributes favourably to ACE2 recognition, and it is conserved between RaTG13 and SARS-CoV-2; its presence in the SARS-CoV-2 RBM may be important for the bat-to-human transmission of SARS-CoV-2 (Extended Data Table 3 and Extended Data Fig. 7). Host and viral factors other than receptor recognition also have important roles in the cross-species transmission of coronaviruses<sup>20,21</sup>. Nevertheless, the identified receptor-binding features of the SARS-CoV-2 RBM may have facilitated SARS-CoV-2 to transmit from bats to humans (Extended Data Fig. 7).

# Article

We then examined whether intermediate hosts were involved in the potential bat-to-human transmission of SARS-CoV-2. Because bat coronavirus RaTG13 binds to human ACE2, one possibility is that there is no intermediate host. Alternatively, pangolins have been proposed to be an intermediate host<sup>22</sup>. The structural information provided in this study enables us to inspect and understand the important RBM residues in coronaviruses isolated from pangolins. Two coronaviruses, CoV-pangolin/GD and CoV-pangolin/GX, have been isolated from pangolins from two different locations in China: Guangdong (GD) and Guangxi (GX), respectively. The RBM of the CoV-pangolin/GD contains Leu455, the 482–485 loop, Phe486, Gln493 and Asn501 (Extended Data Table 3), all of which are favourable for ACE2 recognition. The RBM of CoV-pangolin/GX contains Leu455 and the 482–485 loop, both of which are favourable for ACE2 recognition, and it also contains Leu486, Glu493 and Thr501 (Extended Data Table 3), all of which are less favourable for ACE2 recognition. Therefore, CoV-pangolin/GD potentially recognizes human ACE2 well, whereas CoV-pangolin/GX does not. Hence, pangolins from Guangdong, but not pangolins from Guangxi, could potentially pass coronaviruses to humans. However, many other factors determine the cross-species transmission of coronaviruses<sup>20,21</sup>, and the above analysis will need to be verified experimentally.

Finally, this study helps to inform intervention strategies. First, neutralizing monoclonal antibodies that target the SARS-CoV-2 RBM can prevent the virus from binding to ACE2, and are therefore promising antiviral drugs. Our structure has laid out all of the functionally important epitopes in the SARS-CoV-2 RBM that can potentially be targeted by neutralizing antibody drugs. Thus, this study can help to guide the development and optimization of these antibody drugs. Second, the RBD itself can function as a subunit vaccine<sup>10,23</sup>. The functionally important epitopes in the SARS-CoV-2 RBM that were identified in this study can guide structure-based design of highly efficacious RBD vaccines. Such a structure-based strategy for subunit vaccine design has previously been developed<sup>24</sup>. This strategy may be helpful in designing SARS-CoV-2 RBD vaccines. Overall, this study can help to inform structure-based intervention strategies that target receptor recognition by SARS-CoV-2.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-2179-y>.

- Li, Q. et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N. Engl. J. Med.* **382**, 1199–1207 (2020).
- Huang, C. et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **395**, 497–506 (2020).
- Li, F., Li, W., Farzan, M. & Harrison, S. C. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. *Science* **309**, 1864–1868 (2005).
- Li, W. et al. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* **426**, 450–454 (2003).
- Lee, N. et al. A major outbreak of severe acute respiratory syndrome in Hong Kong. *N. Engl. J. Med.* **348**, 1986–1994 (2003).
- Peiris, J. S. M. et al. Coronavirus as a possible cause of severe acute respiratory syndrome. *Lancet* **361**, 1319–1325 (2003).
- Zhou, P. et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273 (2020).
- Perlman, S. & Netland, J. Coronaviruses post-SARS: update on replication and pathogenesis. *Nat. Rev. Microbiol.* **7**, 439–450 (2009).
- Li, F. Structure, function, and evolution of coronavirus spike proteins. *Annu. Rev. Virol.* **3**, 237–261 (2016).
- Du, L. et al. The spike protein of SARS-CoV — a target for vaccine and therapeutic development. *Nat. Rev. Microbiol.* **7**, 226–236 (2009).
- Li, F. Structural analysis of major species barriers between humans and palm civets for severe acute respiratory syndrome coronavirus infections. *J. Virol.* **82**, 6984–6991 (2008).
- 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.* **287**, 8904–8911 (2012).
- 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.* **94**, e00127-20 (2020).
- Letko, M., Marzi, A. & Munster, V. Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nat. Microbiol.* **5**, 562–569 (2020).
- Hoffmann, M. et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* <https://doi.org/10.1016/j.cell.2020.02.052> (2020).
- Walls, A. C. et al. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* <https://doi.org/10.1016/j.cell.2020.02.058> (2020).
- Lan, J. et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* <https://doi.org/10.1038/s41586-020-2180-5> (2020).
- Pylaeva, S., Brehm, M. & Sebastiani, D. Salt bridge in aqueous solution: strong structural motifs but weak enthalpic effect. *Sci. Rep.* **8**, 13626 (2018).
- Wrapp, D. et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**, 1260–1263 (2020).
- Cui, J., Li, F. & Shi, Z. L. Origin and evolution of pathogenic coronaviruses. *Nat. Rev. Microbiol.* **17**, 181–192 (2019).
- Yang, Y. et al. Receptor usage and cell entry of bat coronavirus HKU4 provide insight into bat-to-human transmission of MERS coronavirus. *Proc. Natl. Acad. Sci. USA* **111**, 12516–12521 (2014).
- Lam, T. T. et al. Identifying SARS-CoV-2 related coronaviruses in Malayan pangolins. *Nature* <https://doi.org/10.1038/s41586-020-2169-0> (2020).
- Du, L. et al. MERS-CoV spike protein: a key target for antivirals. *Expert Opin. Ther. Targets* **21**, 131–143 (2017).
- Du, L. et al. Introduction of neutralizing immunogenicity index to the rational design of MERS coronavirus subunit vaccines. *Nat. Commun.* **7**, 13473 (2016).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

