

# Serine 477 plays a crucial role in the interaction of the SARS-CoV-2 spike protein with the human receptor ACE2.

**Amit Singh**

Institute of Molecular Bioscience, University of Graz, 8010 Graz, Austria

**Georg Steinkellner**

Innophore GmbH, 8010 Graz, Austria

**Katharina Köchl**

Innophore GmbH, 8010 Graz, Austria

**Karl Gruber** (✉ [karl.gruber@uni-graz.at](mailto:karl.gruber@uni-graz.at))

Institute of Molecular Bioscience, University of Graz, 8010 Graz, Austria

**Christian C. Gruber** (✉ [christian.gruber@innophore.com](mailto:christian.gruber@innophore.com))

Innophore GmbH, 8010 Graz, Austria

---

## Research Article

**Keywords:** Covid-19, SARS-CoV-2, RBD, ACE2, SARS-CoV-2 mutations, Spike

**Posted Date:** November 13th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-106969/v2>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

Serine 477 plays a crucial role in the interaction of the SARS-CoV-2 spike protein with the human receptor ACE2.

Amit Singh<sup>1</sup>, Georg Steinkellner<sup>1,2</sup>, Katharina Köchl<sup>2</sup>, Karl Gruber<sup>1,3\*</sup>, Christian C. Gruber<sup>1,2\*</sup>

<sup>1</sup>Institute of Molecular Bioscience, University of Graz, 8010 Graz, Austria

<sup>2</sup>Innophore GmbH, 8010 Graz, Austria

<sup>3</sup>Field of Excellence BioHealth – University of Graz, 8010 Graz, Austria

\*Correspondence should be addressed to (email: [karl.gruber@uni-graz.at](mailto:karl.gruber@uni-graz.at), [christian.gruber@innophore.com](mailto:christian.gruber@innophore.com))

**DISCLAIMER:** This is a PDF file of a paper that has been submitted for peer-review. Currently, most scientific journals require immediate publication of manuscripts before entering the peer-review process. We fully support this mechanism as a tool to distribute Covid-19 relevant information prior to final acceptance. Although unreviewed, the content has been subjected to preliminary formatting and quality control. We are providing this early version of the manuscript as a service to our authors and readers. The text and figures will undergo copyediting and a proof review before the paper is published in its final form after acceptance. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers apply. THE INFORMATION IS PROVIDED “AS IS”, NOT PEER-REVIEWED, WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE INFORMATION.

## **Abstract**

Since the worldwide outbreak of the infectious disease COVID-19, several studies have been published to understand the structural mechanism of the novel coronavirus SARS-CoV-2. During the infection process, the SARS-CoV-2 spike (S) protein plays a crucial role in the receptor recognition and cell membrane fusion process by interacting with the human angiotensin-converting enzyme 2 (hACE2) receptor. However, new variants of these spike proteins emerge as the virus passes through the host reservoir. This poses a major challenge for the design of a potent antibody against these spike proteins. Through structural bioinformatics analysis based on normal mode analysis (NMA) we identified a highly flexible region in the receptor binding domain (RBD) of SARS-CoV-2, starting from residue 475 up to residue 485. Structurally, S477 shows the highest rotational degrees of freedom among them. At the same time, S477 is the world's most frequently mutated residue in the RBD of SARS-CoV-2. Therefore, using advanced MD simulations, we have investigated the role of S477 and its two frequent mutations (S477G and S477N) at the RBD during the binding to hACE2. We found that the amino acid exchanges S477G and S477N strengthen the binding of the SARS-COV-2 spike with the hACE2 receptor.

## **Introduction**

In the recent global outbreak of COVID-19, more than 46 million individuals have as yet been infected with this severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). SARS-CoV-2 is responsible for around 1.25 million deaths as of October 2020. This novel virus is a member of structurally "crowned" viruses, first appreciated and defined as coronavirus in the 1960's when Tyrrell and Bynoe discovered the first human coronaviruses (HCoVs)<sup>1</sup>. These single-stranded RNA viruses have round enveloped virions and are covered by trimeric aggregates of spike (S) proteins. The S protein is a type I fusion protein and is involved in receptor binding, eventually

leading to virus fusion with the host cell. Specifically, the SARS-CoV-2 spike receptor binding domain (RBD) is known to interact with the human angiotensin-converting enzyme 2 (hACE2) receptor<sup>2,3</sup>. A conformational change in the trimeric S protein exposes one of its RBD domains in an active "up" state, and triggers the binding of the receptor binding motif (RBM) to hACE2<sup>4,5</sup>. Subsequently, the S protein is also recognized by the immune cells as a primary antigenic target to neutralize the virus. However, favourable mutations for the virus could be acquired through natural selection, following this as of now hundreds of new mutations were found on different residue positions of SARS-CoV-2 S protein. While preparing potential targets of the SARS-CoV-2 proteome for our recently published large-scale virtual screening<sup>6,7</sup> aiming to identify potential binders and inhibitors, we analysed the diversity of the SARS-CoV-2 genome landscape: Till 30<sup>th</sup> August 2020, among 73042 reported SARS-CoV-2 spike sequences on GISAID we have found that a total of 185 amino acid substitution are reported on the 223 residue long RBD and 68 mutations are reported in the 72 residues long RBM (Fig. 1a)<sup>8</sup>. Within the SARS-CoV-2 spike protein, residue D614 is a highly variable site, whereas residue S477 has the highest number of mutation events within the RBD (Fig.1b). Interestingly, S477N occurs very frequently alongside the D614G mutant which has been reported to be associated with a very high viral load<sup>9</sup>, structurally these two sites are located distinctly in the spike protein, but the very high frequency of co-occurrence among these two sites needs to be investigated.

