Biophysical*Journal*

**Article**

Biophysical Society

# The flexibility of ACE2 in the context of SARS-CoV-2 infection

Emilia P. Barros,[1] Lorenzo Casalino,[1] Zied Gaieb,[1] Abigail C. Dommer,[1] Yuzhang Wang,[2] Lucy Fallon,[2]
Lauren Raguette,[2] Kellon Belfon,[2] Carlos Simmerling,[2,3] and Rommie E. Amaro[1,*]
[1]Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California; [2]Department of Chemistry and [3]Laufer Center for Physical and Quantitative Biology, Stony Brook University, Stony Brook, New York

ABSTRACT The coronavirus disease 2019 (COVID-19) pandemic has swept over the world in the past months, causing significant loss of life and consequences to human health. Although numerous drug and vaccine development efforts are underway, there are many outstanding questions on the mechanism of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) viral association to angiotensin-converting enzyme 2 (ACE2), its main host receptor, and host cell entry. Structural and biophysical studies indicate some degree of flexibility in the viral extracellular spike glycoprotein and at the receptor-binding domain (RBD)-receptor interface, suggesting a role in infection. Here, we perform explicitly solvated, all-atom, molecular dynamics simulations of the glycosylated, full-length, membrane-bound ACE2 receptor in both an apo and spike RBD-bound state to probe the intrinsic dynamics of the ACE2 receptor in the context of the cell surface. A large degree of fluctuation in the full-length structure is observed, indicating hinge bending motions at the linker region connecting the head to the transmembrane helix while still not disrupting the ACE2 homodimer or ACE2-RBD interfaces. This flexibility translates into an ensemble of ACE2 homodimer conformations that could sterically accommodate binding of the spike trimer to more than one ACE2 homodimer and suggests a mechanical contribution of the host receptor toward the large spike conformational changes required for cell fusion. This work presents further structural and functional insights into the role of ACE2 in viral infection that can potentially be exploited for the rational design of effective SARS-CoV-2 therapeutics.

SIGNIFICANCE As the dominant host receptor of SARS-CoV-2, ACE2 has been the subject of extensive structural and antibody design efforts in aims to curtail COVID-19 spread. Here, we perform molecular dynamics simulations of the homodimer ACE2 full-length structure to study the dynamics of this protein in the context of the cellular membrane. The simulations evidence exceptional plasticity in the protein structure due to flexible hinge motions in the head-transmembrane domain linker region and helix mobility in the membrane, resulting in a varied ensemble of conformations distinct from the experimental structures. Our findings suggest a dynamical contribution of ACE2 to the spike glycoprotein shedding required for infection and contribute to the question of stoichiometry of the spike-ACE2 complex.

## INTRODUCTION

Angiotensin-converting enzyme 2 (ACE2) acts as the extracellular receptor for the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (1–3), the virus responsible for the COVID-19 pandemic that has catastrophically affected the world since its first identification in December 2019 (4–7). ACE2 is a membrane protein found in the lungs, kidneys, heart, and intestine cells (8,9) that plays a physiological role in cardiovascular regulation via the cleaving of intermediates in the maturation process of angiotensin, a peptide hormone involved in vasoconstriction control (10–14). ACE2 is a homodimer with a large claw-like extracellular head domain, a small transmembrane domain, and a short intracellular segment (8). The head can be further subdivided into the catalytic zinc-binding peptidase domain (PD; residues 19–615) (15) and the smaller neck domain (residues 616–726), which is where the majority of the homodimer interactions seems to lie (16). The neck domain is further connected to the single-helix transmembrane (TM) domain by a long linker (Fig. 1 *A*). ACE2 can also function as a membrane-trafficking chaperone for $B^0AT1$, an amino acid transporter (17), and it was in fact only in complex

---

*Submitted September 16, 2020, and accepted for publication October 27, 2020.*

*Correspondence: ramaro@ucsd.edu

Editor: Susan Schroeder.

https://doi.org/10.1016/j.bpj.2020.10.036

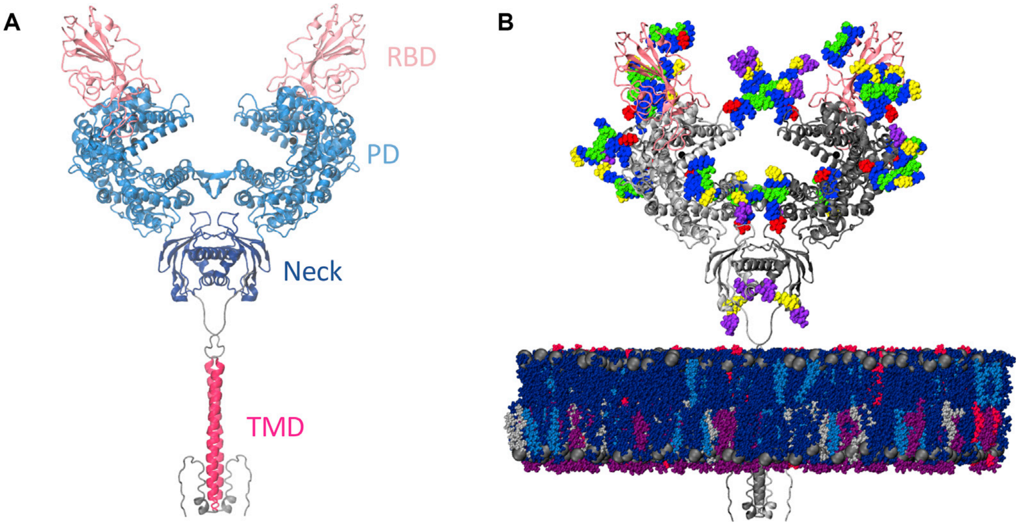


FIGURE 1 Model structure. (*A*) Full-length ACE2 homodimer protein structure in complex with spike protein RBDs. ACE2 peptidase, neck, and transmembrane domains are shown with cartoons highlighted in blue, navy, and magenta, respectively. Spike RBDs are depicted with pink cartoons. (*B*) Fully glycosylated and membrane-embedded model. ACE2 and RBDs are represented with gray and pink cartoons, respectively. Atoms of N- and O-glycans are shown with per-monosaccharide-colored spheres in which GlcNAc is highlighted in blue, mannose in green, fucose in red, galactose in yellow, and sialic acid in purple. Lipid heads (P atoms) are represented with gray spheres, whereas lipid tails are depicted with a licorice representation using the following color scheme: POPC (*navy*), POPI (*violet*), POPE (*silver*), CHL (*blue*), and PSM (*magenta*). To see this figure in color, go online.

with this partner that the single TM helix of ACE2 could be resolved (16).

SARS-CoV (18) (responsible for 8096 cases worldwide in 2002 (19)), and now the closely related SARS-CoV-2, hijack ACE2 as the host cell receptor to its large extracellular spike (S) glycoprotein (1,20). The spike's receptor-binding domain (RBD) in the "up" conformation binds to ACE2's PD with high affinity (21), and the resolved ACE2-RBD complex consists of a dimer of heterodimers, with each monomer of the ACE2 homodimer interacting with one RBD (thus forming a heterodimer; Fig. 1 *A*). Cellular recognition and binding to ACE2's PD via the RBD is proposed to initiate a series of complex conformational transitions in the S homotrimeric protein, leading to the shedding of its S1 subunit and fusion to the host cell membrane driven by the S2 subunit (22–26), ultimately resulting in the infection of the host cell. Downregulation of ACE2 and accumulation of angiotensin II because of spike binding is also associated with acute respiratory distress syndrome (ARDS) and acute lung failure (27–31), contributing to SARS-associated symptoms. As such, the S glycoprotein and ACE2-S complex are considered key targets for drug and antibody development efforts aiming to curtail the virus' remarkable transmissibility and negative effect on human health (1,32–35), including exploiting the ACE2-S high affinity with recombinant soluble ACE2-antibody constructs (35–38).

Experimental and biophysical studies of the SARS-CoV-2 RBD and soluble ACE2's PD complex have suggested structural factors likely responsible for the higher affinity and infectivity of SARS-CoV-2 compared with SARS-CoV (39–41) and revealed significant dynamics at the RBD-PD interface in the form of rocking motions between the two subunits (41,42). At the individual PD level, opening and closing of the active site cleft has been observed in X-ray structures of the extracellular region (43). Additionally, the recent cryogenic electron microscopy (cryoEM) structures of full-length ACE2 indicate that the homodimer can adopt a less populated open conformation, as defined by the distance between the head domains, in addition to the closed conformation shown in Fig. 1 *A* (16), but little else is known about the global flexibility of this membrane protein. The cryoEM and computational studies of the full-length spike glycoprotein have recently suggested a significant degree of flexibility of the spike's stalk and at the ACE2-RBD interface (44,45), evidencing the need to study these macromolecular complexes in the context of the cell surface. Here, we perform all-atom molecular dynamics (MD) simulations of the full-length, membrane-embedded, and glycosylated ACE2 homodimer both in the apo state and in complex with RBD to study the dynamics and molecular origins of the ACE2-S flexibility on the host receptor side that evade experimental characterization (46). Seven complex N-glycans and one O-glycan in ACE2 were modeled according to glycoanalytic data (47–49) as well as glycan N343 in RBD for the RBD-bound simulations (Fig. 1 *B*; Table S1). The $B^0AT1$ transporter solved in the cryoEM structure was not included in our simulations to computationally probe the intrinsic dynamics of ACE2. $B^0AT1$ is mainly expressed in kidneys and intestines (50), whereas ACE2 can also be found in the lungs and heart tissues, supporting the likelihood that ACE2 can be found uncomplexed with $B^0AT1$ upon cellular recognition and binding to S.

The simulations reveal an exceptional structural plasticity of the full-length ACE2 homodimer, pinpointing a large tilting of the head relative to the TM domain as well as profuse mobility of the TM helix in the membrane. Remarkably, the homodimer interface at the level of the neck domains remains stable despite the dramatic motions as well as the ACE2-RBD contacts, emphasizing the high affinity interaction between them. A systematic characterization of glycan-protein and glycan-glycan contacts indicates a possible role of glycan N53 in both homodimer and heterodimer interactions. Overall, the RBD does not seem to significantly affect the dynamics of ACE2 compared with the apo state, although that might differ in the presence of the full-length spike. Taken together, the remarkable ACE2 flexibility indicates a mechanical contribution to the S1 or S2 conformational changes required for cellular fusion and infection and suggests the structural basis for the possibility of finding two or more ACE2 complexes bound to the same S glycoprotein with two or more "RBD-up" conformations.

## METHODS

### ACE2 system construction

Coordinates of the ACE2-RBD complex were taken from the full-length cryoEM structure (Protein Data Bank, PDB: 6M17) (16), removing the coordinates from the co-complexed $B^0AT1$ dimer. Missing C-terminal residues of the ACE2 transmembrane helices were modeled using I-TASSER (51–53) based on the known sequence (residues 769–805), whereas missing N-terminal residue coordinates (residues 19–21) were copied from PDB: 6M0J after the alignment of the N-terminal helix. Zinc coordinating residues and a coordinating water molecule were taken from PDB: 1R42 because the zinc coordination site is poorly resolved in PDB: 6M17.

ACE2 and RBD glycosylation was defined according to glycoanalytic data (47,49) and modeled using the Glycan Reader & Modeler tool (54) integrated into Glycan Reader (55) in CHARMM-GUI (56). In total, seven complex, bi-antennary N-glycans and one O-glycan were added to ACE2, as well as 1 N glycan to RBD (Table S1). Only one O-glycan was included at site 730 because analytic data suggest extremely low stoichiometry at the other O-glycosylation sites (48). The apo ACE2 model was created by deleting the RBDs of the complete ACE2-RBD model.

### Membrane modeling

The plasma membrane modeled in this study was composed of 56% POPC, 20% CHL, 11% POPI, 9% POPE, and 4% PSM. The lipid composition was estimated based on the known lipid compositions of mammalian cellular membranes (57,58). It is hypothesized that phospholipids containing charged headgroups such as PI and PS are more likely to face the cytoplasmic side of the membrane and are additionally thought to aid in tolerance of the increased membrane curvature (57). Using a precedent set by a 2014 coarse-grained molecular dynamics study of the asymmetrical mammalian plasma membrane, the lipids were partitioned according to the outer versus inner leaflet enrichment factors of 2.0, 1.2, 0.0, 0.25, and 2.0 for POPC, CHL, POPI, POPE, and PSM, respectively (58). To reduce the chemical complexity of the system for simulation purposes, PS lipids were not included in these calculations. The small percentage (4%) of PS recorded in the literature is represented in the membrane by PI lipids.

An asymmetric 350 × 350 Å lipid bilayer according to the above specifications was generated using CHARMM-GUI's input generator (56). The lipids were packed to an approximate equilibrium area per lipid of 63 $Å^2$. Before insertion of ACE2 and subsequent trimming, the membrane patch contained a total of 2432 POPC, 870 CHL, 460 POPI, 404 POPE, and 128 PSM lipids.

### System preparation and molecular dynamics simulations

Histidine protonation states at pH 7.0 were verified using PROPKA on Maestro (Schrödinger, New York, NY). The models were parametrized using PSFGEN and CHARMM36 all-atom additive force fields for protein, lipids, and glycans (59) fully solvated in TIP3P water boxes (60) with 150 mM NaCl. The total number of atoms is 738,696 for the apo system (size: 18.7 × 18.9 × 23.7 nm) and 783,954 for the RBD-bound system (size: 18.7 × 18.9 × 25.1 nm).