## Methods

### Data reporting

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

### Plasmids

SARS-CoV-2 spike (GenBank accession number QHD43416.1), SARS-CoV spike (GenBank accession number AFR58740.1), RaTG13 spike (GenBank accession number QHR63300.2) and ACE2 (GenBank accession number NM\_021804) were all synthesized (GenScript Biotech). SARS-CoV-2, SARS-CoV, chimeric RBDs (see Extended Data Fig. 1 for residue ranges of RBDs) and ACE2 ectodomain (residues 1–615) were subcloned into pFastBac vector (Life Technologies) with a N-terminal honeybee melittin signal peptide and a C-terminal His<sub>6</sub>-tag. The ACE2 ectodomain (residues 1–615) with a C-terminal Fc-tag was also constructed.

### Protein expression and purification

All of the proteins were prepared from Sf9 insect cells using the Bac-to-Bac system (Life Technologies) as previously described<sup>3</sup>. In brief, the His<sub>6</sub>-tagged proteins were collected from cell culture medium, purified using a Ni-NTA column, purified further using a Superdex200 gel filtration column (GE Healthcare) and stored in a buffer containing 20 mM Tris pH 7.2 and 200 mM NaCl. The Fc-tagged protein was purified in the same way as the His<sub>6</sub>-tagged proteins, except that the protein A column replaced the Ni-NTA column in the procedure.

### Crystallization and structure determination

To purify the RBD–ACE2 complex, ACE2 and RBD were incubated together, and the complex was purified using Superdex200 gel-filtration chromatography. RBD–ACE2 crystals were grown in sitting drops at room temperature over wells containing 100 mM Tris (pH 8.5), 18–20% PEG 6000 and 100 mM NaCl. Crystals were soaked briefly in 100 mM Tris (pH 8.5), 30% PEG 6000, 100 mM NaCl and 30% ethylene glycol before being flash-frozen in liquid nitrogen. X-ray diffraction data were collected at the Advanced Photon Source beamline 24-ID-E. The structure was determined by molecular replacement using the structure of SARS-CoV RBD complexed with ACE2 as the search template (Protein Data Bank (PDB) 2AJF). Structure data and refinement statistics are shown in Extended Data Table 1.

### Protein–protein binding assay

The SPR assays using a Biacore 2000 system (GE Healthcare) were carried out as described previously<sup>12</sup>. In brief, different RBDs were covalently immobilized to a CM5 sensor chip through their amine groups (GE Healthcare). The running buffer contained 10 mM HEPES pH 7.4, 150 mM NaCl, 3 mM EDTA and 0.05% Tween-20. Serial dilutions of purified recombinant ACE2 were injected ranging in concentration from 5 to 80 nM for the SARS-CoV-2 RBD and chimeric RBD, and from 20 to 320 nM for the SARS-CoV RBD. The resulting data were fit to a 1:1 binding model using Biacore Evaluation Software (GE Healthcare).

The protein pull-down assay was performed using a Dynabeads His-Tag Isolation and Pull-down kit (Invitrogen) and a Dynabeads Protein A for Immunoprecipitation kit (Invitrogen) according to the manufacturers' manual. In brief, 150 µl indicated Dynabeads were washed with phosphate-buffered saline (PBS) and incubated with either 5 µg ACE2–His<sub>6</sub> (ACE2 with a C-terminal His<sub>6</sub>-tag) or 5 µg ACE2–Fc (ACE2 with a C-terminal Fc-tag) on a roller at room temperature for 30 min. After incubation, ACE2-bound beads were washed three times with 1 ml PBST buffer (PBS and 0.05% Tween-20) on a roller for 10 min and then were aliquoted into different tubes for use. To prepare the cell-associated coronavirus spike protein, HEK293T cells were transfected with a pcDNA3.1(+) plasmid encoding coronavirus spike (containing a C-terminal C9-tag); 48 h after transfection, the spike-expressing cells

were lysed using a sonicator in immunoprecipitation assay buffer (20 mM Tris-HCl, pH 7.4, 150 mM NaCl, 1 mM EDTA and 1% Triton X-100, supplemented with protease inhibitors) and centrifuged at 12,000g for 2 min. The supernatants (containing solubilized SARS-CoV-2 spike) were transferred to mix with the ACE2-bound beads in 2-ml tubes separately (spike was in excess of ACE2). After a 1-h incubation on a roller at room temperature, beads were washed three times with PBST buffer and the bound proteins were eluted using elution buffer (300 mM imidazole, 50 mM sodium phosphate pH 8.0, 300 mM NaCl, 0.01% Tween-20 for ACE2–His<sub>6</sub>-bound beads; 0.1 M citric acid pH 2.7 for ACE2–Fc-bound beads). The samples were then subjected to SDS-PAGE and analysed by western blotting using an anti-C9 tag antibody.

### Coronavirus-spike-mediated pseudovirus entry assay

The pseudovirus entry assay was performed as described previously<sup>21</sup>. In brief, HEK293T cells were co-transfected with luciferase-expressing HIV-1 genome plasmid (pNL4-3.luc.RE) and a plasmid encoding SARS-CoV-2 spike or RaTG13 spike. Pseudoviruses were collected 72 h after transfection, and were used to enter recipient cells (HEK293T cells exogenously expressing ACE2). After incubation of pseudoviruses with recipient cells at 37 °C for 6 h, the medium was changed and cells were incubated for an additional 60 h. Cells were then washed with PBS buffer and lysed. Aliquots of cell lysates were transferred to Optiplate-96 (PerkinElmer), followed by the addition of luciferase substrate. Relative light units were measured using an EnSpire plate reader (PerkinElmer). All measurements were carried out on at least three independent biological samples.