Whereas, the rate of mutation in the RBD is posing a greater threat since the interaction of this domain to hACE2 is the key to enter a cell, recent studies are highlighting that several mutations in the RBD strengthened the SARS-CoV-2 infectivity<sup>10-12</sup>. Each mutation is said to have an impact on the protein-protein interaction interface in the heterodimeric complex of hACE2 and the RBD.

However, the site in the RBD which looks very promising to investigate is the residue S477. S477 has the highest frequency of mutation in the RBD (Fig. 1b) with reported substitutions to G(2) I(7) N(3400) R(2) T(2) X(20). Deep mutational scanning of the RBD has shown that a mutation at S477 has a potential to affect the stability of the RBD as well as its binding to the ACE2<sup>13</sup>, whereas the structural effect of residue S477 in context of intra-RBD interactions is lesser understood while it is reported that the residues S477 and T478 in SARS-CoV-2 are responsible for specific and efficient interactions with hACE2 and provide an edge over SARS-CoV<sup>14</sup>. Therefore, we investigate the structural contribution of residue S477 in the RBD in order to understand its role in hACE2 binding. Unbound proteins are predisposed to undergo conformational fluctuations relevant for protein-protein interactions and intra-protein interactions and have a profound role in presenting peripheral residues in inter-protein interactions<sup>15-17</sup>.

The structural analysis of the unbound RBD is therefore a very critical step for designing a potent inhibitor or an antibody to potentially block hACE2 and RBD interactions. In this study, we are investigating the intrinsic motion of the RBD in an unbound state using the normal mode analysis (NMA) technique. NMA is a time dependent harmonic approximation of internal motion in proteins around an energy minimum<sup>18,19,25</sup> and proved to be a very handy tool to study the harmonic motions relating to protein dynamics and conformational fluctuations<sup>18-20</sup>. NMA have been used previously to resolve the domain motion and dynamical correlation among protein residues<sup>20,21</sup>.

## **Results**

### **Flexibility analysis highlighted a highly correlated domain motion in the RBD**

The RBD is very efficient in finding its attachment surface on hACE2 and is therefore SARS-CoV-2's first line tool to explore its potential binding target on a human cell. It has been shown that the flexibility of peripheral residues plays a key role in the recognition of potential binding sites during protein-protein interactions<sup>22-24</sup>. Therefore, we performed a detailed flexibility analysis to identify the “flexible” and “rigid” regions of the SARS-CoV-2 RBD. We found a significant difference among the dynamics of peripheral residues on the RBD, especially in the receptor binding motif (RBM) of the SARS-CoV-2 RBD (Fig. 2). The key flexible region on the RBM covers the residues from 475 to 487, which might have a key role for an induced fit binding with hACE2. It has also been reported that the region, which corresponds to residues 475-483 in RBD, is a site of interest due to a higher frequency of mutations in this region<sup>9,13</sup>.

Normal mode flexibility analyses have shown an array of highly dynamic residues from residue 475 to 487 in the RBM. Surprisingly, these residues have a very high dynamical pair correlation among them with a correlation coefficient of  $\geq 0.9$ . Apart from RBM, RBD has no other highly correlated dynamical pair of residues in it.

However, we have noticed an interesting site among these highly flexible residues in the RBM, the residue S477, preceded by G476. A mutation corresponding to S477G results in two consecutive Gly residues at position 476 and 477, with the resulting arrangement A475-G476-S477G (AGG). This AGG motif might favour easier structural arrangements of side chains which are otherwise restricted<sup>26,27</sup>. Our detailed NMA suggested that there is a flexible domain of highly correlated dynamical residues on RBM, and local mutations S477G and S477N have a very high potential to affect local conformation significantly upon binding. Further using advanced MD simulations techniques, we are resolving a critical loop on RBM, which has a prominent role in binding with hACE2.

### **Volumetric map reveals interaction hotspot on hACE2**

To verify the impact of the mutation S477G in affecting the binding with hACE2, molecular dynamics simulations (MD) have been carried out to obtain volumetric maps to visualize the capture signature of inter atomic contacts. We performed 100 ns MDs with the native hACE2:RBD complex and the hACE2:RBD S477G and S477N mutant complexes. Based on the atomic coordinates from these 100 ns trajectories a volumetric map was created by selecting the hACE2 residues within 5 Å of the RBD. The resulting maps, as shown in (Fig. 3), reveal notable differences: There is a pronounced interaction hotspot on hACE2 opposite to residue S477 of the RBD, while the mutant S477G lacks this hotspot on hACE2. Apparently, this difference seems to be due to the mutation of a serine to a glycine, because glycine lacks any functional group in contrast to the hydroxyl group of the serine. However, the average number of hydrogen bonds per trajectory frame between the hACE2 and the RBD are 5.711, 5.982, 6.657 for native, S477G, S477N mutants, respectively.

### **S477 mutants are showing contrasting change in flexibility upon binding to hACE2**

Using molecular dynamics simulations, root mean square fluctuation (RMSF) is a broadly used tool to identify residue-specific-flexibility<sup>28,29</sup>. Plotting the RMSF of each residue in the RBD, in its unbound form as well as in complex with hACE2, we noticed an overall decrease in flexibility of the RBD upon binding with hACE2 compared to the unbound RBD (Fig. 4).

It is noticeable that S477G and S477N increase the local flexibility with the unbound RBD, but in a complex with hACE2, the S477 mutants have a significantly reduced local flexibility. This contrasting change in the local flexibility of S477 mutants upon binding with hACE2 suggesting an active participation of this residue at the interaction interface.