MD simulations were performed on the Frontera computing system at the Texas Advanced Computing Center (TACC) using NAMD 2.14 (61). The systems were first subjected to 10,000 steps of conjugate gradient minimization in which the protein, glycans, lipid heads (P atoms for POPC, POPI, POPE, and POPS and $O_3$ atoms for CHL), solvent, and ions were kept fixed. The temperature was subsequentially increased incrementally from 10 to 310 K for 0.5 ns at 1 fs/step for the lipids tails to equilibrate using the NVT ensemble. This was followed by simulation at the NPT ensemble at 1.01325 bar and 310 K for 0.5 ns using 2 fs/step. The systems were then simulated for 2500 minimization steps and an additional 0.5-ns simulation with the protein and glycans harmonically restrained at 5 kcal/mol to allow for environment relaxation. Finally, the systems were further equilibrated for 0.5 ns with all restraints released. Production runs were conducted in triplicates of 1 $\mu$s for each system. All simulations were performed using periodic boundary conditions and particle-mesh Ewald (62) with a 12-Å cutoff and a switching distance of 10 Å for evaluation of nonbonded interactions. The SHAKE algorithm (63) was employed to constrain bonds involving hydrogen atoms.

### Analysis

Distributions and average values were calculated considering the total sampling in the three replicas for each system. Except when the dimer structure or interdimer contacts were analyzed, the monomer sampling was accumulated such that the calculations correspond to 6 $\mu$s of combined sampling (three replicas × two monomers/replica). Reported uncertainties of average values correspond to the standard deviation across the replicas.

#### *ACE2 angles and distances calculations*

To quantify the range of motion of ACE2 in the simulations, several angle and distance metrics were developed. Calculation was performed using MDTraj (64) with visualization through VMD (65). The 6M17 cryoEM structure was used as the reference structure.

The head tilt angle relative to the transmembrane domain was calculated by first aligning the dimer's coordinates to the reference cryoEM TM domains, the angle calculated between the centers of mass of the reference's dimer PDs (residues 18–600), reference's TM helices (residues 747–774), and monomer's PD at each frame in the simulation. The helix tilt angle was computed as the angle between a vector defining the membrane's normal and a vector connecting residues 741 and 765 at the extremities of the helix.

Revolution angle was calculated between the center of mass of the monomer's PD in the reference conformation, the center of mass of the reference's TM domain, and the center of mass of the monomer's PD at each frame in the simulation after the alignment of the monomer's TM helices. Buckling angle was calculated using the $xy$ projections of the center of mass of monomer's A PD at frame $f$, the center of mass of the reference dimer's PDs, and center of mass of monomer's B PD at frame $f$ after alignment of the trajectories to the reference dimer neck domains (residues 617–726).

Distance between the monomer's head domains in the homodimer was calculated by determining the distance between each monomer's PD center of mass. Distance between the head domain and membrane corresponds to the minimal distance between the PD's heavy atoms and membrane's phosphorous atoms at each frame of the simulation. Distance between TM helices was calculated based on the distance between their centers of mass.

### Principal component analysis

Principal component analysis was performed using the Scikit-learn library in python 3.6 (66). The monomers in each of the ACE2 homodimer trajectories were saved separately and all aligned to the backbone of the transmembrane helix, residues 747–774. The simulation coordinates of the apo and RBD-bound systems were concatenated and used to fit the transformation function, so that both systems were transformed in the same principal component space.

### Root mean-square fluctuation and secondary structure analysis

Root mean-square fluctuation (RMSF) of C$\alpha$ carbons was calculated using CPPTRAJ (67). Calculations were performed for the head and TM domains separately, after alignment of the backbone atoms of the respective domain. Secondary structure calculation was performed with MDTraj's compute_dssp function (64) using the simplified three-category assignment scheme. The proportion of frames in which each residue's secondary structure was assigned as a coil, helix, or strand was computed and compared with the assignment obtained for the starting cryoEM structure.

### Fraction of native contacts and glycan contacts

Fraction of native contacts was calculated according to Mehdipour and Hummer (68). The 6M17 cryoEM structure was used as the reference structure for identification of native contacts. The ACE2-RBD interface was subdivided into three interacting regions according to the interacting residues pairs listed on Table S2.

A systematic characterization of contacts established by each glycan in the system was performed using MDTraj (64), using a cutoff of 3.5 Å between the heavy atoms.

## S model construction

The spike model was obtained from our previous simulations (69). The simulations included only the solvated spike. All atoms except for the spike protein and glycans were removed, along with the lower part of the stalk region of each protomer (residues 1165–1273). Because the cryoEM model was missing density for portions of the RBD, we replaced the RBD coordinates (residues 355–494 for closed RBD, 339–523 for open RBD) of each protomer in both models with the RBD coordinates from the crystal structure of the RBD bound to ACE2 (PDB: 6M0J (70)). The RBD structure from 6M0J was aligned with the backbone heavy atoms (alpha-carbon, carbonyl-carbon, and nitrogen) of each RBD in the initial spike model. We then grafted the RBD coordinates onto the spike at the hinge region, which resolved missing loops as well as introducing a disulfide bond in the RBD. The remaining disulfides not resolved in the cryoEM structures were assessed based on distance criteria and sequence conservation. The system was built using the *ff14SBonlysc* (71) and GLYCAM (72) force fields for the protein and glycan atoms, respectively. These were explicitly solvated in OPC3 water (73) with a 200 mM NaCl buffer (74). The RBD-up and -down systems both consisted of 1,298,646 atoms, and were simulated on Frontera at TACC, and SDCC at BNL using the pmemd.CUDA module of Amber20. The spike systems were equilibrated using a 10-step protocol. First, the water molecules were minimized for 1000 steps using steepest descent, and then for an additional 9000 steps with conjugate gradient, whereas the rest of the system was positionally restrained with 1 kcal/(mol × Å$^2$) restraints. The systems were then heated to 310 K at constant volume, again with all atoms except hydrogens and waters restrained with 100 kcal/(mol × Å$^2$) positional restraints. The box size and density were then equilibrated over 1 ns with constant pressure, with the same positional restraints as the previous step. The restraints were then lowered to 10 kcal/(mol × Å) for an additional 1 ns of equilibration, before a second minimization. This minimization consisted of 10,000 steps of conjugate gradient with positional restraints now only applied to backbone atoms (alpha-carbon, carbonyl-carbon, and nitrogen), using a force constant of 10 kcal/(mol × Å$^2$). The next three steps of equilibration were MD for 1 ns each at constant NPT with positional restraints on protein backbone atoms at 10, 1, and 0.1 kcal/(mol × Å$^2$), respectively. This was followed by a final 1 ns of unrestrained MD at constant NPT before beginning production.

To generate a three-up model of the prefusion spike protein, steered molecular dynamics (SMD) was used as implemented in the Amber NFE toolkit (75). The initial structure was an all-closed model from the equilibration described above. To generate a structure used as reference for steering, the open monomer from the one-up model was aligned to the other two closed monomers by overlapping S2 domains. The closed RBD domains were then replaced with the open model of the aligned monomer. No equilibration was performed on the resulting reference structure because it was not subjected to simulations, and only used as a steering target. Root mean squared deviation (RMSD) was used as the collective variable (CV) to guide SMD. A separate CV was used for the opening of each RBD. In each CV, the RMSD region includes the RBD (residue 338–517) and three helices in S2 (residue 747–782, 946–966, 987–1034). The three RMSDs were simultaneously gradually decreased by SMD from their initial values to 0 during 20 ns of simulation time at 310 K in the NVT ensemble (Fig. S1), with a 4-fs time step enabled by hydrogen mass repartitioning (76), using a spring constant of 10,000 kcal/mol/Å$^2$. Final RMSD values were 0.19, 0.19 and 0.20 Å for the three monomers. Weak (1 kcal/mol/Å$^2$) positional restraints were applied to the S2 helices, which were relatively stable during SMD simulations.

## Data availability

Additional supporting research data of our models to enable other groups to use and explore this dynamic system in atomic detail (77) may be accessed through the NSF MolSSI and BioExcel COVID-19 Molecular Structure and Therapeutics site at https://covid.molssi.org/.

## RESULTS

## The ACE2 dimer shows pronounced flexibility

Simulations of RBD-bound and apo ACE2 evidenced a striking degree of flexibility in the ACE2 homodimer. With respect to the fairly vertical, extended conformation of the initial cryoEM structure (16), the most striking fluctuation observed during the simulations is characterized by a tilt of the head relative to the long axis of the respective monomer's transmembrane helix. Although each monomer in the reference cryoEM structure displays a tilt angle of 16°, structures in the simulations sample tilt angles that range from 0 to 50° (Fig. 2, *A* and *D*; the distribution is calculated over the combined sampling of the two monomers in the three replicas of each system). This tilt motion, combined with an overall "shrinking" of the initial extended conformation, moves the head toward the membrane, with head-membrane distances varying from 30 to 84 Å and the great majority of conformations (98 and 98.6% of the frames of apo and RBD-bound simulations, respectively) exhibiting the head domain closer to the membrane than the starting cryoEM structure (Fig. 2 *B*). The presence of the RBD does not seem to affect

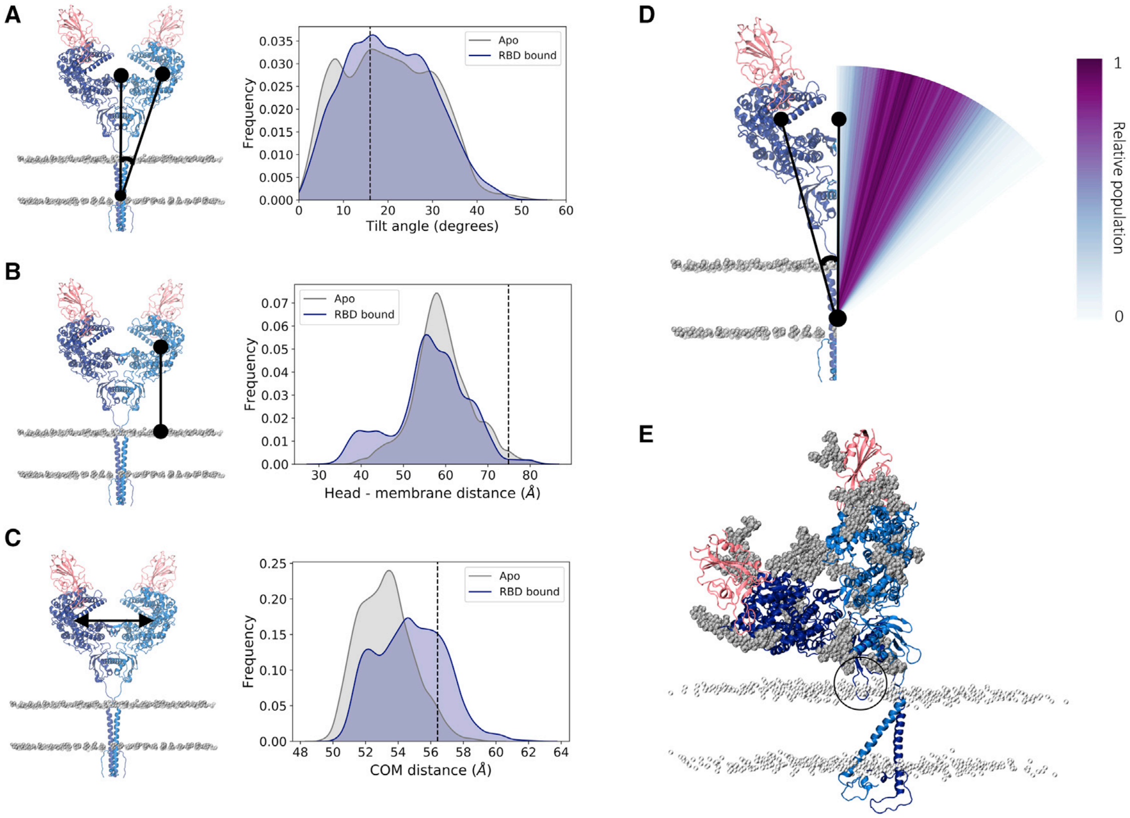


FIGURE 2 Tilt motion of ACE2. (*A*) Right panel shows head tilt angle distribution relative to the transmembrane domain long axis for apo (*gray*) and RBD-bound (*navy*) simulations, accumulated over the two monomers in the three replicas of each system. The angle value corresponding to the cryoEM conformation is indicated by a black line. Left panel shows a representation of the metric, with ACE2 monomers colored dark and light blue, RBDs colored pink, and phosphorus atoms from membrane's lipid heads shown in gray in a van der Waals representation. (*B*) Distribution of the minimal distance between peptidase domain (PD) and membrane. (*C*) Distribution of the ACE2 monomer heads' center of mass distance. (*D*) Visual representation of the tilt angle distribution for the RBD-bound simulations with a color gradient according to the relative population. (*E*) Example of a highly tilted ACE2 homodimer conformation sampled in the simulation. The black circle highlights a membrane-inserted neck domain loop. The ACE2 and RBD glycans are shown in gray, and the membrane's lipid heads are shown in a silver transparent representation. To see this figure in color, go online.

the dynamics, with apo and RBD-bound simulations showing average head-membrane distances of 59.1 ± 6.9 and 56.5 ± 8.8 Å, respectively.

Remarkably, the head tilt motion occurs in a concerted fashion between the monomers, such that as one monomer bends toward the membrane, with large tilt angle values, the other monomer follows this deformation by adopting a more extended conformation with lower angle values (Fig. 2 *E*; Fig. S2). Accordingly, the distance between the head domains fluctuates only slightly, varying by no more than 6 Å from the reference structure (Fig. 2 *C*) and resulting in a stable relative position of the heads within the homodimer. The majority of the conformations display the two heads slightly closer to each other than in the starting cryoEM structure, whereas the presence of the spike RBD shifts the distribution slightly toward more open conformations. In fact, a very small fraction of the conformations access dimer distances equivalent to the head separation observed in the cryoEM open dimer conformation (head center of mass distance of 60 Å (PDB: 6M1D)) (16), but the distribution indicates that this is a rare state.

