

## Loss of pH switch unique to SARS-CoV2 supports unfamiliar virus pathology

Kristina A. Paris, Ulises Santiago, and Carlos J. Camacho\*

*Department of Computational and Systems Biology  
University of Pittsburgh, Pittsburgh, PA, 15260, USA*

\*Correspondence: [ccamacho@pitt.edu](mailto:ccamacho@pitt.edu)

### Abstract

Cell surface receptor engagement is a critical aspect of viral infection. This paper compares the dynamics of virus-receptor interactions for SARS-CoV (CoV1) and CoV2. At low (endosomal) pH, the binding free energy landscape of CoV1 and CoV2 interactions with the angiotensin-converting enzyme 2 (ACE2) receptor is almost the same. However, at neutral pH the landscape is different due to the loss of a pH-switch (His445Lys) in the receptor binding domain (RBD) of CoV2 relative to CoV1. Namely, CoV1 stabilizes a transition state above the bound state. In situations where small external strains are applied by, say, shear flow in the respiratory system, the off rate of the viral particle is enhanced. As a result, CoV1 virions are expected to detach from cell surfaces in time scales that are much faster than the time needed for other receptors to reach out and stabilize virus attachment. On the other hand, the loss of this pH-switch, which sequence alignments show is unique to CoV2, eliminates the transition state and allows the virus to stay bound to the ACE2 receptor for time scales compatible with the recruitment of additional ACE2 receptors diffusing in the cell membrane. This has important implications for viral infection and its pathology. CoV1 does not trigger high infectivity in the nasal area because it either rapidly drifts down the respiratory tract or is exhaled. By contrast, this novel mutation in CoV2 should not only retain the infection in the nasal cavity until ACE2-rich cells are sufficiently depleted, but also require fewer particles for infection. This mechanism explains observed longer incubation times, extended period of viral shedding, and higher rate of transmission. These considerations governing viral entry suggest that number of ACE2-rich cells in human nasal mucosa, which should be significantly smaller for children (and females relative to males), should also correlate with onset of viral load that could be a determinant of higher virus susceptibility. Critical implications for the development of new vaccines to combat current and future pandemics that, like SARS-CoV2, export evolutionarily successful strains via higher transmission rates by viral retention in nasal epithelium are also discussed.

**Keywords:** SARS-CoV; SARS-CoV2; binding free energy; pH; shear flow; infectivity; incubation; shedding; transmission rate; viral entry.

## Introduction

Although accurate assessments are still evolving, reports from the world health organization indicate that infection with SARS-CoV-2 (CoV2) is significantly different relative to infection with previous respiratory viruses. The biggest distinctions are the longer incubation period and increased viral shedding (both likely responsible for higher rates of transmission), strong correlation of infected-fatality rates (IFR) with age and comorbidities, higher IFR for males relative to females, and minimal impact in children. As much as 40% of deaths from CoV2 relate to cardiovascular complications (Akhmerov and Marban, 2020), while cellular and animal models have also revealed inappropriate inflammatory responses along with high chemokine production (Blanco-Melo et al., 2020). Using thermodynamic, kinetic and molecular modeling, we explored potential reasons for these CoV2-unique aspects and aimed to identify determinants of its complex pathology. While it stands to reason that some of the answers may be found in the complex genomic changes triggered by the virus in cells, tissues, and organs, longer incubation and infectivity time scales also suggest that differences could have a biophysical origin.

Work on SARS-CoV (CoV1) has already determined that the virus enters cells via receptor-mediated endocytosis in a pH-dependent manner (Wang et al., 2008) that is characterized by co-translocation of the viral spike glycoprotein and its specific functional receptor, the angiotensin-converting enzyme 2 (ACE2), from the cell surface to early endosomes. Key steps that control the fate of the virus in the early and late endosome are driven in part by lowering the pH from 6.5-to-6.0 and from 6.0-to-5.0 (Bui et al., 1996), respectively; exposure to low pH triggers a spike catalyzed fusion between the viral and endosomal membranes followed by viral genome release. The cell machinery is then hijacked to replicate and assemble new virus particles that eventually exit the cell, primarily through budding. Broadly speaking, this is the same pH-dependent endocytic path followed by the influenza virus (Qin et al., 2019; Yamauchi, 2019), and likely all other coronaviruses (including MERS-CoV).

While infections by CoV1 and CoV2 are mediated by the ACE2 receptor, MERS-CoV (MERS) gains entry to cells through DPP4 (Raj et al., 2013). Figure 1 shows structures for the receptor-binding domains (RBD) in complex with their receptors for all three viruses. Surprisingly and likely significantly, while the RBD of CoV1 (Li et al., 2005) and MERS (Wang et al., 2013) have one Histidine on opposite ends of their binding interface, CoV2 (Wang et al., 2020) does not have His residues in this domain. As we have been able to determine to date, all known strains of CoV2 have mutated away their last His residue that is still present in the RBD of CoV1/MERS and related zoonotic viruses (see below). Because cell surface receptor engagement is a critical aspect of viral infection and life cycle, and sensing pH is relevant for both viral replication and regulation of Histidine protonation, we set to decipher the mechanistic role of the remaining His residue that distinguishes the RBDs of CoV1 and CoV2.

## Results

**pH-switch.** At low (endosomal) pH, the binding free energy landscape of CoV1 and CoV2 interactions with their ACE2 receptor is almost the same. This is important because low pH is critical for the activation of the spike fusion in late endosomes (Martin and Helenius, 1991). In particular, His34 in ACE2 (Fig. 1), located at the core of the RBD binding interface, should play a key role in this process. Indeed, we predict based on the co-crystal structures that His34<sup>+</sup> should readily form a hydrogen bond network that stabilizes the RBD/ACE2 complex in both CoV1 and

CoV2 (Fig. 2). Thus, loss of the pH-switch in CoV2 has no impact on the low pH bound conformation with ACE2.

On the other hand, at neutral pH the landscape is different due to the loss of the pH-switch (His445Lys) in the RBD of the spike protein of CoV2 relative to CoV1. We studied this loss by performing three independent unconstrained molecular dynamics simulations (MDs) of CoV1 PDB 2DD8 (Prabakaran et al., 2006) and CoV2 from the receptor (apo) in PDB 6LZG at both physiological and low pH  $\sim 5.0$  conditions (See Methods in Supplementary Information). In practice, lowering the pH protonates the His residue from neutral to positively charged (His<sup>+</sup>) by the addition of an extra hydrogen. Of note, although not studied here, the pH-switch in the RBD of MERS has previously been observed in pH-dependent crystal structures (Zhang et al., 2018).

MDs revealed two distinct regions in the binding interface of CoV1: (a) a loop motif (F<sub>460</sub>SPDGKPCPTPALNCY<sub>475</sub>) that for His and His<sup>+</sup> shows mostly not-bound- and bound-like conformations, respectively; and, (b) the rest of the binding interface that consistently adopts bound-like conformations that are independent of pH (Fig. 3A-C). Figure 3A-B shows the dominant clusters observed in the His and His<sup>+</sup> simulation of CoV1, as well as the pH-independent simulation of CoV2. Figure 3C-D shows detailed analyses of the corresponding root-mean-square-deviations (RMSDs) of these two regions relative to the co-crystals as a function of time (Additional MDs are shown in Fig. S1). The plots include the equilibration time (between 100-200 ns) in order to emphasize that the distinct trajectories were not biased by different initial conformations. Remarkably, deletion of the pH switch in CoV2 that mutates the last Histidine, His445, in CoV1 by Lys458 in CoV2 generates a motif, which includes T<sub>470</sub>EIYQAG<sub>476</sub>, that yields almost exclusively bound-like conformations (Fig. 3D)—i.e., CoV2 is always ready to bind ACE2. It is interesting to note that the effect of both Lys458 in CoV2 and His<sup>+</sup> in Cov1 (Fig. 3C-D) is to stabilize the bound-like state.