### Analyses of protein contact residues and protein buried surface areas

Protein contact residues were analysed using the LigPlot<sup>+</sup> program (v.1.4.5) (<https://www.ebi.ac.uk/thornton-srv/software/LigPlus/>). Protein buried surface areas were analysed using PDBePISA tool (<http://pdbe.org/pisa/>).

### Reporting summary

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

### Data availability

Coordinates and structure factors have been deposited to the Protein Data Bank with accession number 6VW1.

25. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
26. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D* **75**, 861–877 (2019).
27. Winn, M. D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D* **67**, 235–242 (2011).
28. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
29. Sui, J. et al. Potent neutralization of severe acute respiratory syndrome (SARS) coronavirus by a human mAb to S1 protein that blocks receptor association. *Proc. Natl. Acad. Sci. USA* **101**, 2536–2541 (2004).
30. Li, W. et al. Receptor and viral determinants of SARS-coronavirus adaptation to human ACE2. *EMBO J.* **24**, 1634–1643 (2005).
31. Sun, C. et al. SARS-CoV-2 and SARS-CoV spike-RBD structure and receptor binding comparison and potential implications on neutralizing antibody and vaccine development. Preprint at <https://doi.org/10.1101/2020.02.16.951723> (2020).
32. Pesce, A. J. & Michael, J. G. Artifacts and limitations of enzyme immunoassay. *J. Immunol. Methods* **150**, 111–119 (1992).

**Acknowledgements** This work was supported by NIH grants R01AI089728 and R01AI110700 (to F.L.) and R35GM118047 (to H.A.) and is based on research conducted at the Northeastern Collaborative Access Team beamlines, which are supported by NIH grants P30GM124165 and S10OD021527, and DOE contract DE-AC02-06CH1357. We thank staff at Advanced Photon Source beamline 24-ID-E for assistance in data collection and Y. V. Jiang for statistical consultation and edits to the manuscript.

**Author contributions** J.S. conceptualized the project, expressed and purified proteins, performed crystallization, carried out protein pull-down experiments and the pseudovirus entry assay, and reviewed the manuscript. G.Y. performed crystallization, determined and refined the structure, analysed the structure, performed the SPR experiment, and reviewed the manuscript.

# Article

K.S. collected X-ray diffraction data, determined and refined the structure, analysed the structure, and reviewed the manuscript. Y.W. conceptualized the project, expressed and purified proteins, performed protein pull-down experiments and the pseudovirus entry assay, and reviewed the manuscript. C.L. performed protein pull-down experiments and the pseudovirus entry assay, and reviewed the manuscript. H.A. provided resources, analysed the structure, and reviewed the manuscript. Q.G. performed protein pull-down experiments and the pseudovirus entry assay, and reviewed the manuscript. A.A. expressed and purified proteins, and reviewed the manuscript. F.L. conceptualized and supervised the project, provided resources, guided the experiments and data analysis, and wrote the manuscript.

**Competing interests** The authors declare no competing interests.

**Additional information**

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-020-2179-y>.

**Correspondence and requests for materials** should be addressed to F.L.

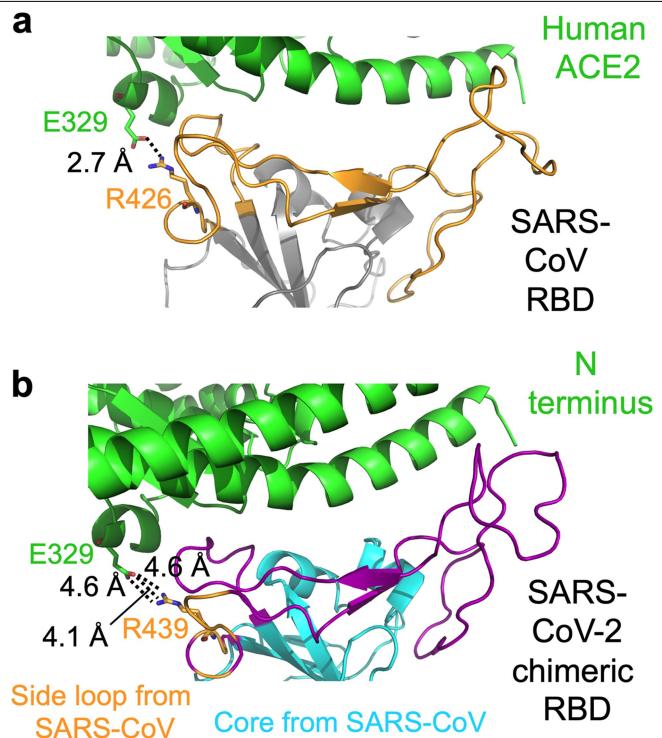
**Peer review information** *Nature* thanks Lijun Rong and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.