In concert with the tilt motion described above, the ACE2 head also undergoes displacement in the *xy* plane around the long axis of the TM helix, as shown for one of the replicas in Fig. 3 *A*. Computation of each monomer's revolution angle suggests a twisting of the flexible linker that connects the neck to the TM domain, with several significant alternative conformations exhibiting almost 180° rotation of the head from its starting position (Fig. 3 *B*). It is important to highlight that the head revolution is measured here for each monomer independently, after an alignment of that monomer's transmembrane domain. Because the monomer's TM helices are not in contact with each other and thus can move independently in the membrane (Fig. 4 *A*), the twisting motion of one of the monomers is not necessarily accompanied by an equivalent twist of the other monomer, avoiding the entanglement of the flexible linkers around each other. Instead, visual observation indicates that the other monomer revolves as a whole around the transmembrane helix of the twisting monomer (Fig. S3), keeping the head dimer interface overall intact. Thus, despite this pronounced motion, the other monomer follows the twist

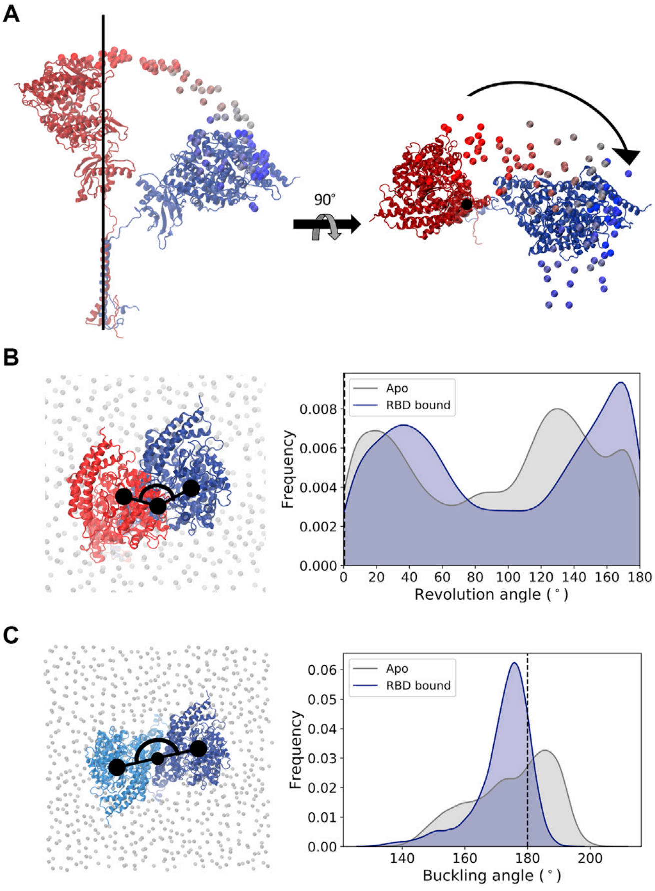


FIGURE 3 ACE2 revolution relative to a plane perpendicular to the transmembrane helix's long axis. (*A*) Representation of a monomer's degree of flexibility in one of the replicas, showing the time evolution of the position of the C$\alpha$ atom of Gln325 colored from dark red (t = 0) to dark blue (t = 1000 ns). Conformations aligned to the cryoEM's reference TM domain C$\alpha$ atoms shown in a van der Waals representation and initial and final monomer conformations shown in a cartoon representation. (*B*) Head revolution angle distribution for apo (*gray*) and RBD-bound (*navy*) simulations. The angle value corresponding to the cryoEM conformation (revolution angle = 0) is indicated by a black line. Left panel shows a representation of the metric, with monomer's head initial position shown in red, the same monomer at a time t in dark blue, and phosphorus atoms from membrane's lipid heads shown in gray in a van der Waals representation. (*C*) Relative orientation of the monomer's head in the heterodimer. The ACE2 monomers are colored dark and light blue. To see this figure in color, go online.

by retaining the heads' symmetry, and the relative angle between the heads in the *xy* plane remains close to the initial 180° (Fig. 3 *C*).

Principal component analysis of the TM-aligned monomers confirm that the tilt and revolution motions correspond to the greatest variance of the ACE2 head coordinates and emphasize that the head dynamics is a combination of these two main motions: principal components 1 and 2 describe the displacement in two orthogonal planes relative to the TM helix axis and account for 51 and 32%, respectively, of the variance (Fig. S4). The head motion occurs at fast timescales in the order of nanoseconds in both of our apo and RBD-bound simulations, although we expect that this tilting timescale would be longer when considering the full spike bound to ACE2.

Besides the above-described global motions, at a residue level, higher fluctuations are observed for head residues located at the homodimer interface, with particularly high fluctuations at the loop regions Pro336-Val343 and Lys625-Tyr633 (Fig. S5 *A*), which will be discussed in more detail in the following sections. Despite the global motions and backbone fluctuations, the protein secondary structure is overall stable, retaining the motifs identified in the cryoEM structure (Figs. S6 and S7).

The simulations indicate that the conformational variability of ACE2 occurs not only because of flexibility at the linker connecting the transmembrane and head domains but also because of the motions of the transmembrane helix in the membrane. In contrast to other multimeric transmembrane domains, such as the coiled coil trimer of S (21,26,69), each ACE2 monomer is anchored to the membrane by a single helix that does not interact with that of the opposing monomer (Fig. 4 *A*) but rather explores a range of tilt angles relative to the membrane's normal (Fig. 4 *B*). This tilt motion, however, does not affect the overall integrity of the transmembrane helix, as indicated by RMSF and secondary structure characterization of the TM-aligned conformations (Figs. S5 *B*, S6 *C*, and S7 *C*).

The overall gaussian distributions of the distances and angles measured here emphasize a continuous sampling of the distinct conformations, with no significant energy barriers hindering the transitions. Combined with the multidirectional tilting of each monomer (Fig. S2), the simulations indicate that the deformations occur transiently and with no preferred direction or conformation. Although some distinctions in the geometric descriptor distributions of the apo and RBD-bound simulations can be observed (such as in Fig. 4 *A*), the absence of clear differences and the pronounced protein dynamics do not allow us to exclude the possibility that these differences are simply due to sampling in the simulations and not necessary functionally important. Taken together, our results suggest that the experimentally resolved extended ACE2 structure is likely not a dominant conformation in solution, and the homodimer displays a large ensemble of conformations in the native state.

## ACE2-RBD interface remains stable despite the large ACE2 motions

Despite the dramatic flexibility of the ACE2 dimer, the RBD included in the RBD-bound model retained a large

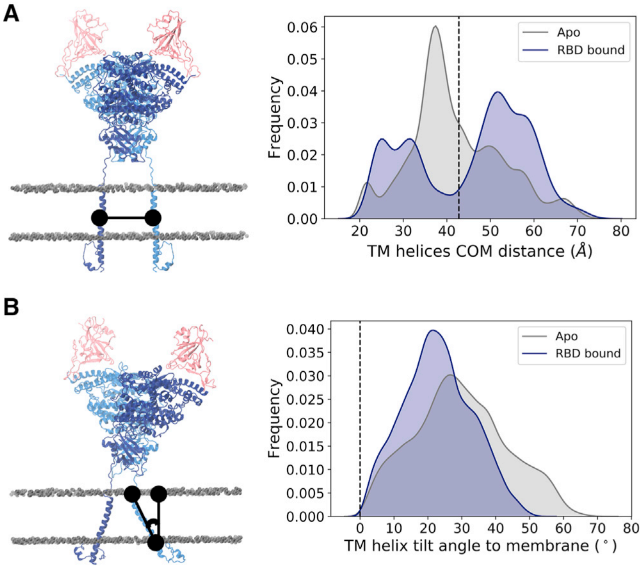


FIGURE 4 Transmembrane helix dynamics. (*A*) Distance between the center of mass of each monomer's TM helix. (*B*) TM helix tilt angle relative to the membrane's normal. The values corresponding to the cryoEM conformation are indicated by a black line. Left panels show representation of the metric, with ACE2 monomers colored dark and light blue, RBDs colored pink, and phosphorus atoms from membrane's lipid heads shown in gray in a van der Waals representation. To see this figure in color, go online.

fraction of the native contacts with ACE2 throughout the simulations, with an average fraction of 0.87 ± 0.11 contacts. Although the interface is thus overall stable and the relatively small RBDs accompany the range of motion of ACE2, dividing the RBD-ACE2 interface into three interacting regions (comprising of the two RBD loop regions at the opposite sides of the dimer interface and the central region containing the two short $\beta$-sheet strands; Fig. 5 *A*; Table S2), we observe that the central region 2 contains the most stable contacts, whereas regions 1 and 3 at the extremities of the interface are less tightly bound, and they sample states with a smaller number of native contacts (Fig. 5 *B*). Analogously, higher backbone fluctuations are observed for the extremities of the RBD binding interface, particularly in the long loop in region 1 (Fig. S5 *C*). Taken together, these observations are in agreement with the rocking motion dynamics at the PD-RBD interface observed in other simulations (41,42).

In addition to protein interactions, five glycans in ACE2 are in close proximity to the RBD and have been suggested to play a role in S binding. In agreement with other studies (48,68), N90, and to a lesser extent N322, of ACE2 establish contacts with RBD. Besides these glycans, we also find that N53 can form a large number of contacts with the RBD residues, whereas the RBD glycan N343 makes very few contacts with ACE2's head protein residues (Fig. 5 *C*) or glycans (Fig. S8).

## N53 is involved in both homodimer and heterodimer contacts

Besides the ACE2-RBD heterodimer interface, we considered the interactions within ACE2 that could contribute to maintaining such a stable homodimer head interface despite the pronounced flexibility of the protein. Experimental structures and simulations suggest that the majority of the protein contacts in the homodimer are located in the neck domain, with only two other interactions, in the form of hydrogen bonds, that are observed in the larger PD (16,68). In agreement with these observations, we find that the dimer interface is mainly held together in the simulations by residues at the neck (Fig. 6 *A* for RBD-bound simulations and Fig. S9 *A* for apo). Computation of the glycan-protein and glycan-glycan contacts enrich the characterization of the inter- and intramonomer interactions and suggest that although the eight ACE2 glycans form several contacts with protein residues located within the same monomer (Fig. 6 *B*), protein-glycan interactions with the opposite monomer are limited to N53 and N690 (Fig. 6 *C*). Additionally, N53 is the only glycan that can be found to form intermonomer glycan-glycan contacts established between the equivalent N53 copies in each of the monomers (Figs. 6 *D* and S10). In this sense, it is interesting that computational predictions indicate that the disruption of the N53 glycosylation motif because of the mutation at T55, leading to removal of the glycan, can have a destabilizing effect on ACE2 stability (78). Similar to the lack of RBD effect on the ACE2 dimer flexibility, we find that the homodimer contact distributions are comparable between apo and RBD-bound states of ACE2 (Fig. S9).

Our systematic analysis of all glycan interactions thus indicate that N53 is the only glycan involved in both heterodimer (ACE2-RBD; Fig. 5 *C*) and homodimer (intra-ACE2 dimer; Fig. 6 *D*) interactions. The opposite positions of the homodimer and RBD interfaces relative to N53 (Fig. 5 *E*) could suggest a competition for N53 contacts within these dimer interfaces. Indeed, we find that when a large number of N53-RBD contacts are formed, the N53 dimer interface is usually abrogated in the RBD-bound simulations and vice versa (Fig. S11 *B*). However, these interactions are not completely mutually exclusive because N53' (the N53 glycan in monomer B) can be found involved in both intermonomer and monomer-RBD interactions in replica 1. This glycan's flexibility is likely the cause of the high backbone fluctuations observed for the neighboring Pro336-Val343 loop region (Fig. S5 *A*) and probably plays a role in conferring a transient nature to the homo- and heterodimer interactions because the N53 homodimer contacts are not consistently formed even in the absence of the competing heterodimer in the apo ACE2 simulations (Fig. S11 *A*). Thus, N53 can still be found highly solvent exposed in the apo state, suggesting optimal conformations for contact with an RBD partner and a role in S binding and infection.