Implications of these findings in the binding free energy landscape are sketched in Fig. 4. Namely, at low (late endosomal) pH, the landscape of CoV1 and CoV2 interactions with the ACE2 receptor are very similar. However, at neutral pH the landscape is different due to the loss of the pH switch in the RBD of CoV2 relative to CoV1. Specifically, the not-bound-like conformations of the pH-dependent loop in CoV1 stabilizes a higher free energy transition state, whereas the persistent bound-like behavior of CoV2 yields a much tighter bond.

**Dynamics of virus-receptor interactions for CoV1 and CoV2.** What are the implications of this transition state found in CoV1 but not in CoV2 at neutral pH? To answer this question, we need to consider that viral particles in the respiratory system are under small external strains from, e.g., shear flow in the respiratory airways, when engaging cell surface receptors.

The cell mechanics of these interactions can be described by applying the reaction-limited kinetics of membrane-to-surface adhesion and detachment first envisioned by Dembo et al (Dembo et al., 1988). Specifically, one can write the free energy of the bound-state (bs)  $\Delta G_{bs}$  under a tensile force as a function of the cell-cell gap width  $L$ , a constant binding free energy  $\Delta G_{bs}^0$  and a “spring” energy such that

	$\Delta G_{bs}(L - L_{bs}) = \Delta G_{bs}^0 + \frac{1}{2}\kappa_{bs}(L - L_{bs})^2,$	(1)
--	---------------------------------------------------------------------------------------	-----

where  $L_{bs}$  is the equilibrium length for the bonded state. A similar equation holds for the free energy of the transition state (ts)  $\Delta G_{ts}$  at the same cell-cell gap,

	$\Delta G_{ts}(L - L_{ts}) = \Delta G_{ts}^0 + \frac{1}{2}\kappa_{ts}(L - L_{ts})^2.$	(2)
--	---------------------------------------------------------------------------------------	-----

This treatment assumes equilibrium between bonded and de-bonded states, so there must be a very slow "ramp" rate for the force of pulling or pushing. If these conditions apply, the equilibrium constant for bond formation can be written as

$$K_{eq}(L - L_{bs}) = e^{-\frac{\Delta G_{bs}^0}{k_b T}} e^{-\frac{\kappa_{bs}(L - L_{bs})^2}{2k_b T}}, \quad (3)$$

where  $k_b T$  is the Boltzmann factor. Note that

$$K_{eq}(0) = e^{-\Delta G_{bs}^0/k_b T} = k_{eq}^{on} / k_{eq}^{off}. \quad (4)$$

According to Arrhenius theory, the de-bonding rate constant at a given gap-width can then be written as

$$K_{off}(L) = k_{off}^0 e^{-[\Delta G_{ts}(L) - \Delta G_{bs}(L)]/k_b T}, \quad (5)$$

The mechanical or structural basis of the RBD/receptor interactions in Fig. S2 can be characterized as a "door-knob" type junction across the gap, as opposed to a gripping or fish-hook bound state. The knob interactions with the receptor entailed two characteristic patches, a large bound-like domain and the smaller switching loop (Fig. 3), which one could model with spring constants  $\kappa_{bl}$  and  $\kappa_{loop}$ , respectively (see red and yellow surfaces in Fig. S2). Then, the elastic constant of the bound- and transition- state can be written as  $\kappa_{bs} \approx \kappa_{bl} + \kappa_{loop}$  and  $\kappa_{ts} \approx \kappa_{bl}$ , respectively, with the equilibrium rest lengths being essentially the same, i.e.,  $L_{bs} = L_{ts} = L_0$ . Thus, virus detachment to the transition state corresponds to an ideal case of the theory, where the only allowed change between the bonded and transition state is in the spring constants, such that

$$K_{off}(L) = k_{off}^0 e^{[\Delta G_{bs}^0 - \Delta G_{ts}^0]/k_b T} e^{(\kappa_{bs} - \kappa_{ts})(L - L_0)^2/2k_b T} \approx K_{off}^0 e^{\kappa_{loop}(L - L_0)^2/2k_b T}, \quad (6)$$

where

$$K_{off}^0 = k_{off}^0 e^{\Delta \Delta G_{loop}^0/k_b T} \text{ and } \Delta \Delta G_{loop}^0 = \Delta G_{bs}^0 - \Delta G_{ts}^0. \quad (7)$$

It is clear that the spring approximation should only apply for small deformations. In general, the springs across the gap undergo a "twisting" motion around its long-axis to reach the transition state from the bonded state. The motion will increase "tightness" of the spring if  $(\kappa_{bs} - \kappa_{ts}) < 0$ , which defines a catch-bond. Here, however, CoV1 always loosens tightness,  $(\kappa_{bs} - \kappa_{ts}) = \kappa_{loop} > 0$ , corresponding to a typical slip-bond whose lifetime is shortened by tensile forces acting in the bond.

**Free energy landscapes and estimates of bond detachment for CoV1/CoV2 and ACE2.** We use the *FastContact* server (Champ and Camacho, 2007) to compute the electrostatic  $\Delta G_{elec}^0$  and desolvation  $\Delta G_{desol}^0$  binding free energies of the bound and transition states using co-crystal structures and chimeras that incorporate changes triggered by low pH structures. Entropies coupled to the unbound state could be somewhat higher for CoV1 relative to CoV2 due to the larger conformational entropy associated with the switching loop in Fig. 3. Other error bars are correlated since interactions are very similar such that  $\Delta \Delta G$ 's have an error bar of  $\pm 1$  kcal/mol. Absolute free energies need to account for size-dependent configurational and vibrational entropy changes upon binding, which for high affinity protein-protein complexes have been estimated to be anywhere between 5-to-15 kcal/mol. However, for flat and rather superficial complexes such as those here (Fig. S2), the entropy loss could be much lower. Finally, the pre-exponential factor  $k_{off}^0$  (Eq. 5) in the absence of a transition state and at equilibrium is exactly  $k_{eq}^{on}$  at 1 M concentration, which for diffusion-limited association can be approximated by  $k_{eq}^{on} \sim 10^6$  s<sup>-1</sup>

(Camacho et al., 2000). Table 1 lists  $\Delta G^0 = \Delta G_{elec}^0 + \Delta G_{desol}^0$  and the baseline rate of detachment for the RBD/ACE2 bond under tension  $K_{off}^0$  to the (a) transition state (ts) of CoV1, and full detachment of the (b) transition (ts) of CoV1 and (c) CoV2

Virus	$\Delta G_{bs}^{exp}$ [kcal/mol]	$\Delta G_{bs}^0$ [kcal/mol]	$\Delta G_{ts}^0$ [kcal/mol]	$\Delta \Delta G_{loop}^0$ $\pm 1$ [kcal/mol]	pH	$K_{off}^0$ [s <sup>-1</sup> ]	RBD/Receptor bond
CoV1	-11.1	-11.3	-8.9	-2.6	7.5	$1.2 \times 10^4$	slip
CoV1	N/A	-8.9 (ts)	N/A	N/A	7.5	0.3	-
CoV2	-11.7	-11.7	N/A	N/A	7.5	0.003	tight

**Table 1. Free energy estimates and baseline rate of detachment of RBD/ACE2 complex.** Estimates of bound-state (bs) are computed by *FastContact* server and co-crystals shown in Figs 1 and 2 (neutral pH). Free energy of transition state in CoV1 was computed by eliminating the contribution of the switching loop 460-475 that acquires not-bound-like conformations from *FastContact* calculation (Fig. 3A-C). Experimental equilibrium binding free energy is from (Walls et al., 2020).