### **S477G delayed the detachment of RBD from hACE2**

Steered molecular dynamics (SMD) is an advanced version of MD simulations, widely used to study the binding and unbinding events among bio-molecules in structural biology<sup>30-33</sup>. Where the dissociation process is being accelerated using an external forced pulling via a harmonic force constant. SMD simulations have proved their worthiness in resolving the effect of single point mutation on protein-protein interaction interfaces<sup>34,35</sup>

To study the impact of the mutation S477G in the detachment and association process of the spike protein to hACE2, a steered molecular dynamics simulation was carried out. The forced dissociation of the RBD from hACE2 depends on the spring constant and tension exerted by the surroundings (Fig. 5). Interestingly, there is a significant difference in the rupture force between the RBD and its mutants S477G and S477N throughout different force constants. The chosen spring constant of 250 kJ/mol/nm<sup>2</sup> allows for a consistence and smooth transition between bound and unbound complex. A final center-of-mass (COM) distance between hACE2 and the RBD of approximately 14 nm was achieved.

Energy and conformational changes show that the rupture force varied greatly between the RBD and its S477 mutants. Van-der-Waals energies show lesser fluctuations in the event of forceful dissociation of S477 RBD mutants from hACE2, whereas the electrostatic forces remained dominant in the native RBD disassociation from hACE2. However, the complex of native RBD and hACE2 start dissociating after 400 ps (Fig. 6), S477N and hACE2 starts dissociating at 390 ps, whereas the S477G RBD mutant delayed the disassociation process with hACE2 for around 100 ps. To further confirm the time point of disassociation of the RBD:hACE2 complex, we plotted the distance between RBD and hACE2 vs. time, indicating that the S477G RBD mutant is significantly delaying the disengagement process (Fig. 7). Potential of mean force calculations (Fig. 8) show that the S477N mutant features the highest binding affinity with hACE2 compared

to its native RBD and S477G (Table 1.) Our results are in line with experimentally reported changes in the binding affinity of S477G and S477N mutants<sup>13</sup>.

<b>System</b>	<b><math>\Delta G_{\text{Bind}}</math> (kcal/mol)</b>	<b><math>\Delta\Delta G_{\text{Bind}}</math> (kcal/mol)</b>
<b>Native</b>	-16.8	-
<b>S477G</b>	-19.6	-2.8
<b>S477N</b>	-23.1	-6.3

**Table 1**

To visualize the changes in the dissociation process we took structural snapshots at the breaking point for both, the RBD and its S477 mutants. At the breaking point, residue S477 is still hydrogen bonded to Q24 of hACE2 in the native RBD, whereas in the analysed mutants no polar contacts at the interaction interface could be observed. However, the mutated residue S477G and S477N is involved in non-polar interactions with hACE2. This structural annotation at the breaking point indicates that the interaction with the highly flexible loop residues of the RBD will remain until the end of the unbinding event.

## **Discussion**

Our flexibility and correlational analyses reveal a highly dynamic loop in the SARS-CoV-2 spike protein RBM from residue 475 to 485. It is known from induced fit theory that the binding event induces a conformational change at the interaction site, and biased dynamics of these highly correlated residues at the RBM are emphasizing their relevance for the hACE2 binding. In fact loops are the most common structural features found at the protein-protein interaction interfaces<sup>36</sup>. Forced pulling dynamics reveal that this highly flexible loop remained interacting to hACE2 till the end of the rupturing event, suggesting that this flexible loop provides the required mobility to mediate the hACE2:RBD interaction. Following this, it is worth to consider that any alteration in the flexibility of this loop could have a strong impact in the binding of the RBD with hACE2. The

comparative analysis of the RBD and its S477 mutants highlights the potential of single amino acid substitution on the protein-protein interaction. Both single mutations of S477 have shown the increased binding affinity for hACE2. On the basis of our flexibility and SMD analysis, we are emphasizing that any mutation subject to a structural change in this region could potentially be critical for the RBD and hACE2 interaction.

## **Methodology**

### **Template Crystal Structure**

As a starting point the recently published crystal structure of the SARS-CoV-2 spike receptor-binding domain bound to the hACE2 receptor obtained from PDB (PDB-Code:6M0J) having a resolution of 2.47 Å<sup>37</sup> was used. Missing heavy atoms, residues and hydrogen were then modelled by using the web based CHARMM package<sup>38,39</sup>. The mutation S477G was modelled by using the standalone CHARMM package.

### **Simulation System Details**

#### *System 1 - Unbound RBD*

The coordinates of the RBD and its S477G mutant were extracted from the above mentioned fully prepared structure. Both structures were subjected to an energy minimization step followed by 100 ps of NVT. A 100 ns equilibration run was performed to relax the interface side-chain conformations constrained by the hACE2. This relaxed unbound RBD structure was used for the NMA on the DynaMut server<sup>40</sup>.

For the in house comparative flexibility analysis, the RBD and its S477 mutants were further undergone three independent 10 ns production runs.

#### *System 2 - Complex of hACE2:RBD*

The coordinates of the hACE2:RBD complex and the complex of hACE2 with S477-RBD mutants (hACE2:S477G-RBD, hACE2:S477N-RBD) were taken from the above mentioned template crystal structure. Both complexes were further subjected to an energy minimization step followed by 100 ps of NVT and an 100 ns equilibration run. These equilibrated structures were used as input for the SMD simulations.

For the in house flexibility analysis, equilibrated complexes of hACE2:RBD and hACE2:S477G-RBD, and hACE2:S477N-RBD were further subjected to three independent 10 ns production runs.

#### **Molecular dynamics Simulation Details**

Each system was placed in the center of a box with a minimum of 10 Å distance from the edge of the box, to avoid imaginary interactions, followed by solvation and system neutralization by adding ions. Potential energy terms for the protein and ions were derived from CHARMM36<sup>41</sup> all atom additive force field, and TIP3P<sup>42</sup> potential was used to represent water molecules.