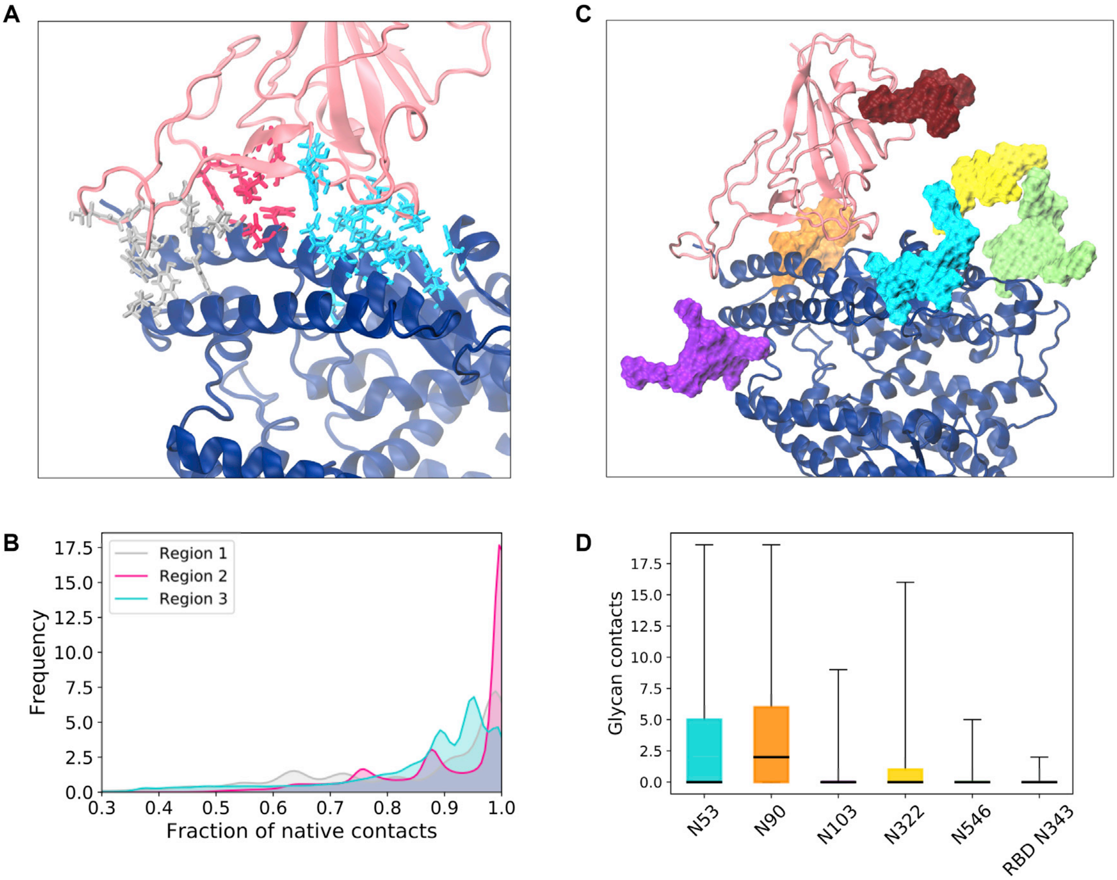


FIGURE 5 ACE2-RBD interactions. (*A*) Residues in the ACE2-RBD interface colored according to the regions classification (*silver*, *magenta*, and *cyan*) and shown with licorice representation. The ACE2 monomer is shown with dark blue cartoons and the RBD with pink cartoons. Glycans have been omitted from (*A*) for clarity. (*B*) Distribution of the fraction of native contacts in each of the interaction regions. The colors are the same as in (*A*). (*C*) Glycans in the ACE2-RBD interface, shown with surface representation with the following color scheme: N53 (*cyan*), N90 (*orange*), N103 (*purple*), N322 (*yellow*), N546 (*lime*), and N343 (*dark red*). (*D*) Box plot of the number of glycan-protein contacts for the interface glycans shown in (*C*), using the same color scheme. The horizontal black lines indicate the mean value, boxes extend to the lower and upper quartiles, and whiskers show the total range of the data. To see this figure in color, go online.

## DISCUSSION

All-atom simulations of apo and RBD-bound, full-length, membrane-embedded ACE2 show a striking degree of fluctuation of the homodimer protein, which can be attributed to hinge motions of the large head domain relative to the transmembrane helices, and tilt of the transmembrane helices relative to the membrane's normal. The head relative motion is due to the flexible linker region connecting the TM helix and the neck domain, whereas the TM helix motion points to a loose anchoring of ACE2 to the membrane. Although the two (head and transmembrane) domains are internally stable, the flexibility of the connecting loop virtually results in a decoupling of these domains' dynamics (Fig. 7 *A*), leading to sampling of conformations strikingly different than the experimentally observed elongated structure (16). The observed highly tilted conformations position the head domains in the vicinity of the membrane. Although the glycans do not make significant contacts with the membrane lipid heads (with the exception of O730, which is located close to the membrane even in the elongated conformation; Fig. S12), insertion of a loop from the neck domain into the membrane can be observed in the simulations (residues 625–633; Figs. 1 *E* and S13 *A*). Curiously, this loop region exhibits the highest head RMSF values (Fig. S5 *A*). The insertion seems to be driven by hydrophobic contacts involving the side chain of Leu628 and the lipid tails and suggest the existence of lipid-protein interactions that can stabilize the bent conformations (Fig. S13 *B*).

In accordance with our findings, a high deformation propensity was also observed for the TM-neck linker upon normal mode analysis of full-length ACE2 (79). These distinct conformations were likely only observed in the simulations because of the removal of $B^0AT1$ from the cocomplexed structure because they seem to bind on opposite sides of the homodimer transmembrane interface and interact with the flexible linker (16). The presence of $B^0AT1$ would thus likely reduce the observed ACE2 flexibility, restraining the sampling of highly tilted conformations. However, the expression profiles of ACE2 and $B^0AT1$ suggest the

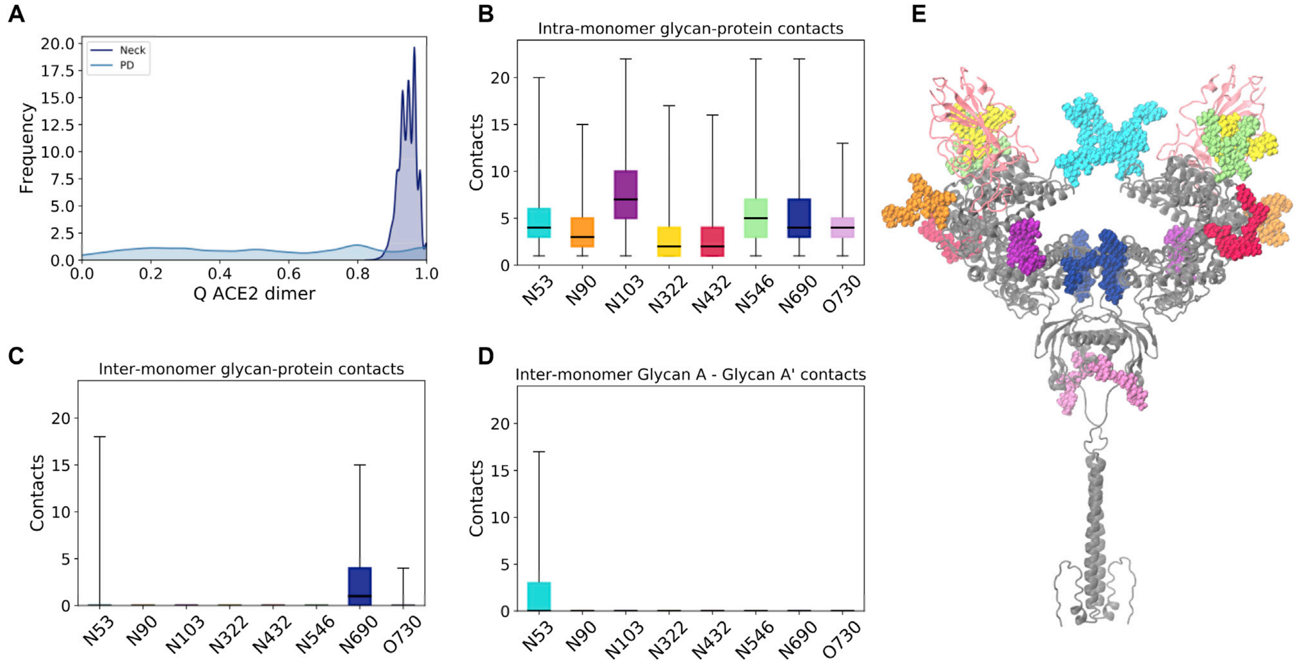


FIGURE 6 ACE2 homodimer contacts for RBD-bound simulations. (*A*) Fraction of the native contacts between ACE2 monomers considering only protein components of the glycoprotein. The neck and PD interacting regions are depicted separately. (*B*) Total glycan-protein interactions formed within each ACE2 monomer, per glycan. The horizontal black lines indicate mean value, boxes extend to the lower and upper quartiles, and whiskers show the total range of the data. (*C*) Glycan-protein contacts between glycans in one of the monomer and protein residues in the opposite monomer. (*D*) Glycan-glycan contacts between the glycan in one of the monomers (glycan A) and its copy in the other monomer (glycan A'). (*E*) The ACE2 dimer with glycans in a van der Waals representation is colored according to (*B*)–(*D*). The ACE2 protein dimer is colored gray and RBDs in light pink. To see this figure in color, go online.

likelihood of ACE2 existing in the apo state, especially in lung and heart tissues where $B^0AT1$ is not expressed (50). The observed high degree of ACE2 flexibility in the apo form can additionally explain the difficulty in the experimental characterization of the full-length structure.

Remarkably, the conformations of the ACE2 peptidase and neck domains remain stable throughout the simulations, and the homodimer heads retain their relative orientation despite the dramatic global homodimer motions. In a similar fashion, the RBDs included in the holo simulations remained tightly bound to ACE2 throughout the simulations, evidencing the high affinity between them. Glycan-glycan and glycan-protein interactions suggest that the ACE2 homodimer interface is maintained not only via protein interactions at the neck domain but also via intermonomer contacts involving N53 at the top of the PD and N690 closer to the neck. Interestingly, N53 also makes extensive contacts with RBD, suggesting a dual and possibly competing role between homodimer (intra-ACE2) and heterodimer (ACE2-RBD) interactions. This dual nature may be dependent on the length of the N53 glycan, but the glycosylation heterogeneity in ACE2 in general and at this position in particular (49) emphasizes the possibility of intermonomer glycan interactions. Even in the absence of RBD, N53 alternates between being sequestered in homodimer contact and being extended and highly solvent accessible, suggesting a role in RBD binding to this glycan.

Because of the stability of the head domain and RBD interface despite ACE2 body motion, this large flexibility would remain invisible in studies that do not take the full-length character of ACE2 into account, looking, for instance, at only the PD-RBD interactions. However, a recent cryoEM study of S-ACE2 PD complex resolved a continuous swing motion of the ACE2-head-RBD relative to the S trimer body (45). These structural characterizations complement our analysis and suggest how the ACE2 motion would translate in the context of full-length spike. Additionally, it has been found from in situ cryoEM and molecular dynamics simulations that the spike glycoprotein also exhibits conformational plasticity, with hinge motions at three different regions of the stalk trimer (44). Large dynamical variations thus seem to be features of these extracellular glycoproteins. The RBD rocking motion and S conformational variability have been proposed as mechanisms for immune evasion and efficient receptor search in the host cell (44,45), but the similar rocking motion of ACE2 we observed also suggests a mechanical aspect to ACE2-S interaction. The process of S conformational transition upon binding to the receptor and cell fusion remains elusive, but ACE2's intrinsic flexibility could promote a large swinging motion of the ACE2-S1 complex, providing a mechanical force for the approximation of the two membranes and shedding of S1 toward fusion of the S2 domains into the receptor cell (Fig. 7 *B*).

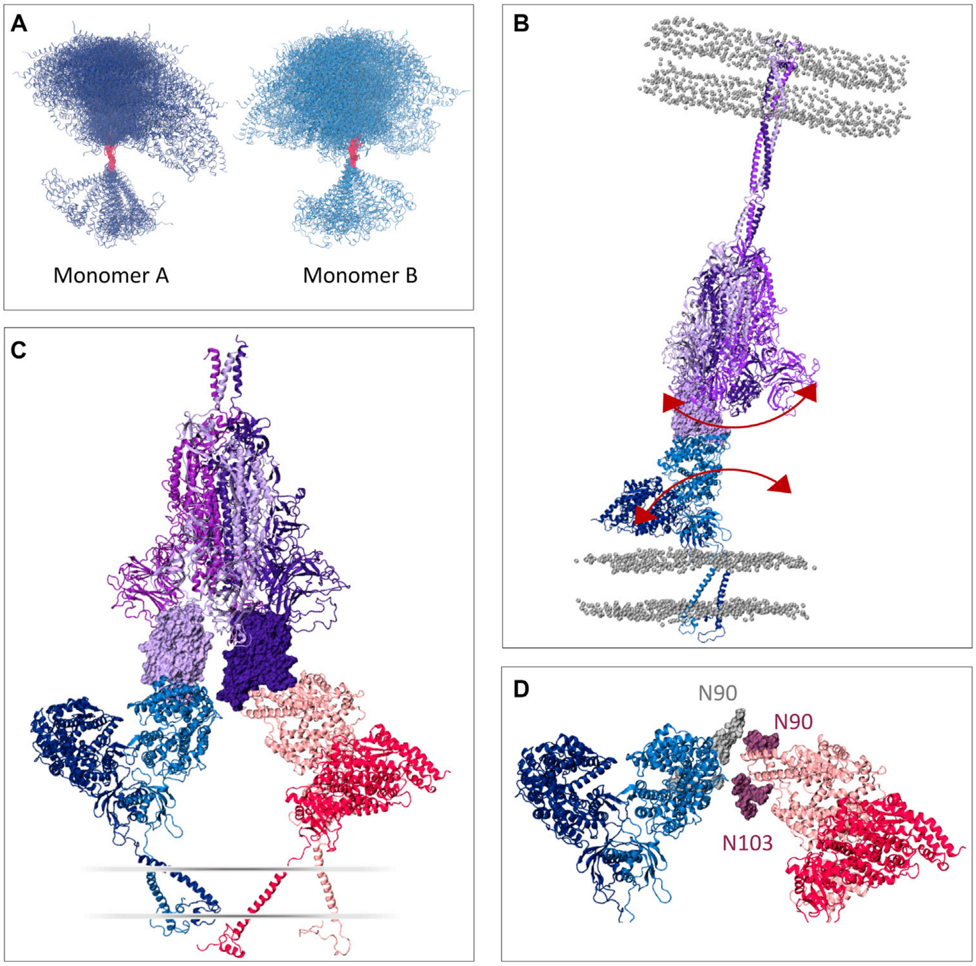


FIGURE 7 ACE2 flexibility's impact on S interaction. (*A*) ACE2 monomer conformations taken from equally spaced frames from the simulations, aligned via the flexible linker. (*B*) Proposed effect of ACE2's flexibility on the spike's dynamics, communicated through the ACE2-RBD complex. The three chains in the spike model are colored in different shades of purple, with the "up" RBD shown in light purple in surface representation. The phosphorus atoms from membrane's lipid heads are shown in gray in a van der Waals representation. The red arrows indicate the proposed effect of ACE2 flexibility in the ACE2-S complex dynamics. (*C*) Proposed complex of two ACE2 dimers bound to a single spike, with two RBDs in the "up" conformation. The ACE2 dimers are shown in dark and light blue and dark and light pink, respectively. The RBDs are shown in a surface representation. A schematic of the membrane is indicated. (*D*) Detailed view of the ACE2 heads in (*C*), with N90 and N103 glycans highlighted. To see this figure in color, go online.