Free energy estimates of bound complexes are fully consistent with experimental data (Walls et al., 2020); alternative measurements have suggested a 10 fold weaker binding (Shang et al., 2020). These differences will only re-scale  $K_{off}^0$  by a factor of 10 but will not be significant to our conclusions. The key observation is that CoV1 stabilizes a transition state by about  $2.6 \text{ kcal/mol}$  above the bound state. As a result, small external strains applied by, say, shear flow in the respiratory system enhance the off rate of the viral particle as shown in Eq. 6. Thus, CoV1 virions are expected to detach from cell surfaces in faster time scales. These binding free energy estimates are depicted in the landscapes in Fig. 4. Only at physiological pH should the landscape of CoV1 display a pH-dependent transition state. Other bonds are expected to break in an all-or-none type of transition.

**Optimal dwelling times and endocytosis.** In principle, the pH-switch in CoV1 could provide a natural mechanism to optimize virus internalization. Namely, CoV1 is expected to “bounce around” cell surfaces many times before cell entry. If the density of receptors is high enough, a “stick-and-slip” approach could be an efficient mechanism to find clusters of receptors randomly distributed on the cell surface. On the other hand, if only a small number of cell surface receptors are available, then receptor diffusion will be the limiting step to accrue the critical number of receptors needed for endocytosis, and longer RBD/ACE2 dwelling times will be required. Of note, tighter binding to ACE2 would also make it easier for a smaller number of CoV2 particles to establish an initial foothold in the respiratory system compared to weak binding where particles could be exhaled out.

Broad estimates of “high” concentration, e.g., in the range of 1000-to-10,000 receptors, yield an average separation between receptors  $D \sim 0.6 - 0.2 \mu\text{m}$  (see Fig. 4B) that is larger than the diameter of the virus  $\sim 0.12 \mu\text{m}$ . Thus, after attachment of the first spike to its receptor, recruitment of a second receptor to stabilized virus attachment will be limited by other receptors circulating in the cell membrane. Lateral protein diffusion in cell membranes is length-scale dependent, varying between  $\sim 0.01 \frac{\mu\text{m}^2}{\text{s}}$  (Spendier et al., 2012) and  $\sim 0.001 \frac{\mu\text{m}^2}{\text{s}}$  (Kusumi et al., 2005) for 40-100 nm and  $>100 \text{ nm}$ , respectively. Thus, diffusion time scales to bring two receptors into close proximity for the above length scales are  $\sim 200 - 5 \text{ s}$ .

It is noteworthy that the number of surface receptors in cells have an upper limit of about 20,000, which in the respiratory airways could limit infectivity to dwelling times of about  $K_{off}^0 \sim 1 \text{ s}^{-1}$  (or  $K_{eq} \sim 10^6 \text{ M}^{-1}$ ) based on  $D \sim 0.15 \mu\text{m}$  (Fig. 4).

## Discussion

**Viral infection and its pathology.** The lifetime of CoV1 RBD/ACE2 bonds at physiological pH ( $\sim 3 \text{ s}$ ) is marginally short-lived for efficiently triggering endocytosis, even at high ACE2 receptor concentrations. As a result, CoV1 virions are expected to detach from cell surfaces in time scales that are much faster than the time needed for other receptors to reach out and stabilize virus attachment. And for human nasal goblet cells, it will be significantly worse since, after each bounce, particles will be biased by gravity to either diffuse down the respiratory tract or be exhaled, where they will not find significant amounts of ACE2 receptors until reaching lung alveolar epithelial cells (Hamming et al., 2004). On the other hand, deletion of the pH switch allows CoV2 to have RBD/ACE2 bonds with dwelling times of about  $\sim 300 \text{ s}$ , commensurate with the diffusion time scales needed to recruit enough ACE2 receptors to trigger endocytosis. This mechanism implies that, for the most part, CoV1 will not co-localize in the nasal cavity. This prediction is consistent with CoV1 being mainly a lower respiratory tract disease, causing complications that include acute respiratory distress (Ding et al., 2003; Hamming et al., 2004).

Viral replication in human mucous gland cells will release viruses back into the same area where they can infect new cells until the supply of ACE2 receptors is depleted below the critical threshold needed for binding and internalization. This process will trap viral particles in the upper respiratory tract, naturally leading to longer incubation times. Similarly, accumulation of viral particles in the nasal mucosa will lead to extended periods of viral shedding. Of note, since viral transit to the lower respiratory tract will be significantly slower for CoV2 relative to CoV1, this period of higher infectivity rates could be for the most part mediated by asymptomatic individuals.

Based on our findings, incubation times should correlate with the number of ACE2-rich cells in the nasal area. It is important to note that children do not have well developed sinuses until adolescence (Henson et al., 2020). Thus, large areas for viral replication will not be available in children, resulting in shorter incubation times due to the faster diffusion down to the lower respiratory tract. Something similar could apply to females who have smaller nasal cavities relative to males (Samolinski et al., 2007). Shorter times in the nasal cavity would lead to a lower viral load in the upper airways and could explain the lower transmission and milder symptoms that are observed in children, as well as the lower IFR in adult women relative to men.

Our proposed mechanism is also consistent with reported loss of sense of smell (anosmia) that may occur by day 3 of a CoV2 infection (Speth et al., 2020), as cells in the nasal cavities support olfactory mucous membranes needed for the perception of smell. Proximity to the brain also suggest that CoV2 infections could impact the brain in ways that other SARS viruses cannot. Moreover, cardiovascular and immunological complications triggered by CoV2 could also be explained on the basis of long-term insult of endothelial cells by viral sequestration of the ACE2 receptor (Gurley and Coffman, 2008).

**pH-switch across species.** Further supporting the observation that CoV2 is unique among other coronaviruses is shown in Table 2 that compares sequence alignments of pH domains in RBDs of both CoV1, CoV2, MERS, as well as other closely related zoonotic viruses. CoV2, and related coronaviruses in one species of pangolin and some bats do not share the pH-switch present in CoV1, instead they share the Lys458 stabilization motif. However, these zoonotic viruses still have

pH-switches that co-localize next to the pH-switch in the RBD of MERS structure (Fig. 1). Interestingly, different bat-infecting strains show putative pH-switches that are closer in both sequence and structure to either MERS or pangolin-associated coronavirus. While we have not yet found the species or strain where the loss of the pH-switch first occurred, these relationships point at the possible zoonotic origin of CoV2 as well as evolutionary pressures to preserve the pH-switch. It is noteworthy that DPP4, the receptor of the MERS RBD, is not found in nasal epithelial cells (Meyerholz et al., 2016).