All classical simulations were performed using the Gromacs 2018<sup>43,44</sup> package with enhanced GPU support. A step size of 2 fs was chosen and the driving force from potential energy of the system having a non-bonding cutoff of 10 Å, a Particle Mesh Ewald (PME)<sup>45</sup> summation was used to deal with long range interactions. All covalent bonds involving hydrogen were constrained using LINCS<sup>46</sup>. The temperature was maintained at 300 K with a Velocity-rescaling algorithm<sup>47</sup> using two indexed groups: protein and non-protein, for effective temperature coupling. Pressure

coupling was applied isotropically at 1 bar with a weak coupling constant 2 ps and isothermal compressibility values of  $4.5 \times 10^{-5} \text{ bar}^{-1}$ .

### **Steered MD Simulation Details**

For SMD, the structures were placed in a rectangular box with the dimensions (100 Åx100 Åx350Å) sufficient to avoid imaginary interactions and providing space for pulling simulations to take place along the Z-axis. As before, followed by solvation and neutralization by adding ions. Subsequently, bad contacts were then removed by energy minimization, followed by a 100 ps NVT and a 500 ps NPT equilibration. After equilibration, position restraints were removed from the RBD whereas hACE2 remained as an immobile reference for the pulling simulations. In each of the 10 different pulling simulations, the RBD was pulled away from the hACE2 along the Z-axis over 1000 ps with a fixed spring constant value, starting from 50 kJ/mol/nm for the first simulation to 500 kJ/mol/nm for the 10<sup>th</sup> simulation, with an increment of 50 kJ/mol/nm in each step between the first and last simulations. A pull rate of 0.1 Å/ps has been used throughout all of the 10 pulling simulations. All pulling simulations were performed with the GROMACS package.

To generate the starting configurations for umbrella sampling we have chosen trajectories resulting from the pulling simulations with the spring constant of 250 kJ/mol/nm<sup>2</sup>. Symmetrical distribution of sampling windows was used with a window spacing of 0.1 nm, up to a 14 nm center-of-mass (COM) separation. Such finer spacing resulted in 90 windows at different COM distances. In each window, 5 ns of MD was performed for a total simulation time of 450 ns utilized for umbrella sampling. Finally, analysis of the umbrella sampling was performed with the weighted histogram method (WHAM)<sup>48</sup> with 300 bootstrapping steps.

## Normal Mode Analysis (NMA)

Normal modes are represented by eigenvectors of the Hessian matrix and eigenvalues are the squares of the associated frequencies. Each eigenvector describes a state of the protein where all C $\alpha$  are oscillating with the same characteristic frequency.

Using characteristic frequency of each mode, atomic fluctuations has been calculated, given by:

$$\langle \Delta x_i^2 \rangle = \frac{k_B T}{m} + \sum_{j=1}^{3N-6} \frac{a_{ij}^2}{\omega_j^2}$$

where  $\langle \Delta x_i^2 \rangle$  is the time averaged square displacement of atom i,  $\omega_j$  is the frequency of mode j,  $a_{ij}$  is the displacement of atom i under mode j, and N is the number of residues.

The DynaMut online server takes advantage of the graph-theory and normal mode analysis (NMA) and has been employed to analyse and visualize the RBD dynamics, to generate a consensus prediction of protein motion and flexibility<sup>40</sup>.

## Root Mean Square Fluctuations (RMSF)

RMSF analysis described the average deviation of a residue over a time from a reference position.

Using RMSF the portions of the structure that are flexible or rigid could be identified. After removing the periodic boundary conditions using *Gromacs trjconv* tool, *gmx rmsf* module has been used to calculate RMSF using following equation:

$$RMSF_i = \sqrt{\frac{1}{T} \sum_{t_j=1}^T |r_i(t_j) - r_i^{ref}|^2}$$

where T is the total time of trajectory,  $t_j$  is the time at j<sup>th</sup> frame of trajectory,  $r_i$  is the position of the atom i and  $r_i^{ref}$  is the initial reference position of the atom i.

### Root Mean Square Deviations (RMSD)

The RMSD of a protein is taken as the root of the average square displacement over all N C $\alpha$  atoms, very often the measure of how much the protein conformation has changed. After removing the periodic boundary conditions using the *Gromacs trjconv* tool, the *gmx rms* module has been used to calculate the RMSD using following equation:

$$RMSD(t) = \sqrt{\frac{1}{M} \sum_{i=1}^N m_i |r_i(t) - r_i^{ref}|^2}$$

where M is the total mass of the protein, N is total amino acids,  $r_i(t)$  is the position of the atom i at time t and  $r_i^{ref}$  is the initial reference position of the atom i.

## References

1. Tyrrell, D. & Bynoe, M. Cultivation of a novel type of common-cold virus in organ cultures. *British medical journal* **1**, 1467 (1965).
2. Shang, J. *et al.* Structural basis of receptor recognition by SARS-CoV-2. *Nature* **581**, 221–224 (2020).
3. Shang, J. *et al.* Cell entry mechanisms of SARS-CoV-2. *Proceedings of the National Academy of Sciences* **117**, 11727–11734 (2020).
4. Wrapp, D. *et al.* Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**, 1260–1263 (2020).
5. Walls, A. C. *et al.* Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* (2020).
6. Gruber, C. C. & Steinkellner, G. Wuhan coronavirus 2019-nCoV – what we can find out on a structural bioinformatics level. 24044224 Bytes (2020)  
doi:10.6084/M9.FIGSHARE.11752749.V3.
7. Gorgulla, C. *et al.* A Multi-Pronged Approach Targeting SARS-CoV-2 Proteins Using Ultra-Large Virtual Screening. [https://chemrxiv.org/articles/A\\_Multi-Pronged\\_Approach\\_Targeting\\_SARS-CoV-2\\_Proteins\\_Using\\_Ultra-Large\\_Virtual\\_Screening/12682316](https://chemrxiv.org/articles/A_Multi-Pronged_Approach_Targeting_SARS-CoV-2_Proteins_Using_Ultra-Large_Virtual_Screening/12682316) (2020) doi:10.26434/chemrxiv.12682316.
8. Shu, Y. & McCauley, J. GISAID: Global initiative on sharing all influenza data—from vision to reality. *Eurosurveillance* **22**, 30494 (2017).
9. Korber, B. *et al.* Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. *Cell* **182**, 812-827.e19 (2020).