Finally, one can speculate that the flexibility of the host receptor might allow the accommodation of more than one ACE2 dimer bound to a single S with two or more RBDs in the up conformation. A high efficiency of ACE2 usage was suggested to contribute to SARS-CoV transmissibility (15,80) and thus could be at play for SARS-CoV-2 as well. To investigate this possibility, we extracted a range of ACE2 conformations from the RBD-bound simulations covering different tilt angles and explored the alignment of these structures to a "three RBD-up" spike model. Indeed, we find that two ACE2 dimers can sterically be accommodated by a single spike with interdimer backbone distances no smaller than 10 Å (Fig. 7 *C*). The flexibility of the homodimers could potentially allow for even three ACE2's per S, opening the possibility of multireceptor usage by the spike glycoprotein for host cell infection. Explicitly considering the glycans in this aligned model evidences that N103 and especially N90 are in close proximity to the neighboring ACE2 dimer (Fig. 7 *D*). Genomic, mutagenesis, and binding studies of ACE2 variants have demonstrated that removal of the N90 glycan due to mutations to the glycosylation motif leads to increased S-ACE2 binding affinity (80–82), and these observations can thus provide the structural basis for the negative effect of N90 on RBD binding.

## CONCLUSIONS

All-atom molecular dynamics simulations of the full-length ACE2 inserted in a mammalian-inspired lipid membrane uncover a significant degree of flexibility of the ACE2 homodimer with consequences for S-ACE2 interaction and SARS-CoV-2 infection and suggest the structural basis for glycan N90's negative effect on RBD binding. Additionally, we identify the involvement of glycan N53 in ACE2 homodimer and ACE2-RBD heterodimer contacts. Taken together, our findings shed further light onto the mechanisms of viral binding and cell entry required for rational design of effective SARS-CoV-2 therapeutics.

## SUPPORTING MATERIAL

Supporting Material can be found online at https://doi.org/10.1016/j.bpj.2020.10.036.

## AUTHOR CONTRIBUTIONS

E.P.B., L.C., and Z.G. built the model of the ACE2 and RBD. A.C.D. built the membrane bilayer. E.P.B. performed the MD simulations and analysis. Y.W. built the three-up RBD, L.F. converted the spike and glycans from Charmm to AMBER and GLYCAM force fields, L.F. performed the MD of the spike structures, and K.B. did the structure grafting. R.E.A. and C.S. designed and oversaw the research project. E.P.B. wrote the article, with contributions from all authors.

## ACKNOWLEDGMENTS

We are grateful for the efforts of the TACC Frontera team and for the compute time made available through a Director's Discretionary Allocation. We thank Prof. Jason McLellan and his group at the University of Texas, Austin, for helpful discussions.

This work was supported by National Institutes of Health GM132826, National Science Foundation RAPID MCB-2032054, an award from the Research Corporation for Science Advancement, and a University of California, San Diego, Moore's Cancer Center 2020 SARS-COV-2 seed grant. L.C. is funded by a Visible Molecular Cell Consortium fellowship. The Director's Discretionary Allocation was made possible by the National Science Foundation award OAC-1818253.

## REFERENCES

1. Hoffmann, M., H. Kleine-Weber, …, S. Pöhlmann. 2020. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell.* 181:271–280.e8.
2. Zhou, P., X. L. Yang, …, Z. L. Shi. 2020. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature.* 579:270–273.
3. Wang, Q., Y. Zhang, …, J. Qi. 2020. Structural and functional basis of SARS-CoV-2 entry by using human ACE2. *Cell.* 181:894–904.e9.
4. Huang, C., Y. Wang, …, B. Cao. 2020. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet.* 395:497–506.
5. Wang, C., P. W. Horby, …, G. F. Gao. 2020. A novel coronavirus outbreak of global health concern. *Lancet.* 395:470–473.
6. Zhu, N., D. Zhang, …, W. Tan; China Novel Coronavirus Investigating and Research Team. 2020. A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* 382:727–733.
7. Chan, J. F.-W., S. Yuan, …, K. Y. Yuen. 2020. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet.* 395:514–523.
8. Donoghue, M., F. Hsieh, …, S. Acton. 2000. A novel angiotensin-converting enzyme-related carboxypeptidase (ACE2) converts angiotensin I to angiotensin 1-9. *Circ. Res.* 87:E1–E9.
9. Zhang, H., Z. Kang, …, H. Xu. 2020. Digestive system is a potential route of COVID-19: an analysis of single-cell coexpression pattern of key proteins in viral entry process. *Gut.* 69:1010–1018.
10. Turner, A. J., J. A. Hiscox, and N. M. Hooper. 2004. ACE2: from vasopeptidase to SARS virus receptor. *Trends Pharmacol. Sci.* 25:291–294.
11. Zisman, L. S., R. S. Keller, …, C. C. Canver. 2003. Increased angiotensin-(1-7)-forming activity in failing human heart ventricles: evidence for upregulation of the angiotensin-converting enzyme Homologue ACE2. *Circulation.* 108:1707–1712.
12. Raizada, M. K., and A. J. Ferreira. 2007. ACE2: a new target for cardiovascular disease therapeutics. *J. Cardiovasc. Pharmacol.* 50:112–119.
13. Crackower, M. A., R. Sarao, …, J. M. Penninger. 2002. Angiotensin-converting enzyme 2 is an essential regulator of heart function. *Nature.* 417:822–828.
14. Tikellis, C., S. Bernardi, and W. C. Burns. 2011. Angiotensin-converting enzyme 2 is a key modulator of the renin-angiotensin system in cardiovascular and renal disease. *Curr. Opin. Nephrol. Hypertens.* 20:62–68.
15. Li, F., W. Li, …, S. C. Harrison. 2005. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. *Science.* 309:1864–1868.
16. Yan, R., Y. Zhang, …, Q. Zhou. 2020. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science.* 367:1444–1448.
17. Kowalczuk, S., A. Bröer, …, S. Bröer. 2008. A protein complex in the brush-border membrane explains a Hartnup disorder allele. *FASEB J.* 22:2880–2887.
18. Zhong, N. S., B. J. Zheng, …, Y. Guan. 2003. Epidemiology and cause of severe acute respiratory syndrome (SARS) in Guangdong, People's Republic of China, in February, 2003. *Lancet.* 362:1353–1358.
19. de Wit, E., N. van Doremalen, …, V. J. Munster. 2016. SARS and MERS: recent insights into emerging coronaviruses. *Nat. Rev. Microbiol.* 14:523–534.
20. Li, W., M. J. Moore, …, M. Farzan. 2003. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature.* 426:450–454.
21. Wrapp, D., N. Wang, …, J. S. McLellan. 2020. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science.* 367:1260–1263.
22. Simmons, G., J. D. Reeves, …, P. Bates. 2004. Characterization of severe acute respiratory syndrome-associated coronavirus (SARS-CoV) spike glycoprotein-mediated viral entry. *Proc. Natl. Acad. Sci. USA.* 101:4240–4245.
23. Belouzard, S., V. C. Chu, and G. R. Whittaker. 2009. Activation of the SARS coronavirus spike protein via sequential proteolytic cleavage at two distinct sites. *Proc. Natl. Acad. Sci. USA.* 106:5871–5876.
24. Simmons, G., P. Zmora, …, S. Pöhlmann. 2013. Proteolytic activation of the SARS-coronavirus spike protein: cutting enzymes at the cutting edge of antiviral research. *Antiviral Res.* 100:605–614.
25. Song, W., M. Gui, …, Y. Xiang. 2018. Cryo-EM structure of the SARS coronavirus spike glycoprotein in complex with its host cell receptor ACE2. *PLoS Pathog.* 14:e1007236.
26. Walls, A. C., Y. J. Park, …, D. Veesler. 2020. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell.* 181:281–292.e6.
27. Kuba, K., Y. Imai, …, J. M. Penninger. 2005. A crucial role of angiotensin converting enzyme 2 (ACE2) in SARS coronavirus-induced lung injury. *Nat. Med.* 11:875–879.
28. Imai, Y., K. Kuba, …, J. M. Penninger. 2005. Angiotensin-converting enzyme 2 protects from severe acute lung failure. *Nature.* 436:112–116.
29. Kuba, K., Y. Imai, …, J. M. Penninger. 2010. Trilogy of ACE2: a peptidase in the renin-angiotensin system, a SARS receptor, and a partner for amino acid transporters. *Pharmacol. Ther.* 128:119–128.

30. Glowacka, I., S. Bertram, …, S. Pöhlmann. 2010. Differential downregulation of ACE2 by the spike proteins of severe acute respiratory syndrome coronavirus and human coronavirus NL63. *J. Virol.* 84:1198–1205.

31. Liu, Y., Y. Yang, …, L. Liu. 2020. Clinical and biochemical indexes from 2019-nCoV infected patients linked to viral loads and lung injury. *Sci. China Life Sci.* 63:364–374.

32. Li, F. 2016. Structure, function, and evolution of coronavirus spike proteins. *Annu. Rev. Virol.* 3:237–261.

33. Watanabe, Y., T. A. Bowden, …, M. Crispin. 2019. Exploitation of glycosylation in enveloped virus pathobiology. *Biochim. Biophys. Acta, Gen. Subj.* 1863:1480–1497.

34. Monteil, V., H. Kwon, …, J. M. Penninger. 2020. Inhibition of SARS-CoV-2 infections in engineered human tissues using clinical-grade soluble human ACE2. *Cell.* 181:905–913.e7.

35. Lei, C., K. Qian, …, S. Hu. 2020. Neutralization of SARS-CoV-2 spike pseudotyped virus by recombinant ACE2-Ig. *Nat. Commun.* 11:2070.

36. Guo, L., W. Bi, …, B. Dang. 2020. Engineered trimeric ACE2 binds and locks "three-up" spike protein to potently inhibit SARS-CoVs and mutants. *bioRxiv* https://doi.org/10.1101/2020.08.31.274704.

37. Chan, K. K., D. Dorosky, …, E. Procko. 2020. Engineering human ACE2 to optimize binding to the spike protein of SARS coronavirus 2. *Science.* 269:1261–1265.

38. Glasgow, A., J. Glasgow, …, J. A. Wells. 2020. Engineered ACE2 receptor traps potently neutralize SARS-CoV-2. *bioRxiv* https://doi.org/10.1101/2020.07.31.231746.

39. Spinello, A., A. Saltalamacchia, and A. Magistrato. 2020. Is the rigidity of SARS-CoV-2 spike receptor-binding motif the hallmark for its enhanced infectivity? Insights from all-atom simulations. *J. Phys. Chem. Lett.* 11:4785–4790.

40. Brielle, E. S., D. Schneidman-Duhovny, and M. Linial. 2020. The SARS-CoV-2 exerts a distinctive strategy for interacting with the ACE2 human receptor. *Viruses.* 12:1–10.

41. Wang, Y., M. Liu, and J. Gao. 2020. Enhanced receptor binding of SARS-CoV-2 through networks of hydrogen-bonding and hydrophobic interactions. *Proc. Natl. Acad. Sci. USA.* 117:13967–13974.

42. Lupala, C. S., X. Li, …, X. Su. 2020. Computational simulations reveal the binding dynamics between human ACE2 and the receptor binding domain of SARS-CoV-2 spike protein. *bioRxiv* https://doi.org/10.1101/2020.03.24.005561.

43. Guo, X., Z. Chen, …, H. Li. 2020. Investigation of the genetic variation in ACE2 on the structural recognition by the novel coronavirus (SARS-CoV-2). *J. Transl. Med.* 18:321.

44. Turoňová, B., M. Sikora, …, M. Beck. 2020. In situ structural analysis of SARS-CoV-2 spike reveals flexibility mediated by three hinges. *Science.* 370:203–208.

45. Xu, C., Y. Wang, …, Y. Cong. 2020. Conformational dynamics of SARS-CoV-2 trimeric spike glycoprotein in complex with receptor ACE2 revealed by cryo-EM. *bioRxiv* https://doi.org/10.1101/2020.06.30.177097.

46. Amaro, R. E., and A. J. Mulholland. 2020. Biomolecular simulations in the time of COVID19, and after. *Comput. Sci. Eng.* 22:30–36.

47. Sun, Z., K. Ren, …, L. Li. 2020. Mass spectrometry analysis of newly emerging coronavirus HCoV-19 spike S protein and human ACE2 reveals camouflaging glycans and unique post-translational modifications. *bioRxiv* https://doi.org/10.1101/2020.04.29.068098.

48. Zhao, P., J. L. Praissman, …, L. Wells. 2020. Virus-receptor interactions of glycosylated SARS-CoV-2 spike and human ACE2 receptor. *Cell Host Microbe.* 28:586–601.e6.

49. Shajahan, A., S. Archer-Hartmann, …, P. Azadi. 2020. Comprehensive characterization of N- and O- glycosylation of SARS- CoV-2 human receptor angiotensin converting enzyme 2. *bioRxiv* https://doi.org/10.1101/2020.05.01.071688.

50. Kleta, R., E. Romeo, …, A. Koizumi. 2004. Mutations in SLC6A19, encoding B0AT1, cause Hartnup disorder. *Nat. Genet.* 36:999–1002.

51. Zhang, Y. 2008. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics.* 9:40.

52. Roy, A., A. Kucukural, and Y. Zhang. 2010. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* 5:725–738.

53. Yang, J., A. Roy, and Y. Zhang. 2013. Protein-ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. *Bioinformatics.* 29:2588–2595.

54. Park, S. J., J. Lee, …, W. Im. 2019. CHARMM-GUI Glycan Modeler for modeling and simulation of carbohydrates and glycoconjugates. *Glycobiology.* 29:320–331.