	CoV1 pH-switch	
Civet	AWNTRNIDATSTG-NYNYKYRYLR	RGKLRPFERDISNVPFSSDGKPCTP-PAPNCYWPLRGYGFYTTSGIGYQP
CoV1	AWNTRNIDATSTG-NYNYKYRYLR	RGKLRPFERDISNVPFSPDGKPCTP-PALNCYWPLNDYGFYTTTGIGYQP
CoV2	435 AWNSNNLDSKVG	G-NYNYLYRFRKSNLKPFFERDIS <span style="border: 1px solid black; padding: 2px;">TEIYQAG</span> STPCNGVEGFNCFYPLQSYGFQPTNGVGYQP 507
Pangolin	AWNSVKQDALTTGG-NYGYLYRFRK	SKLKPFFERDIS <span style="border: 1px solid black; padding: 2px;">TEIYQAG</span> STPCNGQVGLNCYYPLERYGFHPTTGVNYP
Bat1	AWNSKHIDAKEGG-NFNYLYRFRK	ANLKPFFERDIS <span style="border: 1px solid black; padding: 2px;">TEIYQAG</span> SKPCNGQTGLNCYYPLRYGFYPTDGVGHQP
Bat2	AWNTNSKDSSTSG-NYNYLYRWVR	SKLNPERDLNDIYSPGGQSCSA-VGPNCYNPLRPYGFFTTAGVGHQP
MERS	ATVPHNLTITKPLKYSYINKCSR	-----LLSDDRTEVLQLVN PLEGGGWLVASGSTVAM
	MERS pH-switch	pH-switches co-localized with MERS

**Table 2. Sequence alignments of pH domains in RBD CoV2 and related viruses.** The His→Lys mutation is highlighted in red. Conserved motif relevant for interaction with K458 is enclosed in a rectangle for CoV2, Pangolin and Bat1. NCBI Locus AAU04646 (Civet); 6CRV\_C (CoV1); 6VSB\_A (CoV2); QIQ54048 (Pangolin); QHR63300 (Bat1); AGZ48806 (Bat2); QBM11748 (MERS).

**Outlook.** This newly discovered difference in protein sequence in the receptor binding domain of the spike glycoprotein and its impact on receptor binding reveals a mechanism that allows SARS-CoV2 internalization to take advantage of the high expression of ACE2 in the nasal epithelium—resulting in increased retention times in the upper respiratory tract and augmented infectivity. This mechanism reconciles observed epidemiological traits and pathologies specific to SARS-CoV2 and explains differences with those associated with SARS-CoV, which due to its stick-and-slip pH-switch is unable to efficiently undergo endocytosis in the nasal cavity.

SARS-CoV also has a higher infected-fatality rate than SARS-CoV2. While the evolutionary advantage of higher infectivity by SARS-CoV2 in the nasal area is clear, this property comes at the expense of an important regulatory mechanism that would have allowed this virus to more readily move in other organs and tissues. In fact, the life-cycle of SARS-CoV2 is significantly slower than that of SARS-CoV because CoV2 is essentially immobilized at its initial cell receptor contact. Thus, it seems unlikely that the diffusion limited recruitment of ACE2 receptors affecting the virus in the respiratory airways would also be the limiting step in tissues.

After internalization, the virus is encapsulated in a vesicle supported by RBD/ACE2 complexes. The actual final number of complexes in each vesicle should vary above a given threshold, though not much is known about the details of this process. Contrary to SARS-CoV, CoV2 complexes would be expected to have greater difficulty slipping and breaking. It is not difficult to imagine that for vesicles compressed by an excess of receptors the fusion with the early endosome might be hindered, hosting a population of viruses that could stay latent or activate at much later times. This simple mechanism could underlie the still anecdotal evidence for infection recurrence (Chen et al., 2020), as well as extremely long-term of viral shedding.

Collectively, our studies provide insight pertinent to the molecular basis of viral infectivity and, at the same time, validate this form of thermodynamic and molecular modeling as an approach to probe the evolution of the next SARS-mediated pandemic. From a therapeutic perspective, our findings linking viral pathology with long-term viral infection/retention in nasal epithelium of the upper respiratory tract suggest that vaccine development should not just concentrate on fighting systemic infection through induction of IgG responses, but should instead aim to elicit high titers

of secretory IgA antibodies capable of neutralizing the virus in the nasal mucosa. Therefore, intranasal delivery of a vaccine with strong IgA producing potential is a logical approach to consider as the next step in countering the current and future pandemics that, like SARS-CoV2, export evolutionarily successful strains via higher transmission rates.

**Acknowledgements.** This work was supported by NIH GM097082, NS043277 and T32EB009403. CJC is grateful to Dr. Jose Padial for his inspiration in starting this project, to Dr. Dana Ascherman for helping in the immunobiology component of the model, to Drs. Marc Herant and Harinder Singh for successfully challenging earlier interpretations of the model, to Dr. AR Carvunis for highlighting evolution, and to Dr. Olja Finn for her enthusiasm for our model and helpful suggestions.

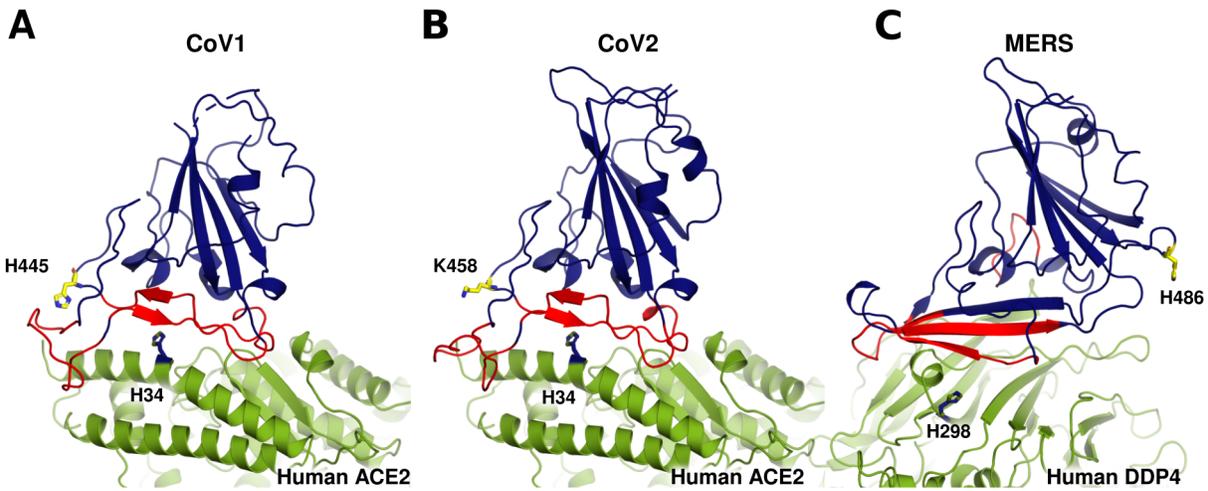
**Author Contributions.** CJC developed the mechanistic ideas and structural modeling, and wrote the paper. KAP and US performed and analyzed the molecular dynamics simulations.