10. Chen, J., Wang, R., Wang, M. & Wei, G.-W. Mutations strengthened SARS-CoV-2 infectivity. *arXiv preprint arXiv:2005.14669* (2020).
11. Ozono, S. *et al.* Naturally mutated spike proteins of SARS-CoV-2 variants show differential levels of cell entry. *bioRxiv* (2020).
12. Ou, J. *et al.* Emergence of RBD mutations in circulating SARS-CoV-2 strains enhancing the structural stability and human ACE2 receptor affinity of the spike protein. *bioRxiv* (2020).
13. Starr, T. N. *et al.* Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. *Cell* S0092867420310035 (2020) doi:10.1016/j.cell.2020.08.012.
14. Wang, Q. *et al.* Receptor utilization of angiotensin-converting enzyme 2 (ACE2) indicates a narrower host range of SARS-CoV-2 than that of SARS-CoV. *Transbound Emerg Dis* tbed.13792 (2020) doi:10.1111/tbed.13792.
15. Li, X., Keskin, O., Ma, B., Nussinov, R. & Liang, J. Protein–protein interactions: hot spots and structurally conserved residues often locate in complemented pockets that pre-organized in the unbound states: implications for docking. *Journal of molecular biology* **344**, 781–795 (2004).
16. Tobi, D. & Bahar, I. Structural changes involved in protein binding correlate with intrinsic motions of proteins in the unbound state. *Proceedings of the National Academy of Sciences* **102**, 18908–18913 (2005).
17. Jayashree, S., Murugavel, P., Sowdhamini, R. & Srinivasan, N. Interface residues of transient protein-protein complexes have extensive intra-protein interactions apart from inter-protein interactions. *Biology direct* **14**, 1 (2019).

18. Levitt, M., Sander, C. & Stern, P. S. Protein normal-mode dynamics: trypsin inhibitor, crambin, ribonuclease and lysozyme. *Journal of molecular biology* **181**, 423–447 (1985).
19. Ben-Avraham, D. Vibrational normal-mode spectrum of globular proteins. *Physical Review B* **47**, 14559 (1993).
20. Case, D. A. Normal mode analysis of protein dynamics. *Current Opinion in Structural Biology* **4**, 285–290 (1994).
21. Hinsen, K. Analysis of domain motions by approximate normal mode calculations. *Proteins: Structure, Function, and Bioinformatics* **33**, 417–429 (1998).
22. Rajamani, D., Thiel, S., Vajda, S. & Camacho, C. J. Anchor residues in protein–protein interactions. *Proceedings of the National Academy of Sciences* **101**, 11287–11292 (2004).
23. Smith, G. R., Sternberg, M. J. & Bates, P. A. The relationship between the flexibility of proteins and their conformational states on forming protein–protein complexes with an application to protein–protein docking. *Journal of molecular biology* **347**, 1077–1101 (2005).
24. Jubb, H., Blundell, T. L. & Ascher, D. B. Flexibility and small pockets at protein–protein interfaces: new insights into druggability. *Progress in biophysics and molecular biology* **119**, 2–9 (2015).
25. Tasumi, M., Takeuchi, H., Ataka, S., Dwivedi, A. & Krimm, S. Normal vibrations of proteins: Glucagon. *Biopolymers: Original Research on Biomolecules* **21**, 711–714 (1982).
26. Vila, R., Ponte, I., Jiménez, M. A., Rico, M. & Suau, P. An inducible helix–Gly–Gly–helix motif in the N-terminal domain of histone H1e: A CD and NMR study. *Protein science* **11**, 214–220 (2002).

27. Ho, B. K. & Brasseur, R. The Ramachandran plots of glycine and pre-proline. *BMC structural biology* **5**, 1–11 (2005).
28. Shaw, D. E. *et al.* Atomic-level characterization of the structural dynamics of proteins. *Science* **330**, 341–346 (2010).
29. Maiorov, V. N. & Crippen, G. M. Significance of root-mean-square deviation in comparing three-dimensional structures of globular proteins. (1994).
30. Izrailev, S. *et al.* Steered molecular dynamics. in *Computational molecular dynamics: challenges, methods, ideas* 39–65 (Springer, 1999).
31. Isralewitz, B., Gao, M. & Schulten, K. Steered molecular dynamics and mechanical functions of proteins. *Current opinion in structural biology* **11**, 224–230 (2001).
32. Cuendet, M. A. & Michielin, O. Protein-protein interaction investigated by steered molecular dynamics: the TCR-pMHC complex. *Biophysical journal* **95**, 3575–3590 (2008).
33. Rodriguez, R. A., Yu, L. & Chen, L. Y. Computing protein–protein association affinity with hybrid steered molecular dynamics. *Journal of chemical theory and computation* **11**, 4427–4438 (2015).
34. Xiao, B.-L. *et al.* Steered molecular dynamic simulations of conformational lock of Cu, Zn-superoxide dismutase. *Scientific reports* **9**, 1–10 (2019).
35. Mino, G., Baez, M. & Gutierrez, G. Effect of mutation at the interface of Trp-repressor dimeric protein: a steered molecular dynamics simulation. *European Biophysics Journal* **42**, 683–690 (2013).
36. Gavenonis, J., Sheneman, B. A., Siegert, T. R., Eshelman, M. R. & Kritzer, J. A. Comprehensive analysis of loops at protein-protein interfaces for macrocycle design. *Nat Chem Biol* **10**, 716–722 (2014).