55. Jo, S., K. C. Song, …, W. Im. 2011. Glycan Reader: automated sugar identification and simulation preparation for carbohydrates and glycoproteins. *J. Comput. Chem.* 32:3135–3141.

56. Jo, S., T. Kim, …, W. Im. 2008. CHARMM-GUI: a web-based graphical user interface for CHARMM. *J. Comput. Chem.* 29:1859–1865.

57. van Meer, G., D. R. Voelker, and G. W. Feigenson. 2008. Membrane lipids: where they are and how they behave. *Nat. Rev. Mol. Cell Biol.* 9:112–124.

58. Casares, D., P. V. Escribá, and C. A. Rosselló. 2019. Membrane lipid composition: effect on membrane and organelle structure, function and compartmentalization and therapeutic avenues. *Int. J. Mol. Sci.* 20:2167.

59. Huang, J., and A. D. MacKerell, Jr. 2013. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *J. Comput. Chem.* 34:2135–2145.

60. Jorgensen, W. L., J. Chandrasekhar, …, M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–932.

61. Phillips, J. C., R. Braun, …, K. Schulten. 2005. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* 26:1781–1802.

62. Darden, T. A., D. York, and L. Pedersen. 1993. Particle-mesh Ewald: an N.log(N) method for Ewald sums in large systems. *J. Chem. Phys.* 98:10089–10092.

63. Ryckaert, J. P., G. Ciccotti, and H. J. C. Berendsen. 1977. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* 23:327–341.

64. McGibbon, R. T., K. A. Beauchamp, …, V. S. Pande. 2015. MDTraj: a modern open library for the analysis of molecular dynamics trajectories. *Biophys. J.* 109:1528–1532.

65. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* 14:27–28, 33–38.

66. Pedregosa, F., V. Michel, …, É. Duchesnay. 2011. Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* 12:2825–2830.

67. Roe, D. R., and T. E. Cheatham, III. 2013. PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.* 9:3084–3095.

68. Mehdipour, A. R., and G. Hummer. 2020. Dual nature of human ACE2 glycosylation in binding to SARS-CoV-2 spike. *bioRxiv* https://doi.org/10.1101/2020.07.09.193680.

69. Casalino, L., Z. Gaieb, …, R. E. Amaro. 2020. Beyond shielding: the roles of glycans in the SARS-CoV-2 spike protein. *ACS Cent. Sci.* 6:1722–1734.

70. Lan, J., J. Ge, …, X. Wang. 2020. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature.* 581:215–220.

71. Maier, J. A., C. Martinez, …, C. Simmerling. 2015. Ff 14SB: improving the accuracy of protein side chain and backbone parameters from ff 99SB. *J. Chem. Theory Comput.* 11:1696–1713.

72. Kirschner, K. N., A. B. Yongye, …, R. J. Woods. 2008. GLYCAM06: a generalizable biomolecular force field. Carbohydrates. *J. Comput. Chem.* 29:622–655.

73. Izadi, S., and A. V. Onufriev. 2016. Accuracy limit of rigid 3-point water models. *J. Chem. Phys.* 145:074501.

74. Li, P., L. F. Song, and K. M. Merz, Jr. 2015. Systematic parameterization of monovalent ions employing the nonbonded model. *J. Chem. Theory Comput.* 11:1645–1657.

75. Case, D. A., K. Belfon, …, P. A. Kollman. 2020. AMBER 2020. University of California, San Francisco, San Francisco, CA.

76. Hopkins, C. W., S. Le Grand, …, A. E. Roitberg. 2015. Long-time-step molecular dynamics through hydrogen mass repartitioning. *J. Chem. Theory Comput.* 11:1864–1874.

77. Amaro, R. E., and A. J. Mulholland. 2020. A community letter regarding sharing biomolecular simulation data for COVID-19. *J. Chem. Inf. Model.* 60:2653–2656.

78. Benetti, E., R. Tita, …, A. M. Pinto; GEN-COVID Multicenter Study. 2020. ACE2 gene variants may underlie interindividual variability and susceptibility to COVID-19 in the Italian population. *Eur. J. Hum. Genet.* 28:1602–1614.

79. Calcagnile, M., P. Forgez, …, P. Alifano. 2020. ACE2 polymorphisms and individual susceptibility to SARS-CoV-2 infection: insights from an *in silico* study. *bioRxiv* https://doi.org/10.1101/2020.04.23.057042.

80. Li, W., C. Zhang, …, M. Farzan. 2005. Receptor and viral determinants of SARS-coronavirus adaptation to human ACE2. *EMBO J.* 24:1634–1643.

81. Procko, E. 2020. The sequence of human ACE2 is suboptimal for binding the S spike protein of SARS coronavirus 2. *bioRxiv* https://doi.org/10.1101/2020.03.16.994236.

82. Stawiski, E. W., D. Diwanji, …, S. Seshagiri. 2020. Human ACE2 receptor polymorphisms predict SARS-CoV-2 susceptibility. *bioRxiv* https://doi.org/10.1101/2020.04.07.024752.

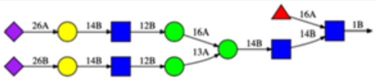
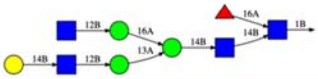
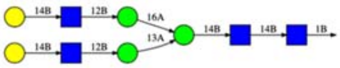
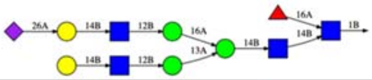
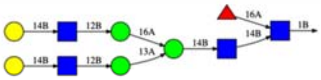
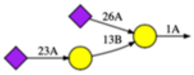
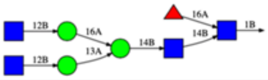
**Biophysical Journal, Volume *120***

## Supplemental information

# The flexibility of ACE2 in the context of SARS-CoV-2 infection

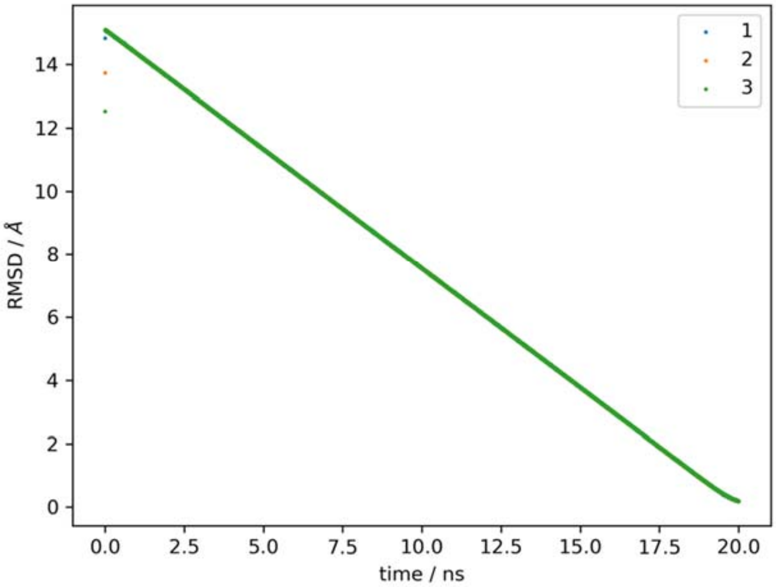
**Emilia P. Barros, Lorenzo Casalino, Zied Gaieb, Abigail C. Dommer, Yuzhang Wang, Lucy Fallon, Lauren Raguette, Kellon Belfon, Carlos Simmerling, and Rommie E. Amaro**

**Supplementary Table 1.** Glycan composition for ACE2 and RBD systems.

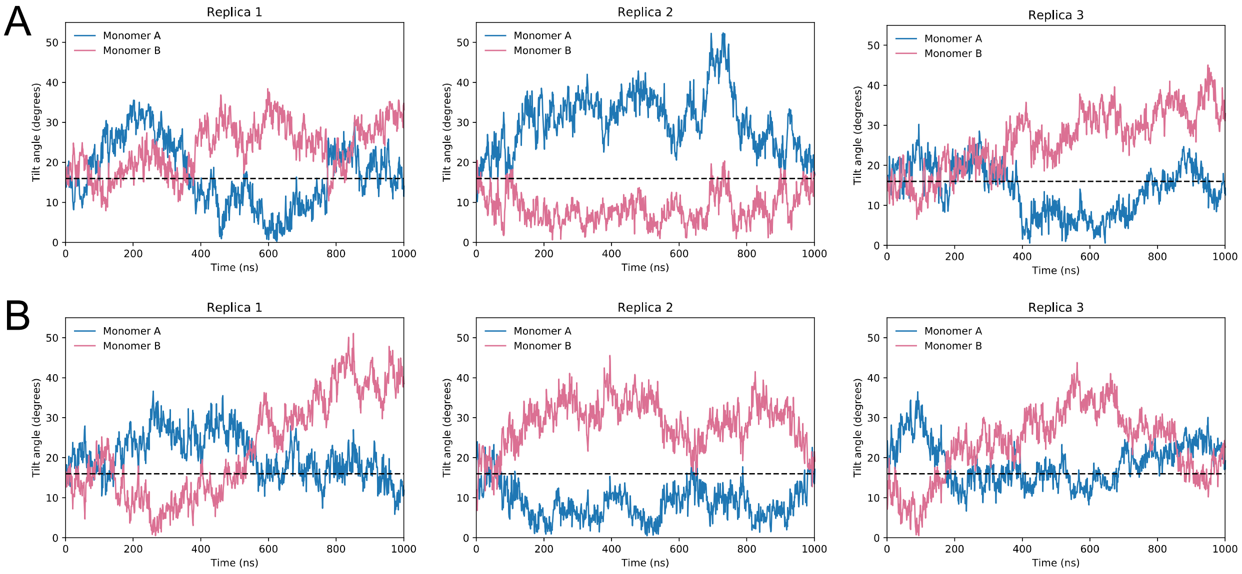
| Protein | Site | Structure |
|---|---|---|
| ACE2 | N53 |  |
| ACE2 | N90 |  |
| ACE2 | N103 |  |
| ACE2 | N322 |  |
| ACE2 | N432 |  |
| ACE2 | N546 |  |
| ACE2 | N690 |  |
| ACE2 | O730 |  |
| RBD | N343 |  |

**Supplementary Table 2.** Composition of ACE2-RBD interface regions.

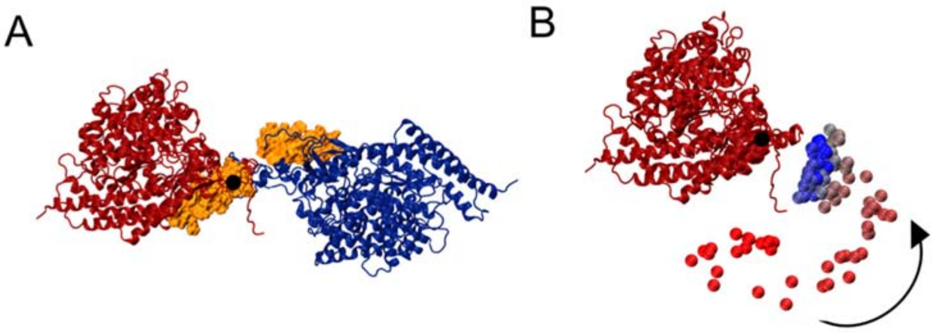
| Region | ACE2 residue | RBD residue |
|---|---|---|
| 1 | GLN24 | ALA475 |
| | GLN24 | GLY276 |
| | GLN24 | SER477 |
| | GLN24 | ASN487 |
| | THR27 | ALA475 |
| | THR27 | TYR489 |
| | PHE28 | TYR489 |
| | LYS31 | TYR489 |
| | MET82 | PHE486 |
| | TYR83 | PHE486 |
| | TYR83 | ASN487 |
| 2 | THR27 | PHE456 |
| | ASP30 | LEY455 |
| | ASP30 | PHE456 |
| | LYS31 | PHE456 |
| | LYS31 | GLN493 |
| | HIS34 | TYR453 |
| | HIS34 | LEU455 |
| | GLY35 | GLN493 |
| 3 | GLU37 | TYR505 |
| | ASP38 | TYR449 |
| | ASP38 | TYR495 |
| | ASP38 | GLY496 |
| | TYR41 | GLN498 |
| | TYR41 | THR500 |
| | TYR41 | ASN501 |
| | GLN42 | TYR449 |
| | GLN42 | GLN498 |
| | LEU45 | GLN498 |
| | ASN330 | THR500 |
| | LYS353 | GLY496 |
| | LYS353 | ASN501 |
| | LYS353 | GLY502 |
| | LYS353 | TYR505 |
| | GLY354 | GLY502 |
| | ASP355 | THR500 |
| | ARG357 | THR500 |
| | ARG393 | TYR505 |



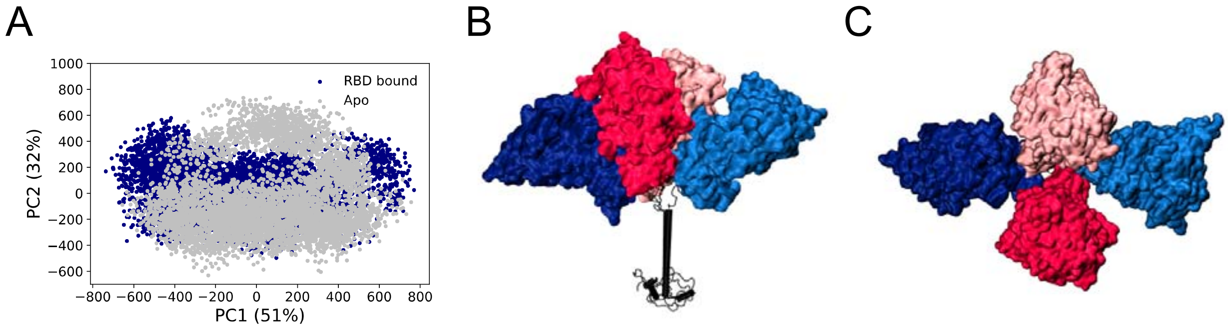
**Supplementary Figure 1.** RMSD *versus* time for the steered molecular dynamics of the RBD in each spike monomer, relative to the "RBD-up" in the reference structure. The three initial points correspond to the initial values in each monomer.