**Declaration of Interests.** The authors declare no competing interests.

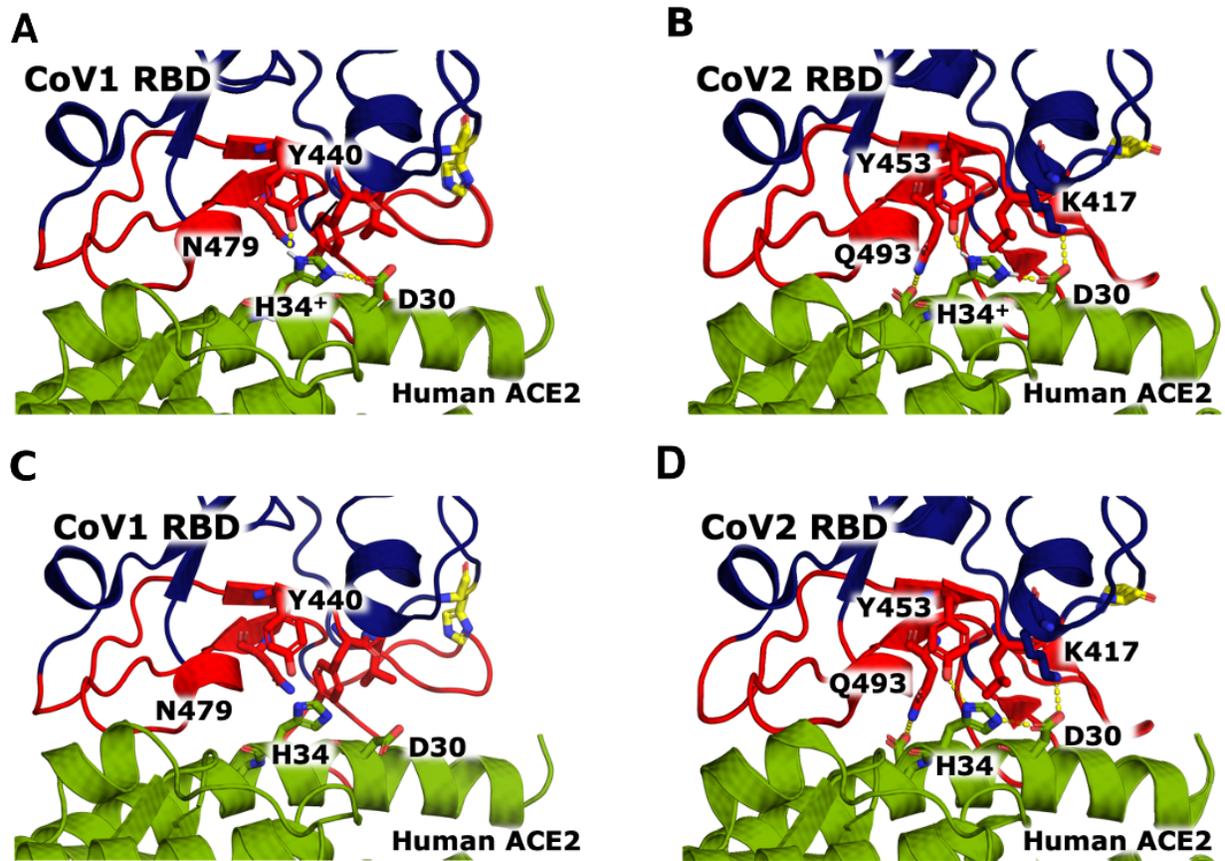
## REFERENCES

- Akhmerov, A., and Marban, E. (2020). COVID-19 and the Heart. *Circ Res* *126*, 1443-1455.
- Blanco-Melo, D., Nilsson-Payant, B.E., Liu, W.C., Uhl, S., Hoagland, D., Moller, R., Jordan, T.X., Oishi, K., Panis, M., Sachs, D., *et al.* (2020). Imbalanced Host Response to SARS-CoV-2 Drives Development of COVID-19. *Cell* *181*, 1036-1045 e1039.
- Bui, M., Whittaker, G., and Helenius, A. (1996). Effect of M1 protein and low pH on nuclear transport of influenza virus ribonucleoproteins. *J Virol* *70*, 8391-8401.
- Camacho, C.J., Kimura, S.R., DeLisi, C., and Vajda, S. (2000). Kinetics of desolvation-mediated protein-protein binding. *Biophys J* *78*, 1094-1105.
- Champ, P.C., and Camacho, C.J. (2007). FastContact: a free energy scoring tool for protein-protein complex structures. *Nucleic Acids Res* *35*, W556-560.
- Chen, D., Xu, W., Lei, Z., Huang, Z., Liu, J., Gao, Z., and Peng, L. (2020). Recurrence of positive SARS-CoV-2 RNA in COVID-19: A case report. *Int J Infect Dis* *93*, 297-299.
- Dembo, M., Torney, D.C., Saxman, K., and Hammer, D. (1988). The reaction-limited kinetics of membrane-to-surface adhesion and detachment. *Proc R Soc Lond B Biol Sci* *234*, 55-83.
- Ding, Y., Wang, H., Shen, H., Li, Z., Geng, J., Han, H., Cai, J., Li, X., Kang, W., Weng, D., *et al.* (2003). The clinical pathology of severe acute respiratory syndrome (SARS): a report from China. *J Pathol* *200*, 282-289.
- Gurley, S.B., and Coffman, T.M. (2008). Angiotensin-converting enzyme 2 gene targeting studies in mice: mixed messages. *Exp Physiol* *93*, 538-542.
- Hamming, I., Timens, W., Bulthuis, M.L., Lely, A.T., Navis, G., and van Goor, H. (2004). Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis. *J Pathol* *203*, 631-637.
- Henson, B., Drake, T.M., and Edens, M.A. (2020). Anatomy, Head and Neck, Nose Sinuses. In *StatPearls* (Treasure Island (FL)).
- Kusumi, A., Nakada, C., Ritchie, K., Murase, K., Suzuki, K., Murakoshi, H., Kasai, R.S., Kondo, J., and Fujiwara, T. (2005). Paradigm shift of the plasma membrane concept from the two-dimensional continuum fluid to the partitioned fluid: high-speed single-molecule tracking of membrane molecules. *Annu Rev Biophys Biomol Struct* *34*, 351-378.
- Li, F., Li, W., Farzan, M., and Harrison, S.C. (2005). Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. *Science* *309*, 1864-1868.
- Martin, K., and Helenius, A. (1991). Transport of incoming influenza virus nucleocapsids into the nucleus. *J Virol* *65*, 232-244.
- Meyerholz, D.K., Lambert, A.M., and McCray, P.B., Jr. (2016). Dipeptidyl Peptidase 4 Distribution in the Human Respiratory Tract: Implications for the Middle East Respiratory Syndrome. *Am J Pathol* *186*, 78-86.
- Prabakaran, P., Gan, J., Feng, Y., Zhu, Z., Choudhry, V., Xiao, X., Ji, X., and Dimitrov, D.S. (2006). Structure of severe acute respiratory syndrome coronavirus receptor-binding domain complexed with neutralizing antibody. *J Biol Chem* *281*, 15829-15836.
- Qin, C., Li, W., Li, Q., Yin, W., Zhang, X., Zhang, Z., Zhang, X.E., and Cui, Z. (2019). Real-time dissection of dynamic uncoating of individual influenza viruses. *Proc Natl Acad Sci U S A* *116*, 2577-2582.
- Raj, V.S., Mou, H., Smits, S.L., Dekkers, D.H., Muller, M.A., Dijkman, R., Muth, D., Demmers, J.A., Zaki, A., Fouchier, R.A., *et al.* (2013). Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. *Nature* *495*, 251-254.