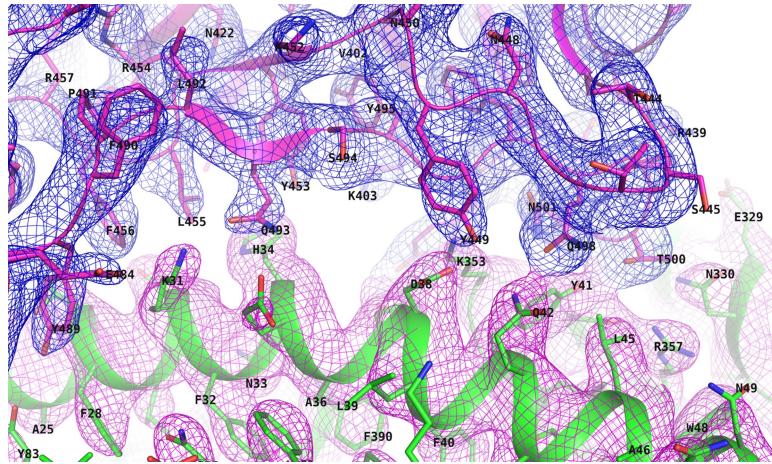
SARS-human	306	RVVPS	GDVVRFPPNIT	NLCPPGEVFN	ATKPFPSVYAW	ERKKISNCVA	DYSVLYNST 360
SARS-civet	306	RVVPS	GDVVRFPPNIT	NLCPPGEVFN	ATKPFPSVYAW	ERKRISNCVA	DYSVLYNST 360
CoV-pangolin/GD	319	RVQPT	ESIVRFPNIT	NLCPPGEVFN	ATTPFASVYAW	NRKRISNCVA	DYSVLYNST 373
CoV-pangolin/GX	319	RVQPT	ISIVRFPNIT	NLCPPGEVFN	ASKFASVYAW	NRKRISNCVA	DYSVLYNST 373
Rs3367-bat	307	RVAPS	KEVVRFPNIT	NLCPPGEVFN	ATTPFASVYAW	ERKRISNCVA	DYSVLYNST 361
RaTG13-bat	319	RVQPT	DSIVRFPNIT	NLCPPGEVFN	ATTPFASVYAW	NRKRISNCVA	DYSVLYNST 373
<b>SARS-CoV-2</b>	319	RVQPT	ESIVRFPNIT	NLCPPGEVFN	ATTPFASVYAW	NRKRISNCVA	DYSVLYNSTS 373
		**.*:	..:*****	***:*****	*.:*:*****	:*:*****	*****:***
SARS-human	FSTFKCYGVGS	ATKLNDLCLS	NVYADSFVVK	GDDVRQIAPG	QTGVIADYN	KLPDDFTMGCV	420
SARS-civet	FSTFKCYGVGS	ATKLNDLCLS	NVYADSFVVK	GDDVRQIAPG	QTGVIADYN	KLPDDFTMGCV	420
CoV-pangolin/GD	FSTFKCYGVGS	PTKLNDLCFT	NVYADSFVVK	GDEVRQIAPG	QTGRIADYN	KLPDDFTGCV	433
CoV-pangolin/GX	FSTFKCYGVGS	PTKLNDLCFT	NVYADSFVVK	GDEVRQIAPG	QTGVIADYN	KLPDDFTGCV	433
Rs3367-bat	FSTFKCYGVGS	ATKLNDLCLS	NVYADSFVVK	GDDVRQIAPG	QTGVIADYN	KLPDDFTGCV	421
RaTG13-bat	FSTFKCYGVGS	PTKLNDLCFT	NVYADSFVIT	GDEVRQIAPG	QTGKIADYN	KLPDDFTGCV	433
<b>SARS-CoV-2</b>	FSTFKCYGVGS	PTKLNDLCFT	NVYADSFVIR	GDEVRQIAPG	QTGKIADYN	KLPDDFTGCV	433
	*****:*****:	*****:*****:	*****:*****	*****:*****	*****:*****	*****:*****	*****:***
SARS-human	LAWNT <u>RHID</u> A T <u>S</u> T <u>G</u> NNYKY R <u>Y</u> LR <u>H</u> G <u>K</u> L <u>R</u> P FERDI <u>S</u> NV <u>V</u> P SP <u>D</u> G <u>K</u> C <u>T</u> P-P <u>A</u> L <u>N</u> C <u>Y</u> W <u>P</u> L <u>I</u> D 480						
SARS-civet	LAWNT <u>RHID</u> A T <u>S</u> T <u>G</u> NNYKY R <u>Y</u> LR <u>H</u> G <u>K</u> L <u>R</u> P FERDI <u>S</u> NV <u>V</u> P SP <u>D</u> G <u>K</u> C <u>T</u> P-P <u>A</u> L <u>N</u> C <u>Y</u> W <u>P</u> L <u>I</u> D 480						
CoV-pangolin/GD	I <u>A</u> W <u>N</u> S <u>N</u> L <u>D</u> S K <u>V</u> G <u>G</u> NNY <u>Y</u> L R <u>Y</u> LR <u>K</u> S <u>N</u> L <u>K</u> P FERDI <u>S</u> TE <u>I</u> Y QAG <u>S</u> T <u>P</u> C <u>N</u> GV <u>E</u> G <u>F</u> N <u>C</u> Y <u>F</u> P <u>L</u> QS 494						
CoV-pangolin/GX	I <u>A</u> W <u>N</u> S <u>Y</u> K <u>O</u> D <u>A</u> L <u>T</u> G <u>D</u> NN <u>Y</u> Y <u>L</u> R <u>Y</u> LR <u>K</u> S <u>K</u> L <u>K</u> P FERDI <u>S</u> TE <u>I</u> Y QAG <u>S</u> T <u>P</u> C <u>N</u> Q <u>V</u> G <u>L</u> N <u>C</u> Y <u>P</u> L <u>E</u> R 494						
Rs3367-bat	I <u>A</u> W <u>N</u> T <u>R</u> H <u>I</u> D <u>A</u> T <u>Q</u> T <u>G</u> NNY <u>Y</u> K Y <u>R</u> SL <u>R</u> H <u>G</u> K <u>L</u> R P FERDI <u>S</u> NV <u>V</u> P SP <u>D</u> G <u>K</u> C <u>T</u> P-P <u>A</u> F <u>N</u> C <u>Y</u> W <u>P</u> L <u>I</u> D 481						
RaTG13-bat	I <u>A</u> W <u>N</u> S <u>X</u> K <u>H</u> I <u>D</u> A K <u>E</u> G <u>G</u> NN <u>Y</u> Y <u>L</u> R <u>Y</u> LR <u>K</u> A <u>N</u> L <u>K</u> P FERDI <u>S</u> TE <u>I</u> Y QAG <u>S</u> K <u>P</u> C <u>N</u> Q <u>T</u> G <u>L</u> N <u>C</u> Y <u>P</u> L <u>Y</u> 494						
<b>SARS-CoV-2</b>	I <u>A</u> W <u>N</u> S <u>N</u> L <u>D</u> S K <u>V</u> G <u>G</u> NNY <u>Y</u> L R <u>Y</u> LR <u>K</u> S <u>N</u> L <u>K</u> P FERDI <u>S</u> TE <u>I</u> Y QAG <u>S</u> T <u>P</u> C <u>N</u> GV <u>E</u> G <u>F</u> N <u>C</u> Y <u>F</u> P <u>L</u> QS 494						
	*****:.*: . ..:.*:.* . :*:.*:.*: . *****:.*: . :*:.*:.*: . :*:.*: . :*:.*: . :*:.*: . :*:.*: .						
SARS-human	YGFYTT <u>TG</u> IG YQP <u>Y</u> RVVVL <u>S</u> FELLNAPAT <u>V</u> CGPKL 515						
SARS-civet	YGFYTT <u>S</u> G <u>I</u> G YQP <u>Y</u> RVVVL <u>S</u> FELLNAPAT <u>V</u> CGPKL 515						
CoV-pangolin/GD	YGFHPT <u>NG</u> VG YQP <u>Y</u> RVVVL <u>S</u> FELLNAPAT <u>V</u> CGPKQ 529						
CoV-pangolin/GX	YGFHPT <u>T</u> GV <u>N</u> YQP <u>Y</u> RVVVL <u>S</u> FELLNAPAT <u>V</u> CGPKL 229						
Rs3367-bat	YGFYIT <u>NG</u> IG YQP <u>Y</u> RVVVL <u>S</u> FELLNAPAT <u>V</u> CGPKL 516						
RaTG13-bat	YGFYPT <u>D</u> GV <u>G</u> HQP <u>Y</u> RVVVL <u>S</u> FELLNAPAT <u>V</u> CGPKK 529						
<b>SARS-CoV-2</b>	YGFQPT <u>NG</u> VG YQP <u>Y</u> RVVVL <u>S</u> FELLHAPAT <u>V</u> CGPKK 529						
	*****:.*: . :*:*****: . :*:*****: . :*:*****: . :*:*****: . :*:*****: . :*:*****: .						