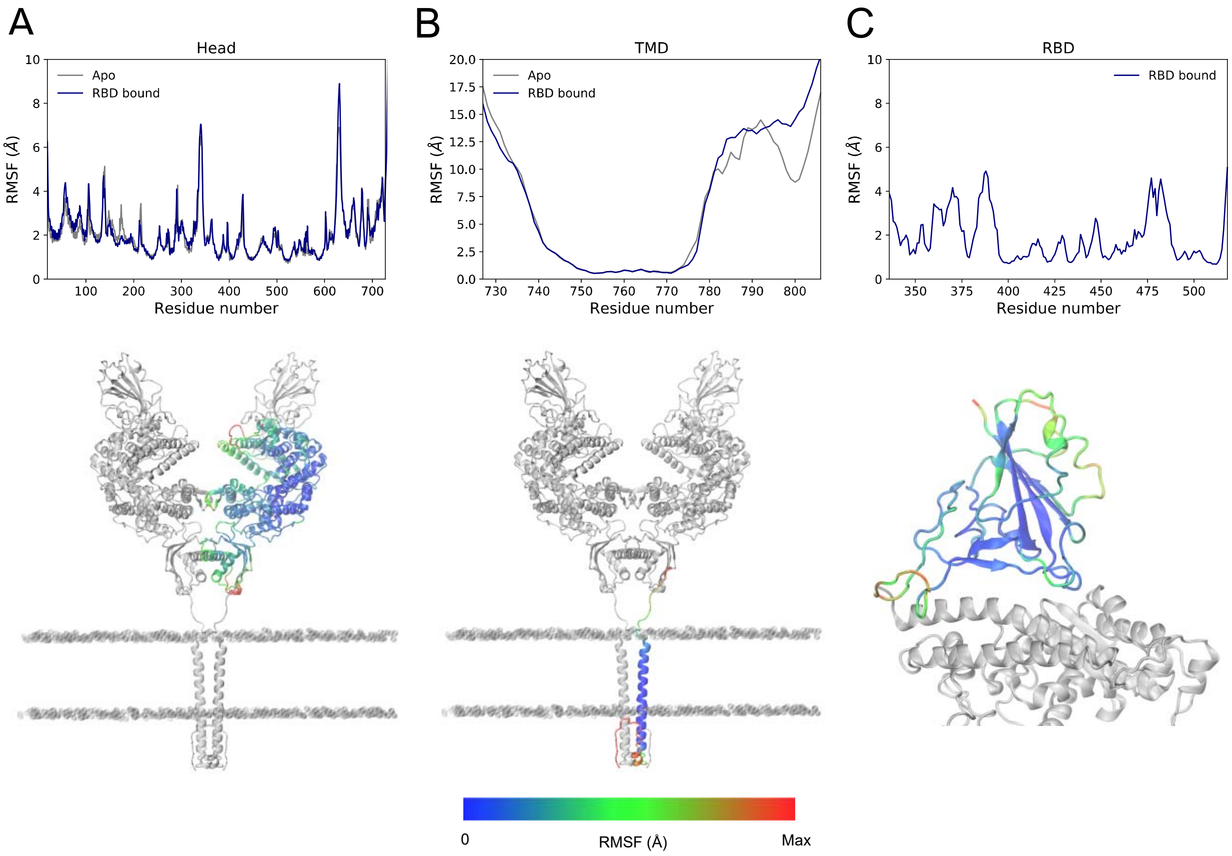
**Supplementary Figure 2.** Time evolution of the head tilt angle for each monomer in the **(a)** apo and **(b)** RBD-bound simulations. Tilt angle of the reference cryo-EM structure indicated by a black horizontal line.



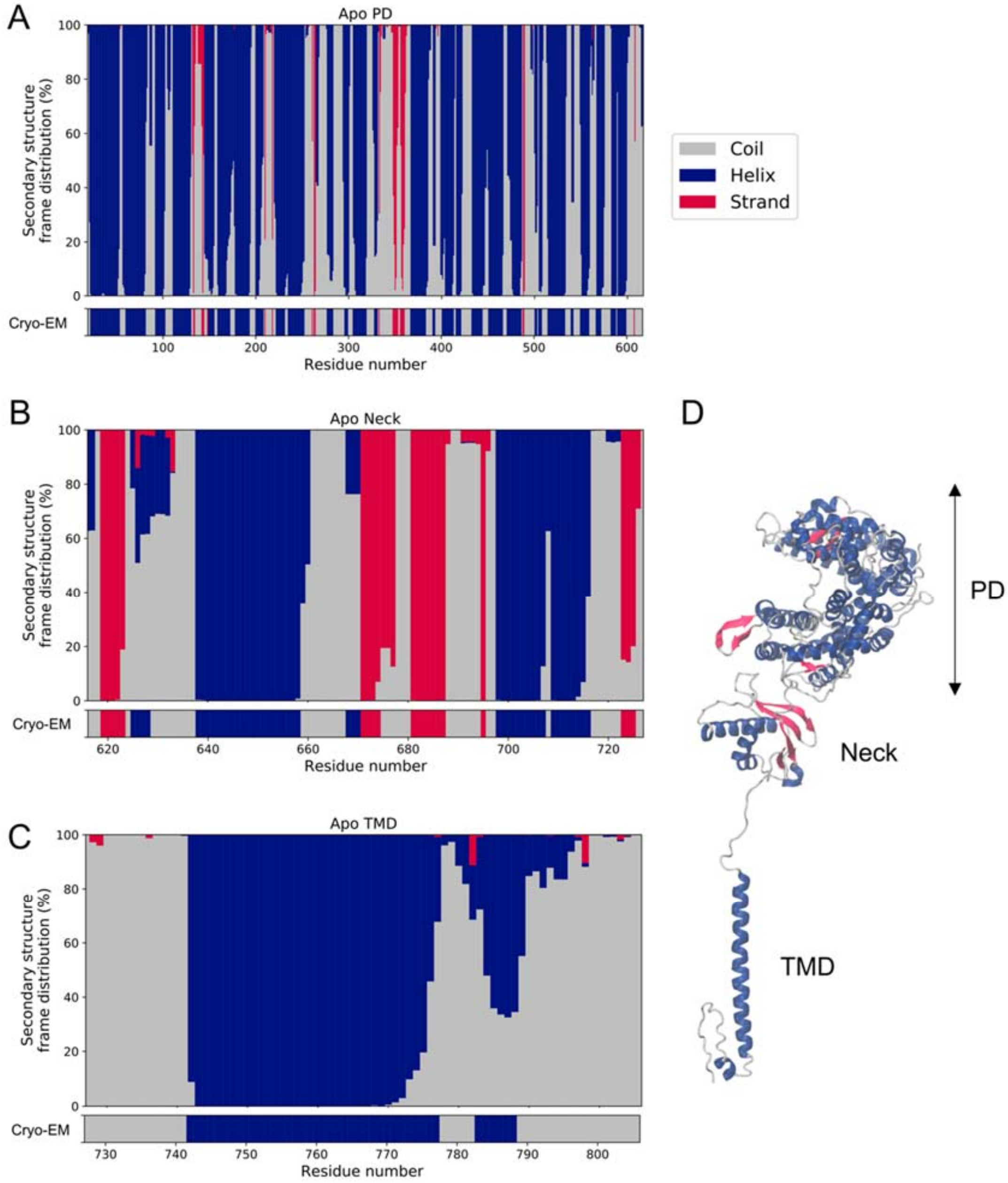
**Supplementary Figure 3. (a)** Representation of the change in position of the ACE2 homodimer binding interface (in orange surface representation) for the same MD initial and final monomer conformations shown in Figure 2a. **(b)** Displacement of the TM domain of the second monomer following motion of the monomer indicated in (a). Time evolution of the position of the Cα atom of residue 760 in the TM helix colored from dark red (t=0) to dark blue (t=1000 ns).



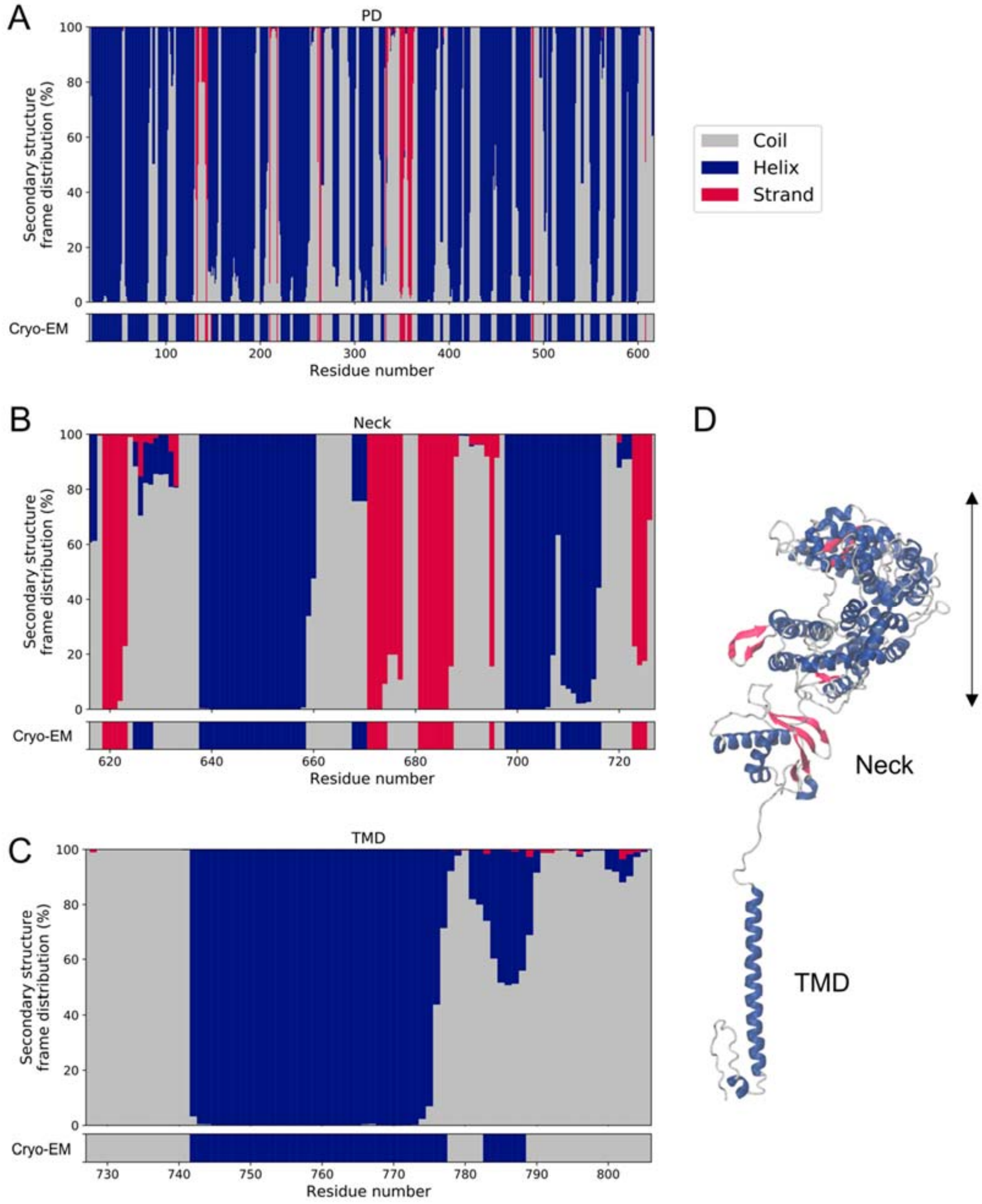
**Supplementary Figure 4. (a)** PCA plot showing PC1 and PC2 distribution of the apo (grey) and RBD-bound (navy) head coordinates in the simulations. **(b)** Side view and **(c)** top view of the conformations corresponding to the RBD-bound PC1 (in dark and light blue) and PC2 (in dark and light pink) extreme values. Head domains shown in surface representation, with the aligned transmembrane domains in black cartoon.



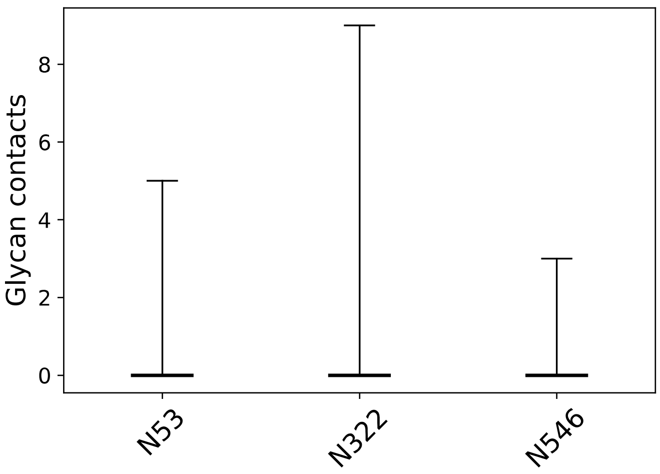
**Supplementary Figure 5.** Root-mean square fluctuations of backbone atoms for **(a)** ACE2's head, **(b)** ACE2's trans-membrane domain and **(c)** spike's RBD. Quantitative results from the accumulated apo (grey) and RBD-bound (navy) simulations shown at the graphs at the top, with the projection of the RMSF values on the protein structure, colored according to each domain's range of values, shown at the bottom (blue low RMSF, red high RMSF). The other regions of the complex structure are shown in grey for reference.