37. Lan, J. *et al.* Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* 1–6 (2020).
38. Jo, S., Kim, T., Iyer, V. G. & Im, W. CHARMM-GUI: a web-based graphical user interface for CHARMM. *Journal of computational chemistry* **29**, 1859–1865 (2008).
39. Brooks, B. R. *et al.* CHARMM: the biomolecular simulation program. *Journal of computational chemistry* **30**, 1545–1614 (2009).
40. Rodrigues, C. H., Pires, D. E. & Ascher, D. B. DynaMut: predicting the impact of mutations on protein conformation, flexibility and stability. *Nucleic acids research* **46**, W350–W355 (2018).
41. Huang, J. & MacKerell Jr, A. D. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *Journal of computational chemistry* **34**, 2135–2145 (2013).
42. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *The Journal of chemical physics* **79**, 926–935 (1983).
43. Abraham, M., Spoel, D., Lindahl, E., Hess, B. & Team, T. *GROMACS user manual version 2018*. (2018).
44. Van Der Spoel, D. *et al.* GROMACS: fast, flexible, and free. *Journal of computational chemistry* **26**, 1701–1718 (2005).
45. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An  $N \cdot \log(N)$  method for Ewald sums in large systems. *The Journal of chemical physics* **98**, 10089–10092 (1993).
46. Hess, B., Bekker, H., Berendsen, H. J. & Fraaije, J. G. LINCS: a linear constraint solver for molecular simulations. *Journal of computational chemistry* **18**, 1463–1472 (1997).

47. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *The Journal of chemical physics* **126**, 014101 (2007).
48. Li, Z., Fichthorn, K. A. & Milner, S. T. Surfactant binding to polymer–water interfaces in atomistic simulations. *Langmuir* **32**, 7519–7529 (2016).

## **Acknowledgements**

Financial support was provided by the Austrian Science Funds (FWF) through the doc.funds project DOC-46 "Catalog" and the Doctoral Academy Graz of the University of Graz.

## **Author Contributions**

A.S. designed and performed the simulations. A.S. wrote the manuscript. A.S. and K.K. performed the genome analysis. K.G. plotted the figures. G.S., K.G., K.K. and C.G. supervised the simulations and methods and reviewed the manuscript.

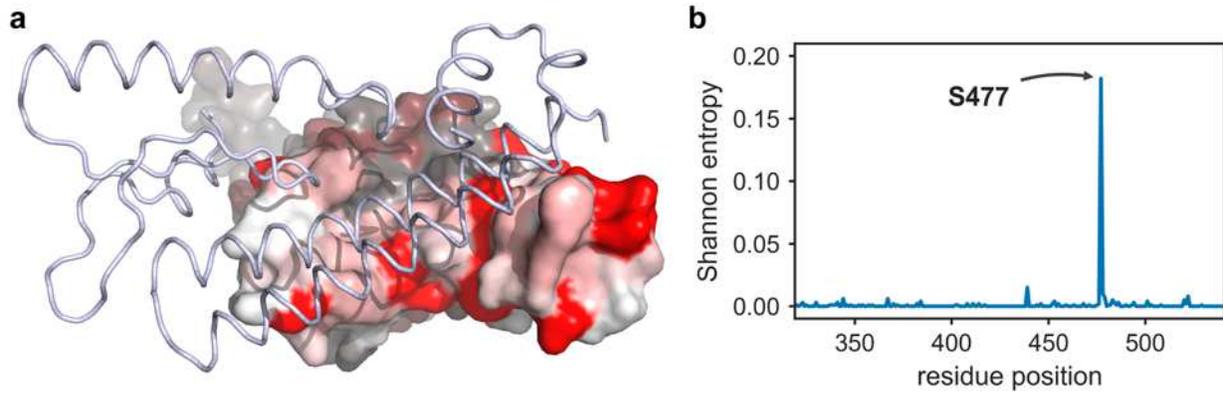
## **Additional Information**

### **Supplementary information**

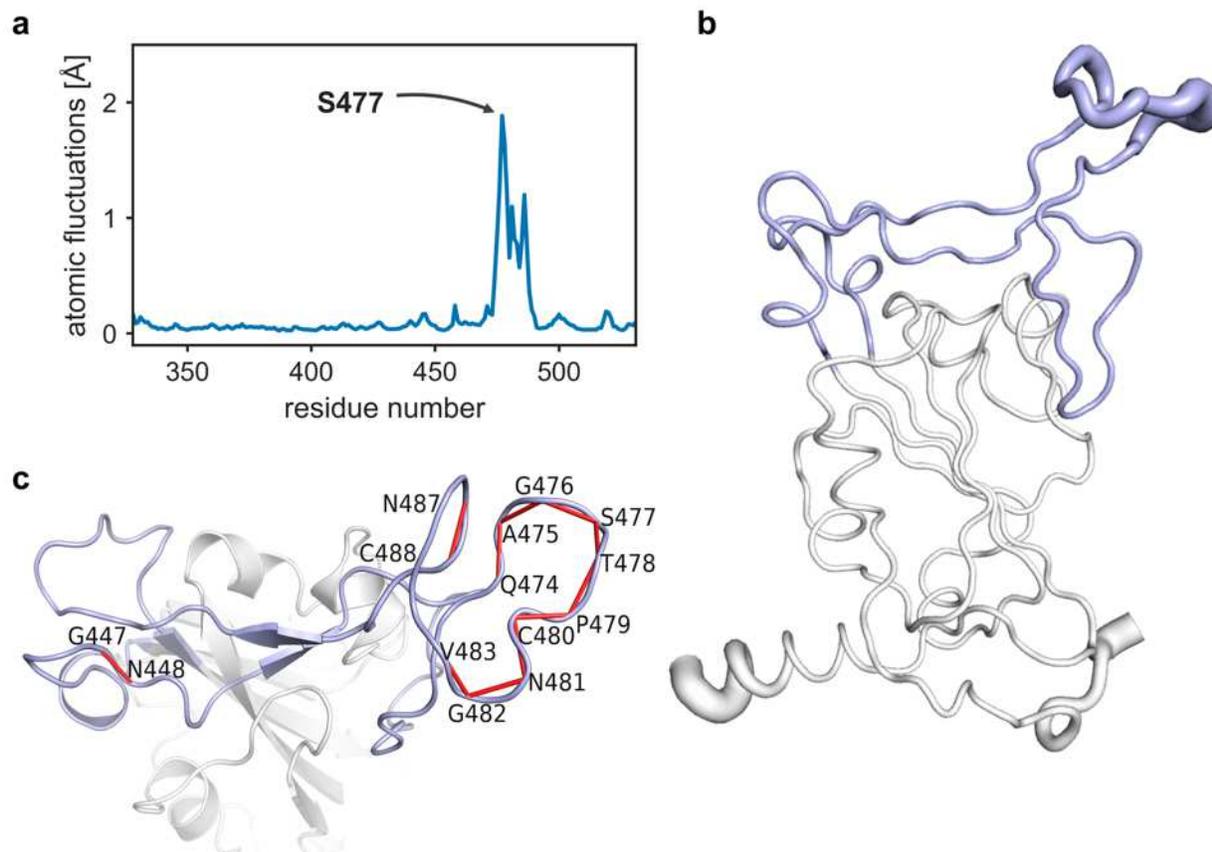
Supplementary movie 1: A movie showing the centre of mass (COM) separation between the hACE2 and RBD using steered MD simulation at a force constant of 250 kJ/mol/nm. hACE2 is shown in grey cartoon representation and is fixed while RBD showed in blue cartoon pulled away. Whereas, the RBD flexible loop residues [475:485] shown in orange colour.