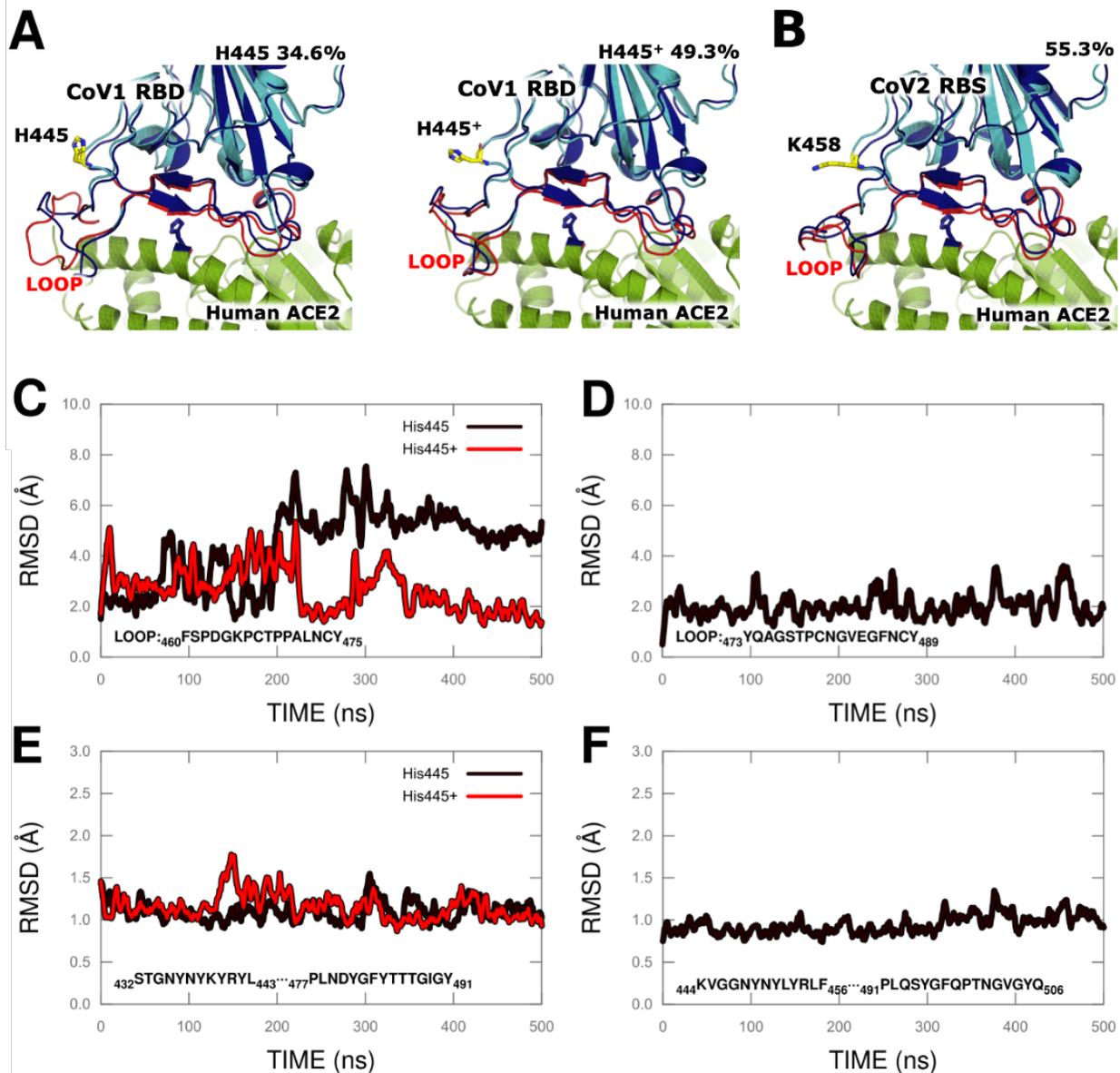
- Samolinski, B.K., Grzanka, A., and Gotlib, T. (2007). Changes in nasal cavity dimensions in children and adults by gender and age. *Laryngoscope* *117*, 1429-1433.
- Shang, J., Ye, G., Shi, K., Wan, Y., Luo, C., Aihara, H., Geng, Q., Auerbach, A., and Li, F. (2020). Structural basis of receptor recognition by SARS-CoV-2. *Nature* *581*, 221-224.
- Spendier, K., Lidke, K.A., Lidke, D.S., and Thomas, J.L. (2012). Single-particle tracking of immunoglobulin E receptors (FcεpsilonRI) in micron-sized clusters and receptor patches. *FEBS Lett* *586*, 416-421.
- Speth, M.M., Singer-Cornelius, T., Obere, M., Gengler, I., Brockmeier, S.J., and Sedaghat, A.R. (2020). Olfactory Dysfunction and Sinonasal Symptomatology in COVID-19: Prevalence, Severity, Timing, and Associated Characteristics. *Otolaryngol Head Neck Surg*, 194599820929185.
- Walls, A.C., Park, Y.J., Tortorici, M.A., Wall, A., McGuire, A.T., and Veesler, D. (2020). Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* *181*, 281-292 e286.
- Wang, H., Yang, P., Liu, K., Guo, F., Zhang, Y., Zhang, G., and Jiang, C. (2008). SARS coronavirus entry into host cells through a novel clathrin- and caveolae-independent endocytic pathway. *Cell Res* *18*, 290-301.
- Wang, N., Shi, X., Jiang, L., Zhang, S., Wang, D., Tong, P., Guo, D., Fu, L., Cui, Y., Liu, X., *et al.* (2013). Structure of MERS-CoV spike receptor-binding domain complexed with human receptor DPP4. *Cell Res* *23*, 986-993.
- Wang, Q., Zhang, Y., Wu, L., Niu, S., Song, C., Zhang, Z., Lu, G., Qiao, C., Hu, Y., Yuen, K.Y., *et al.* (2020). Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell* *181*, 894-904 e899.
- Yamauchi, Y. (2019). Quantum dots crack the influenza uncoating puzzle. *Proc Natl Acad Sci U S A* *116*, 2404-2406.
- Zhang, S., Zhou, P., Wang, P., Li, Y., Jiang, L., Jia, W., Wang, H., Fan, A., Wang, D., Shi, X., *et al.* (2018). Structural Definition of a Unique Neutralization Epitope on the Receptor-Binding Domain of MERS-CoV Spike Glycoprotein. *Cell Rep* *24*, 441-452.



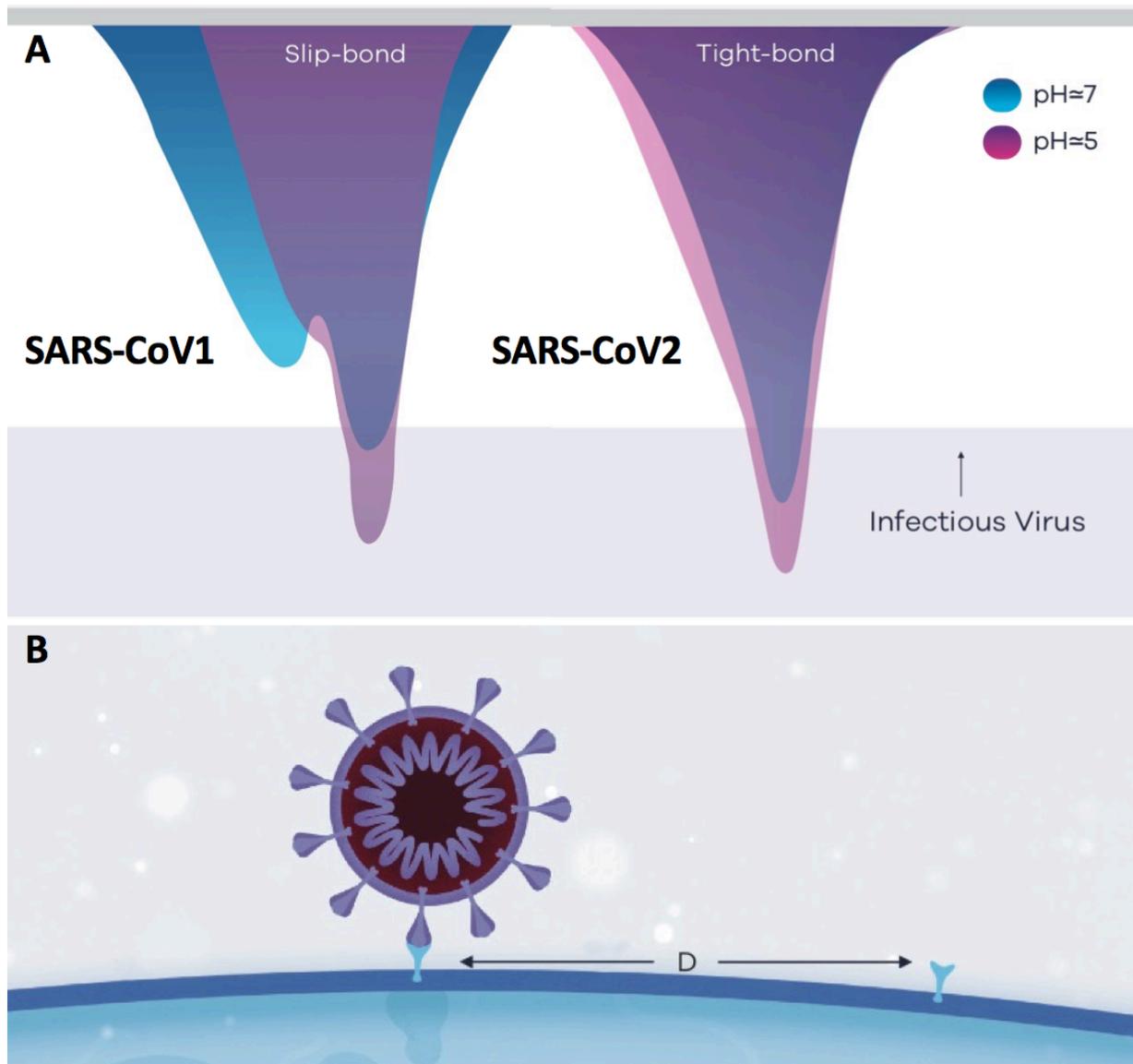
**Figure 1. Deletion of Histidine residue in receptor binding domain (RBD) of CoV2.** Co-crystal structures of the receptor binding domain (RBD) and the receptor for (A) SARS-Cov (PDB 2AJF; CoV1), (B) SARS-CoV2 (PDB 6LZG; CoV2) and (C) MERS-CoV (PDB 4L72; MERS). RBD of the viruses is in blue with binding interface in red. Receptor of Cov1 and Cov2 is ACE2, and of MERS is DPP4. RBD's of Cov1 and MERS have one histidine (H), H445 and H486, respectively, shown in yellow. ACE2 and DPP4 have several histidine residues, only the closest to the binding interface is shown in blue.