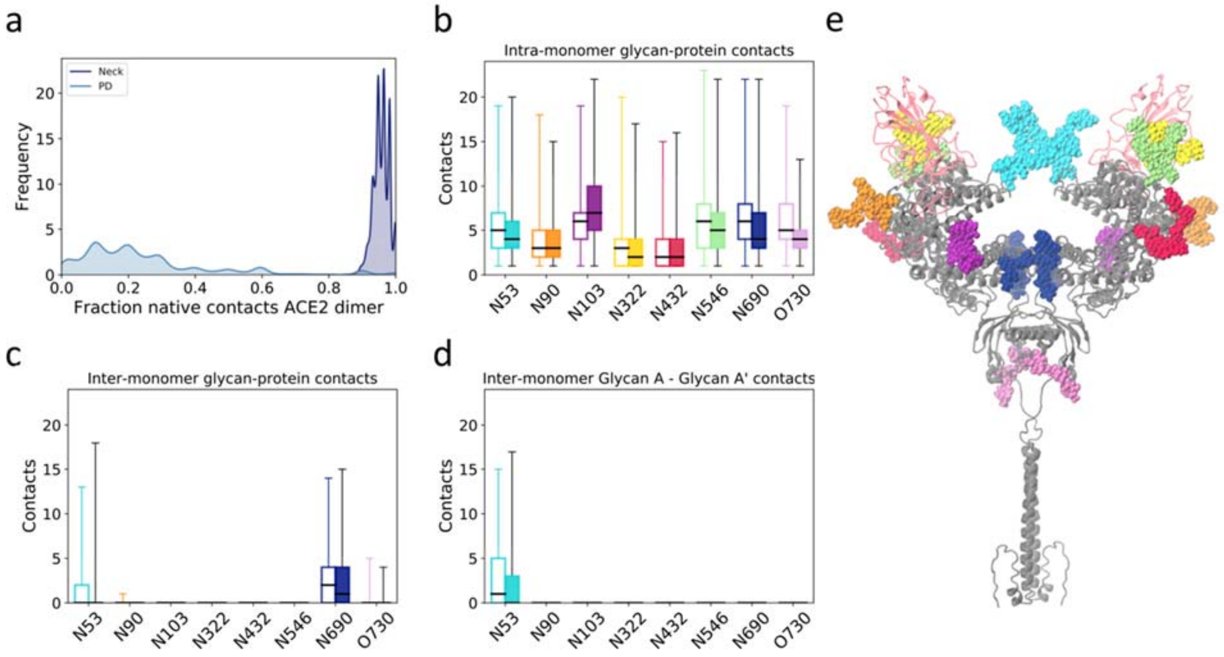
**Supplementary Figure 6.** Apo simulations secondary structure analysis for **(a)** peptidase (PD), **(b)** neck and **(c)** transmembrane (TMD) domains. The top, large graphs in each panel represent the frame distribution of the secondary structure assignments for each protein residue, colored according to the structure assignment: coil (grey), helix (navy) and strand (magenta). The bottom graph indicates the secondary structure assignment of the reference cryo-EM structure. **(d)** Visual representation of the secondary structure motifs in the initial cryo-EM structure, colored according to the same color scheme.



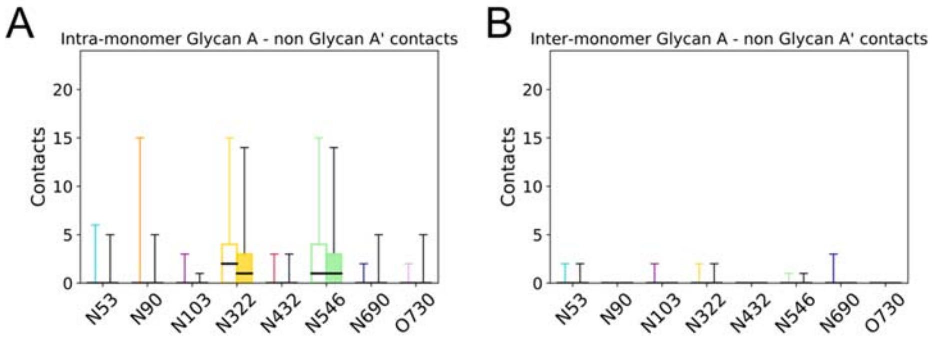
**Supplementary Figure 7.** RBD-bound simulations secondary structure analysis for **(a)** peptidase (PD), **(b)** neck and **(c)** transmembrane (TMD) domains. The top, large graphs in each panel represent the frame distribution of the secondary structure assignments for each protein residue, colored according to the structure assignment: coil (grey), helix (navy) and strand (magenta). The bottom graph indicates the secondary structure assignment of the reference cryo-EM structure. (d) Visual representation of the secondary structure motifs in the initial cryo-EM structure, colored according to the same color scheme.



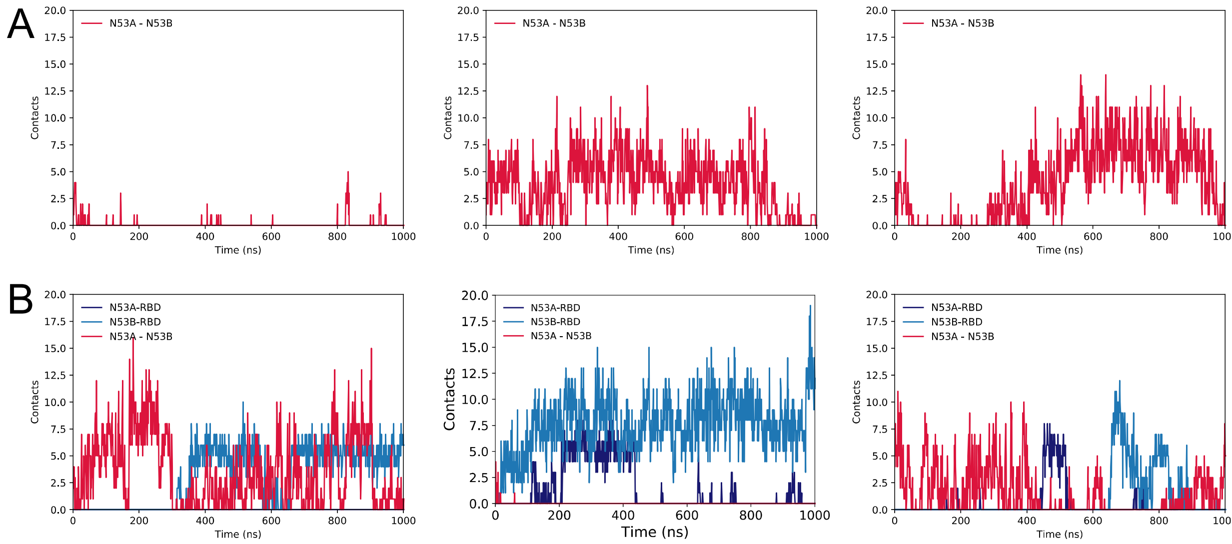
**Supplementary Figure 8.** N343-ACE2 glycan contacts. Horizontal black lines indicate mean value and whiskers show the total range of the data.



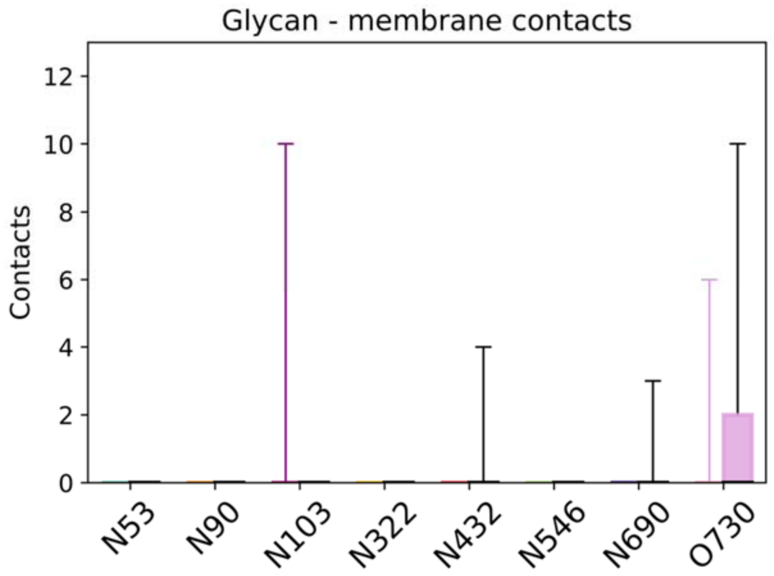
**Supplementary Figure 9.** Comparison of apo and RBD-bound glycan contacts. **(a)** Fraction of native contacts between ACE2 monomers for the apo simulations, considering only protein components of the glycoprotein. Neck and peptidase domain (PD) interacting regions are indicated separately. **(b)** Total glycan-protein interactions formed within each ACE2 monomer. Results from apo simulations shown in white boxes with colored contour and whiskers, RBD-bound simulations in colored boxes with black whiskers. Horizontal black lines indicate mean value, boxes extend to the lower and upper quartiles, and whiskers show the total range of the data. **(c)** Glycan-protein contacts between glycans in one of the monomer and protein residues from the opposite monomer. **(d)** Glycan-glycan contacts between glycan in one of the monomers and its copy in the opposite monomer (glycan A'). **(e)** ACE2 dimer with glycans in van der Waals representation colored according to figures b-d. ACE2 protein dimer colored grey and RBDs in light pink.



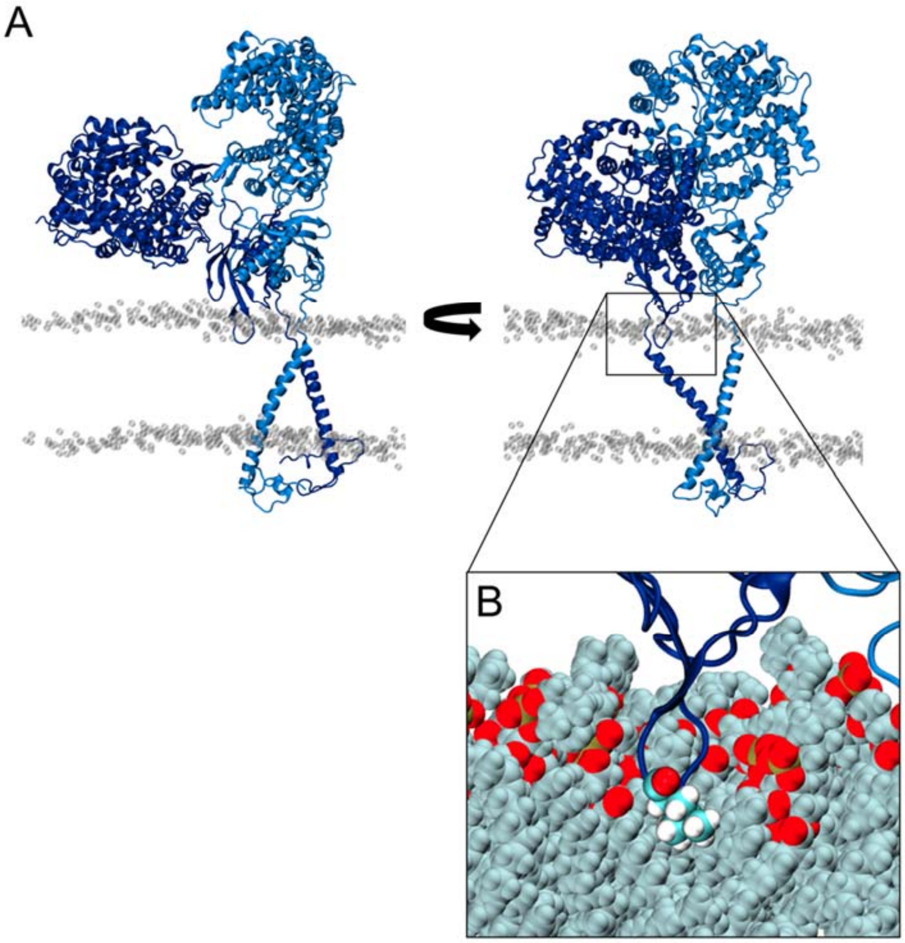
**Supplementary Figure 10.** ACE2 glycan-glycan contacts from apo and RBD-bound simulations. **(a)** Intra-monomer glycan-glycan contacts. **(b)** Inter-monomer glycan-glycan contacts, excluding contacts between equivalent glycans in the dimer. Results from apo simulations shown in white boxes with colored contour and whiskers, RBD-bound simulations in colored boxes with black whiskers. Horizontal black lines indicate mean value, boxes extend to the lower and upper quartiles, and whiskers show the total range of the data.



**Supplementary Figure 11.** Time evolution of N53 contacts. **(a)** N53A-N53B contacts in the homodimer during each replica of the apo ACE2 simulations. **(b)** N53-RBD contacts in each of the heterodimers as well as N53A-N53B contacts in the homodimer during each replica of the RBD-bound simulations.



**Supplementary Figure 12.** Glycan-membrane contacts for apo and RBD-bound simulations. Results from apo simulations shown in white boxes with colored contour and whiskers, RBD-bound simulations in colored boxes with black whiskers. Horizontal black lines indicate mean value, boxes extend to the lower and upper quartiles, and whiskers show the total range of the data.



**Supplementary Figure 13. (a)** Two views of a tilted ACE2 conformation, highlighting the insertion of the neck loop in the membrane. ACE2 monomers colored dark and light blue, and phosphorus atoms from membrane's lipid heads shown in grey, transparent van der Waals representation. (b) Detail of the inserted loop, showing Leu628 (carbon atoms in cyan, hydrogen atoms in white and oxygen atoms in red), and membrane lipids (oxygens in red, phosphorous atoms in ochre and tails in light cyan) in van der Waals representation.