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

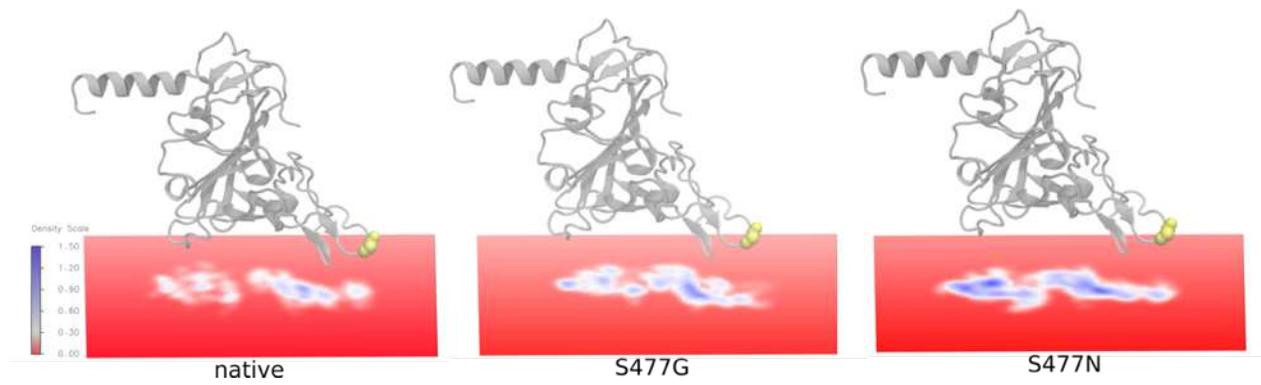
## Figure Captions



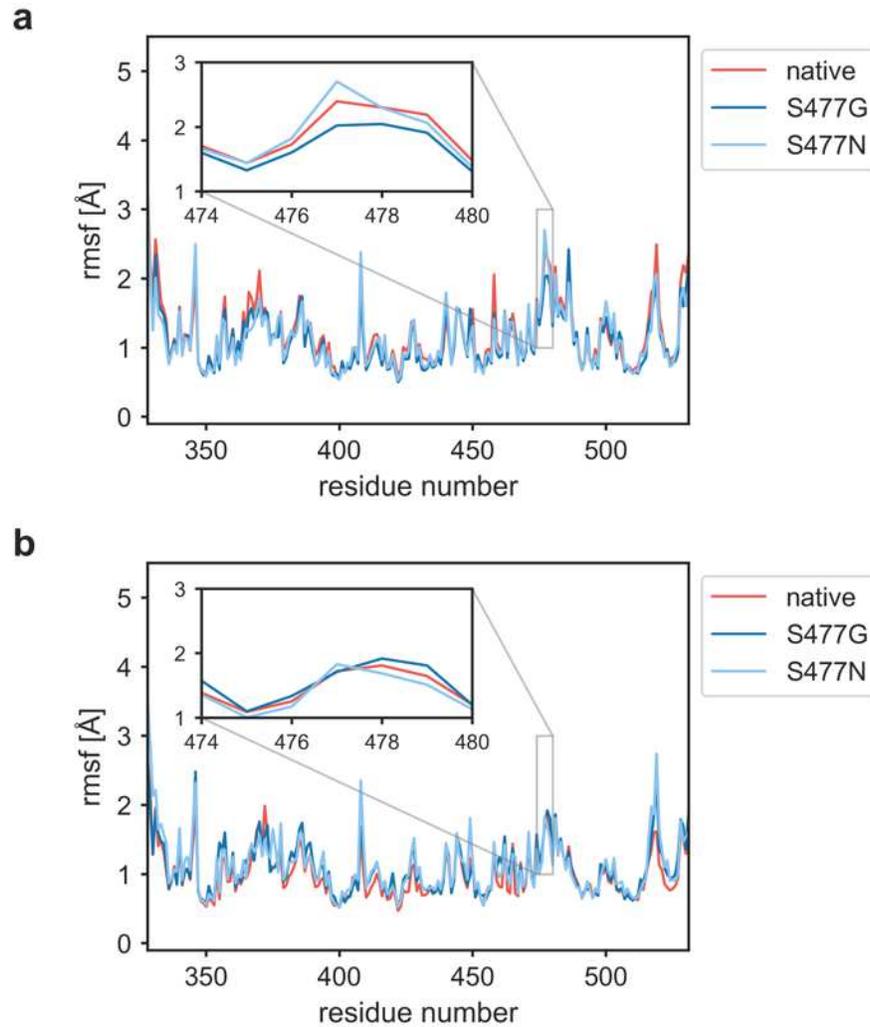
**Figure 1. Sequence analysis for reported mutations in RBD.** a) Structural annotation of the reported mutations via conservation score of the residue position. Intensities of red and white colour is representing variable and conserved region, respectively, in RBD represented as surface. Whereas, hACE2 shown in light blue ribbon. b) The reported mutations in RBD shown as position vs Shannon entropy scores. The Shannon entropy is given by  $H = -\sum_{i=1}^M P_i \ln_2 P_i$ . Where M is the total number of residue positions and  $P_i$  corresponds to probability of residue at position i. Higher the entropy value the lesser conserved a position is.