**Figure 2. Protonation of the Human ACE2 H34 further stabilizes binding interface.** Protonated state of H34<sup>+</sup> is predicted to form stable H-bond network with D30 in ACE2, and Y440 and Y453 in (A) CoV1 and (B) CoV2, respectively. H-bond network is based on rotamers already observed in co-crystals of the unprotonated forms: (C) PDB 2AJF for CoV1, and, (D) PDB 6LZG for CoV2 (see also PDB 6M0J). Unprotonated co-crystal structures of CoV2 assigned H34 with a N $\delta$ H making a bond with backbone oxygen of D30, which is already making a bond in the  $\alpha$ -helix. Our MDs indicate that even in the unprotonated form the rotamer should be rotated 180° having N $\delta$  and N $\epsilon$  more readily interacting with D30 and Y453 (as shown in panel D).



**Figure 3. Role of pH switch in CoV1 relative to CoV2.** (A) and (B) show the same co-crystals as in Fig. 1 superimposed with centroid of largest 1.8 Å RMSD cluster of conformations of pH-switching loop from MDs shown in (C) for CoV1 and (D) for CoV2. Also indicated is the size of the corresponding cluster relative to simulation time. (C) Root-mean-squared-deviation (RMSD) of 16 amino acid loop as a function of time that switches between not-bound-like (~ 5.8 Å) to bound-like (~ 1.9 Å) relative to co-crystal (PDB 2AJF), for His and His<sup>+</sup>, respectively; (D) Same analysis for CoV2 homologous loop shows most conformations under 2 Å RMSD relative to PDB 6LZG. Binding interface, not including loop, stays in a bound-like conformation for 100% of the simulation time for both (E) CoV1 and (F) CoV2.



**Figure 4. Free energy binding landscapes of RBD of SARS-CoV1 and SARS-CoV2 and ACE2 receptor.** (A) At physiological pH, CoV1 has a pH-dependent transition state that significantly weakens the RBD/ACE2 bond when small external strains are applied. Although CoV2 conformations do not depend on pH, protonation of ACE2 further stabilizes the complexes as depicted for pH  $\approx$  5. (B) Cartoon of virus attached to one cell receptor in respiratory airways, recruitment of a second receptor depends on balance between dwelling time of the bond and diffusion time scale by a distance  $D$  of protein in cell membrane, where  $D$  is average distance between receptors.

## Supplementary Information

### Methods

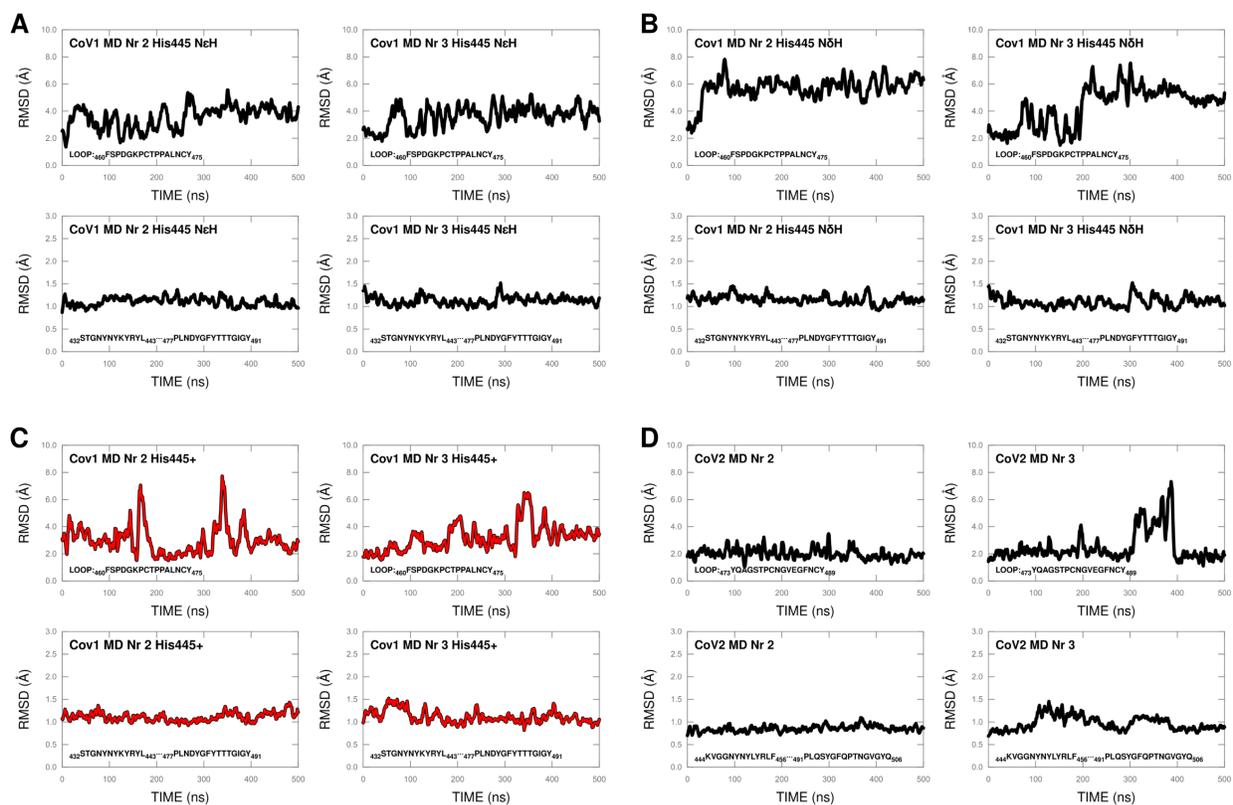
Atomic coordinates for starting structures were acquired from the Protein Data Bank [1]: 2DD8 was used for CoV1 RBD and 6LZG was used for the CoV2 RBD. The RBD from 2DD8 (bound to neutralizing antibody) was chosen instead of that in PDB ID 2AJF (bound to ACE2) as a starting structure as it includes otherwise missing portions of the domain. Modification of His to different tautomeric or protonation states was done with PyMol's Mutagenesis Wizard [2]. Molecular dynamics simulations (MDs) were carried out with pmemd.cuda from AMBER18 [3-5] using AMBER ff14SB force field [6] and Generalized Amber Force Field (GAFF) [7]. We used tLeap binary (part of AMBER18) for solvating the structures in a cubed TIP3P water box with a 10 Å distance from structure surface to the box edges, and closeness parameter of 0.75 Å. The system was neutralized and solvated. Simulations were carried out after minimizing the system, gradually heating the system from 0 K to 300K over 50 ps, and equilibrating the system for 1 ns at NPT. 500 ns of production was then carried out using NPT at 300 K with the Langevin thermostat, a non-bonded interaction cut off of 8 Å, time step of 2 fs, and the SHAKE algorithm to constrain all bonds involving hydrogens.