**Extended Data Fig. 1 | Sequence alignment of the RBDs from SARS-CoV and SARS-like viruses.** RBM is shown in magenta. Previously identified critical ACE2-binding residues are shown in blue. The seven RBM residues that differ between the SARS-CoV-2 wild-type RBD and SARS-CoV-2 chimeric RBD are shaded. A critical arginine on the side loop of the SARS-CoV RBM that forms a strong salt bridge with ACE2 is shown in green. Another arginine in the core structure that interacts with glycan is shown in cyan. The residues on the variable loop between two disulfide-bond-forming cysteines in the

ACE2-binding ridge are shown in red. The important motif changes in the ACE2-binding ridge are underlined. GenBank accession numbers are: QHD43416.1 for SARS-CoV-2 spike; AFR58742 for human SARS-CoV spike; AY304486.1 for civet SARS-CoV spike; MG916901.1 for bat Rs3367 spike; QHR63300.2 for bat RaTG13 spike. Two coronaviruses, CoV-pangolin/GD and CoV-pangolin/GX, were isolated from pangolins at two different locations in China, Guangdong and Guangxi; their RBD sequences were from a previous study<sup>22</sup>.



**Extended Data Fig. 2 | Interface between SARS-CoV-2 or SARS-CoV RBM and ACE2.** **a**, The interface between the SARS-CoV RBD and ACE2, showing a strong salt bridge between Arg426 on the side loop of the RBM and Glu329 of ACE2. The core structure is shown in grey. The RBM is shown in orange. **b**, The interface between the SARS-CoV-2 chimeric RBD and ACE2, showing a weaker, but still energetically favourable, N–O bridge between Arg439 on the side loop of the RBM and Glu329 of ACE2. The interaction between Arg439 on the side loop of the RBM and Glu329 of ACE2 is non-natural in SARS-CoV-2 (and is a result of the design of the SARS-CoV-based chimaera).



**Extended Data Fig. 3 | Composite omit map of the interface between the SARS-CoV-2 RBM and ACE2.** Contour level is 1 $\sigma$ .

# Article

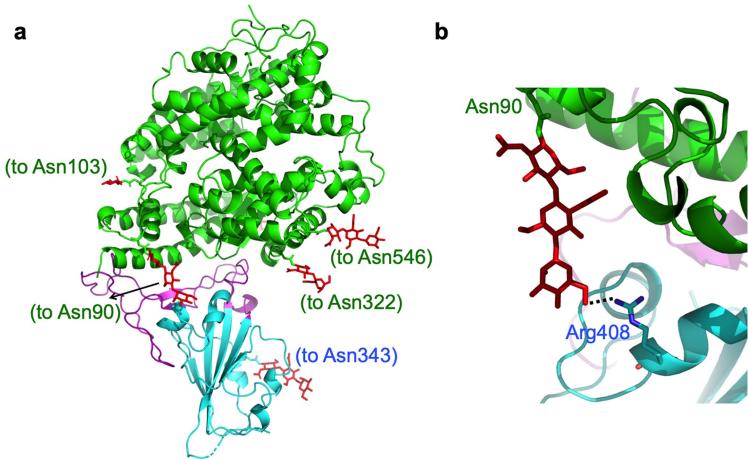
<b>a</b>	Buried surface (Å <sup>2</sup> )	Complex 1	Complex 2	<b>c</b>	RBD	K <sub>d</sub> (nM)	k <sub>off</sub> (s <sup>-1</sup> )	k <sub>on</sub> (M <sup>-1</sup> s <sup>-1</sup> )
SARS-CoV-2	895.9	860.9		SARS-CoV-2 (wild type)	<b>44.2</b>	7.75 × 10 <sup>-3</sup>	1.75 × 10 <sup>5</sup>	
SARS-CoV-2 (chimeric)	924.2	883.5		SARS-CoV-2 (chimeric)	<b>23.2</b>	4.23 × 10 <sup>-3</sup>	1.82 × 10 <sup>5</sup>	
SARS-CoV	849.2	829.7		SARS-CoV	<b>185</b>	3.70 × 10 <sup>-2</sup>	2.01 × 10 <sup>5</sup>	

<b>b</b>		19	24	27	28	31	34	35	37	38	41	42	45	79	82	83	325	329	330	353	354	355	357	human ACE2
S	Q	T	F	K	H	E	E	D	Y	Q	L	L	M	Y	Q	E	N	K	G	D	R			
N473	Y475	Y475	Y442	Y440			Y436	Y484	Y436	Y484			L472	N473	I489	R426	T486	G482	G488	T486	T486	SARS		
																						G488 Y491		
																						T487		
																						Y491		
A475	G476	F456	Y489	F456	Y453	Q493	Y505	Y449	Q498	Q498	Q498	Q498	F486	F486	F486	N487							<b>R439</b>	
N487	A475				Q493	L455							T500				T500	G496	G502	T500	T500	SARS-2		
																		N501	Y505				chimeric	
																						G502		
																						Y505		

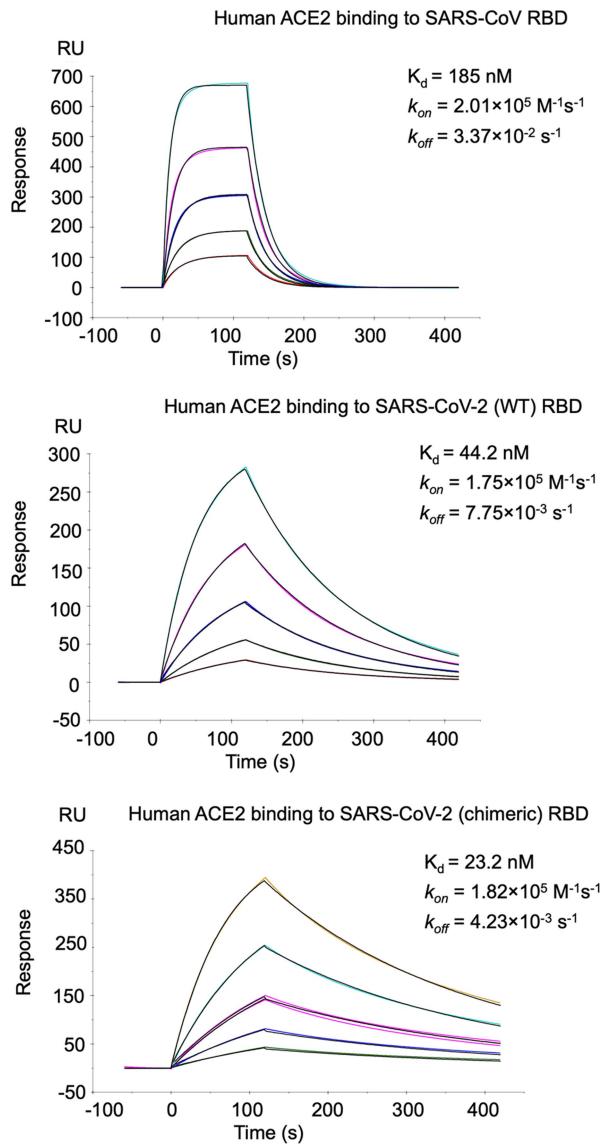
**Extended Data Fig. 4 | Comparison of ACE2 binding by the SARS-CoV RBD, SARS-CoV-2 wild-type RBD and SARS-CoV-2 chimeric RBD.** **a**, Buried surface areas at SARS-CoV RBM–ACE2 and SARS-CoV-2 RBM–ACE2 interfaces. In the crystals for both the SARS-CoV RBD–ACE2 complex and chimeric RBD–ACE2 complex, two copies of each complex were present in one asymmetric unit. Numbers for both copies of the complexes are shown. The interaction between

Arg439 on the side loop of the RBM and Glu329 of ACE2 was excluded from the calculation of the buried surface area of SARS-CoV-2. **b**, List of contact residues from RBM and ACE2 that are directly involved in RBM–ACE2 binding. The engineered Arg439 in the chimeric RBD is shown in orange. Contact residues of the SARS-CoV RBM–ACE2 complex are taken from PDB 2AJF. **c**, Binding affinities between the RBDs and ACE2 measured using SPR.



**Extended Data Fig. 5 | Glycans built into the SARS-CoV-2 chimeric RBD–ACE2 structure.** **a**, Distribution of glycans in the structure. Glycans are shown in red. The residues to which the glycans attach are indicated in parentheses. **b**, Interaction between a glycan attached to Asn90 of ACE2 and Arg408 from the RBD.