Clustering was completed using CPPTRAJ [8] and H-bond and RMSD calculations were done with VMD [9].

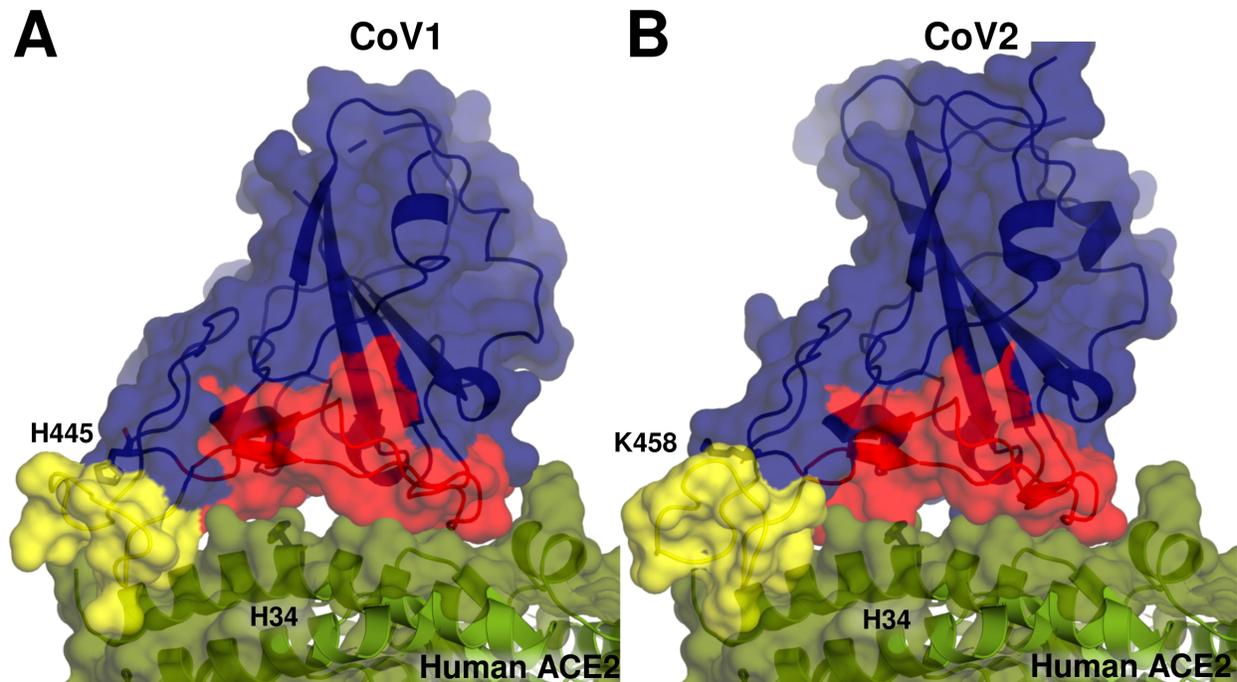
All figures were drawn using PyMOL [2] and GNUPLOT.

## References

1. H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.B. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne. The Protein Data Bank. *Nuc. Acids. Res.* 2000; 28 (1). 235-242.
2. The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC.
3. A.W. Goetz, M.J. Williamson, D. Xu, D. Poole, S. Le Grand, and R.C. Walker. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 1. Generalized Born. *J. Chem. Theory Comput.* 2012; 8 (5) 1542-1555. DOI: 10.1021/ct200909j
4. R. Salomon-Ferrer, A.W. Goetz, D. Poole; S. Le Grand, and R.C. Walker. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* 2013; 9 (9) 3878-3888. DOI: 10.1021/ct400314y
5. D.A. Case, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, D. Ghoreishi, M.K. Gilson, H. Gohlke, A.W. Goetz, D. Greene, R Harris, N. Homeyer, S. Izadi, A. Kovalenko, T. Kurtzman, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D.J. Mermelstein, K.M. Merz, Y. Miao, G. Monard, C. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C.L. Simmerling, J. Smith, R. Salomon-Ferrer, J. Swails, R.C. Walker, J. Wang, H. Wei, R.M. Wolf, X. Wu, L. Xiao, D.M. York and P.A. Kollman. AMBER 2018, 2018, University of California, San Francisco.
6. J.A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K.E. Hauser, C. Simmerling. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* 2015; 11 (8) 3696-3713. DOI: 10.1021/acs.jctc.5b00255
7. J. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollman, D.A. Case, Development and testing of a general AMBER force field. *J. Comput. Chem.*, 2004; 1157-1174. DOI: 10.1002/jcc.20035
8. D.R. Roe and T.E. Cheatham, III. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data *J. Chem. Theory Comput.* 2013; 9. 3084-3095.
9. W. Humphrey, A. Dalke, and K. Schulten. VMD - Visual Molecular Dynamics. *J. Molec. Graphics.* 1996; 14. 33-38.



**Figure S1. Dynamics of binding interface – additional MDs.** Switching Loop and remainder of the binding interface (not including Loop residues) RMSD relative to co-crystal structures as a function of time as described in Fig 3. **A.** Two independent MDs of CoV1 with unprotonated H445 and hydrogen on N $\epsilon$ . Top: RMSD of Loop shows unstable not-bound-like conformations. Bottom: RMSD of the remainder of the binding interface show consistent bound-like phenotype. **B.** Two extra MDs of CoV1 with unprotonated H445 and hydrogen on N $\delta$ . Top: RMSD of Loop shows not-bound-like phenotype. Bottom: RMSD of the remainder of the binding interface show stable bound-like conformations. **C.** Two additional MDs of CoV1 with protonated H445<sup>+</sup> shows significant bound-like behavior (RMSD  $\sim$  2-3 Å) for Loop (Top); and, RMSD  $\sim$  1 Å for binding interface without the loop (Bottom). **D.** Two additional MDs of CoV2: RMSD as a function of time shows consistent bound-like behavior for both the Loop region homologous to CoV1 (Top), and the remainder of the binding interface (Bottom).



**Figure S2. “Door-knob” binding interface of CoV1 (PDB 2AJF) and CoV2 (PDB 6LZG).** Surface representation of the co-crystals reveal two characteristic lobes with flat and mostly superficial contacts. Yellow surface corresponds to pH-switch loop and red surface indicates remaining of binding interface.