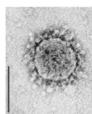
# Article



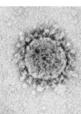
**Extended Data Fig. 6 | Measurement of the binding affinities between the RBDs and ACE2 by SPR assay using Biacore.** Purified recombinant RBDs were covalently immobilized on the sensor chip through their amine groups and purified recombinant ACE2 flowed over the RBDs. ACE2 was diluted to different concentrations (from 5 to 80 nM for SARS-CoV-2 RBD and chimeric RBD, and from 20 to 320 nM for SARS-CoV RBD) before being injected. The resulting data were fit to a 1:1 binding model. Each experiment was repeated independently twice with similar results. Each time, five different protein concentrations were used to calculate the  $K_d$  values.

# SARS-CoV-2: Where did it come from?

Bats: the natural reservoir?



Evolution of RBM

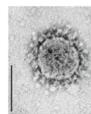


SARS-CoV-2 RBM residues	Adapted to what human ACE2 residues?
GVEG (482-485)	N-terminal helix of human ACE2
F486	M82 in human ACE2
Q493/L455	K31 and E35 in human ACE2
N501	K353 in human ACE2

Possible interim reservoir

2019-?:  
Over 2 million infections; over 100,000 fatalities; highly infectious

Humans



**Extended Data Fig. 7 | Summary of ACE2 adaptation and evolution of SARS-CoV-2.** Several structural features of the SARS-CoV-2 RBM contribute favourably to the ability of the virus to bind to human ACE2. Each of these

structural features of the SARS-CoV-2 RBM matches well with one or more structural features of human ACE2. This figure establishes the correlations among these structural features of SARS-CoV-2 RBM and ACE2.

# Article

**Extended Data Table 1 | Crystallization data collection and refinement statistics**

## Data collection

Space group	P12 <sub>1</sub> 1
Unit cell dimensions	
<i>a, b, c</i> (Å)	80.435, 118.034, 112.075
$\alpha, \beta, \gamma$ (°)	90, 93.12, 90
Resolution (Å)	59–2.68 (2.78–2.68)*
$R_{\text{sym}}$ or $R_{\text{merge}}$	0.0807 (1.47)
$I / \sigma I$	12.08 (1.12)
Completeness (%)	98.96 (98.97)
Redundancy	3.9 (4.0)

## Refinement

Resolution (Å)	59–2.68 (2.78–2.68)*
No. reflections	58219 (5774)
$R_{\text{work}} / R_{\text{free}}$	0.197/0.228
No. atoms	13180
Protein	12782
Ligand/ion	372
Water	26
<i>B</i> -factors	108.85
Protein	108.05
Ligand/ion	138.01
Water	82.32
R.m.s. deviations	
Bond lengths (Å)	0.003
Bond angles (°)	0.60

Data processing was carried out using HKL2000<sup>25</sup>. Molecular replacement and model refinement were performed using PHENIX and CCP4<sup>26,27</sup>. Model building was carried out using COOT<sup>28</sup>. Structural figures were made using PYMOL (The PyMOL Molecular Graphics System, v.2.0 Schrödinger). We used 26 crystals for X-ray data collection. Each crystal resulted in one set of X-ray data. The best dataset (as judged by data statistics) was used for structure determination and refinement.

\*Values in parentheses are for the highest-resolution shell.

**Extended Data Table 2 | Summary of spike–ACE2 binding affinities measured by different studies**

Protein coated	K <sub>d</sub> (M)	Coating method	Detection method	References
SARS-CoV-S1 -Fc tag	1.7×10 <sup>-9</sup>	Covalently immobilized (via amine group) to sensor chip	SPR	29
SARS-CoV-RBD -Fc tag	1.62×10 <sup>-8</sup>	Non-covalently immobilized (via Fc tag) to sensor chip	SPR	30
<b>SARS-CoV-RBD -His tag</b>	<b>1.52×10<sup>-7</sup></b>	<b>Covalently immobilized (via amine group) to sensor chip</b>	SPR	12
<b>Human ACE2 -His tag</b>	<b>2.09 ×10<sup>-8</sup></b>	<b>Covalently immobilized (via amine group) to sensor chip</b>	SPR	
SARS-CoV-RBD -His tag	3.258×10 <sup>-7</sup>	Non-covalently immobilized (via His tag) to sensor chip	SPR	19
SARS-CoV-2 -spike-His tag	1.47×10 <sup>-8</sup>	Non-covalently immobilized (via His tag) to sensor chip	SPR	
SARS-CoV-spike -His tag	7.7 ×10 <sup>-9</sup>	Non-covalently immobilized (via His tag) to sensor chip	Blitz	16
SARS-CoV-2 -spike-His tag	2.9 ×10 <sup>-9</sup>	Non-covalently immobilized (via His tag) to sensor chip	Blitz	
SARS-CoV-S1	Similar binding affinity	Serial dilution coated on plates	ELISA	31
SARS-CoV-2-S1		Serial dilution coated on plates	ELISA	
<b>SARS-CoV-RBD -His tag</b>	<b>1.85×10<sup>-7</sup></b>	<b>Covalently immobilized (via amine group) to sensor chip</b>	SPR	Current study
SARS-CoV-2 -RBD-His tag	4.42×10 <sup>-8</sup>	Covalently immobilized (via amine group) to sensor chip	SPR	

Protein–protein binding affinities are more accurately measured using SPR than using enzyme-linked immunosorbent assay (ELISA)<sup>12,16,19,29–31</sup>, as ELISA often causes artefacts in protein binding<sup>32</sup>. K<sub>d</sub> values measured using SPR depend on how the proteins are coated. Non-covalently immobilized proteins using Fc or His tags (on the opposite side to ligand-binding sites) have the advantage over covalently immobilized proteins using amine groups because the former have the ligand-binding sites fully exposed. However, non-covalently immobilized proteins risk dissociating from sensor chips, leading to under-evaluated K<sub>d</sub> values. Covalently immobilized proteins using amine groups do not dissociate from sensor chips, but they are attached to sensor chips in many orientations; for some of these orientations, the ligand-binding sites are not approachable, leading to under-evaluated K<sub>d</sub> values. Compared with large proteins, the ligand-binding sites on covalently immobilized small proteins are more likely to be buried, leading to under-evaluated K<sub>d</sub> values. Compared with RBD–ACE2 binding, the spike protein–ACE2 binding is more complex: the RBD in the spike switches between standing up (to expose the RBM for ACE2 binding) and lying down (to hide the RBM) conformations<sup>16,19</sup>, complicating the interpretation of measured K<sub>d</sub> values. Therefore, K<sub>d</sub> values measured by different SPR studies vary, depending on which protein is coated as well as the size and shape of proteins. In a previous study<sup>12</sup>, the K<sub>d</sub> value was higher when the RBD was coated than when the ACE2 was coated. In the present study, we could not coat ACE2 because ACE2 dissociated from sensor chips in regeneration buffer (for unknown reasons). We therefore coated the RBD, and the measured K<sub>d</sub> value was comparable to that from the previous study<sup>12</sup>.

# Article

**Extended Data Table 3 | Critical ACE2-binding residues in SARS-CoV-2 and SARS-CoV RBMs**

Viral RBD	Year	442	468-471	472	479	480	487
SARS-human	2002	Y	P-PA	L	N	D	T
SARS-civet	2002	Y	P-PA	L	K	D	S
CoV-pangolin/GD	2020	L (455)	GVEG (482-485)	F (486)	Q (493)	S (494)	N (501)
CoV-pangolin/GX	2020	L (455)	GQVG (482-485)	L (486)	E (493)	R (494)	T (501)
Rs3367-bat	2013	S (443)	P-PA (469-472)	F (473)	N (480)	D (481)	N (488)
RaTG13-bat	2020	L (455)	GQTG (482-485)	L (486)	Y (493)	R (494)	D (501)
<b>SARS-CoV-2</b>	<b>2019</b>	<b>L (455)</b>	<b>GVEG (482-485)</b>	<b>F (486)</b>	<b>Q (493)</b>	<b>S (494)</b>	<b>N (501)</b>

Residues in the SARS-CoV-2 RBM are labelled in red.

Corresponding author(s): Fang Li

Last updated by author(s): Mar 12, 2020

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give P values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection

Beam line at Advanced Photon Source beamline 24-ID-E is controlled by in house developed "Console 6.2.0" suite of programs.  
Automated data processing is enabled by locally developed software suite called RAPD.

Data analysis

HKL2000, CCP4 7.0, PHENIX-1.17.1, PyMol 2.0, LigPlot+ program and PDBePISA web server Ver.1.48

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Coordinates and structure factors have been deposited to the Protein Data Bank with accession number 6VW1.

### Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

 Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](http://nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample-size calculation was performed. For the protein expressions in insect cells, 2 liters cell culture (about 2-3x10 <sup>6</sup> cells/ml) were used each time.
Data exclusions	No data were excluded from the analyses
Replication	We have successfully repeated the crystallization condition more than 20 times. Pull-down assay and pseudovirus assay were each repeated 3 times.
Randomization	Randomization was not relevant to our study. Because there's no allocation of samples/organisms/participants involved in our study.
Blinding	Investigators were not blinded to group allocation during data collection and/or analysis. Because there's no group allocation involved in this study.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

- |                                     |   |
|-------------------------------------|---|
| n/a                                 | Involved in the study                                     |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Antibodies            |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Eukaryotic cell lines |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology                    |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Animals and other organisms      |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Human research participants      |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data                    |

### Methods

- |                                     |   |
|-------------------------------------|---|
| n/a                                 | Involved in the study                           |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq               |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry         |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |

## Antibodies

### Antibodies used

Primary antibody for C9 tag detection: rhodopsin (1D4). Its supplier: Santa Cruz Biotechnology. Its catalog number: sc-57432. Its clone name: 1D4. Its lot #: E0819.  
 Primary antibody for HIV-1 p24 detection: HIV-1 p24 (24-4). Its supplier: Santa Cruz Biotechnology. Its catalog number: sc-69728. Its clone name: 24-4. Its lot #: F1417.  
 Peroxidase-conjugated secondary antibody was also used for Western blotting (WB). Its supplier: Jackson ImmunoResearch. Its catalog number: 115-035-062. Its lot #: 139773

### Validation

Anti-rhodopsin Antibody (1D4) is a mouse monoclonal IgG1, which is recommended for detection of rhodopsin of mouse, rat and human origins by WB, IP, IF, IHC(P) and ELISA; also reactive with additional species, including bovine. The dilution ratio is 1:1,000 for WB.  
 Anti-HIV-1 p24 Antibody (24-4) is a mouse monoclonal IgG2b which is recommended for detection of Gag p24 of HIV-1 origin by WB, IP, IF and FCM. The dilution ratio is 1:1,000 for WB.  
 Peroxidase-conjugated secondary antibody is a goat anti-mouse IgG (H+L) which is recommended for WB with a dilution ratio of 1:10,000 - 1:20,000.

## Eukaryotic cell lines

Policy information about [cell lines](#)

### Cell line source(s)

sf9 insect cells were purchased from ATCC (ATCC® CRL-1711™).  
 HEK293T cells were purchased from ATCC (ATCC® CRL-3216™).  
 ESF 921 Insect Cell Culture Medium were purchased from ThermoFisher Scientific (catalog #: 96-001-01).  
 DMEM (Dulbecco's Modified Eagle Medium) were purchased from Gibco (catalog #: 11965092).

### Authentication

Cell lines used were not authenticated

### Mycoplasma contamination

Cell lines used were not tested for mycoplasma contamination

### Commonly misidentified lines (See [ICLAC](#) register)

No commonly misidentified cell lines were used