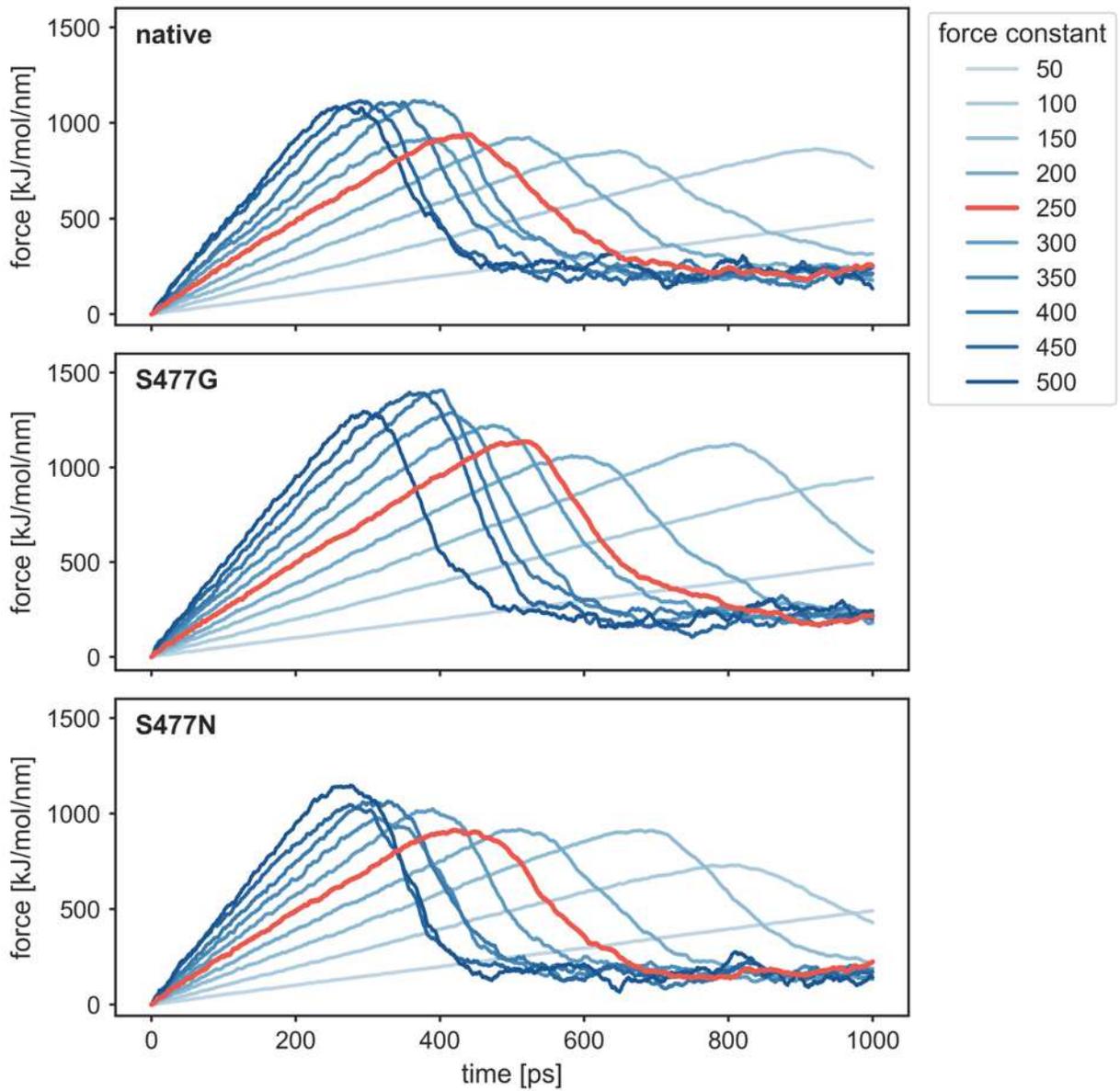
**Figure 2. DynaMut flexibility analysis** performed over the first 10 non-trivial modes of the molecule a) showing the residue wise atomic fluctuations in RBD, whereas the inset figure highlighting the residue wise atomic fluctuations in RBM b) Annotation of atomic fluctuations in RBD shown by the magnitude of the fluctuation represented by thin to thick tube, where blue colour corresponds to RBM residues c) Annotation of dynamical cross correlational among residues of RBD (dimmed in grey color), whereas RBM is shown in light blue cartoon. Only the highly correlated pairs of residues with a correlation coefficient of ( $\geq 0.9$ ) are annotated and connected by red lines.



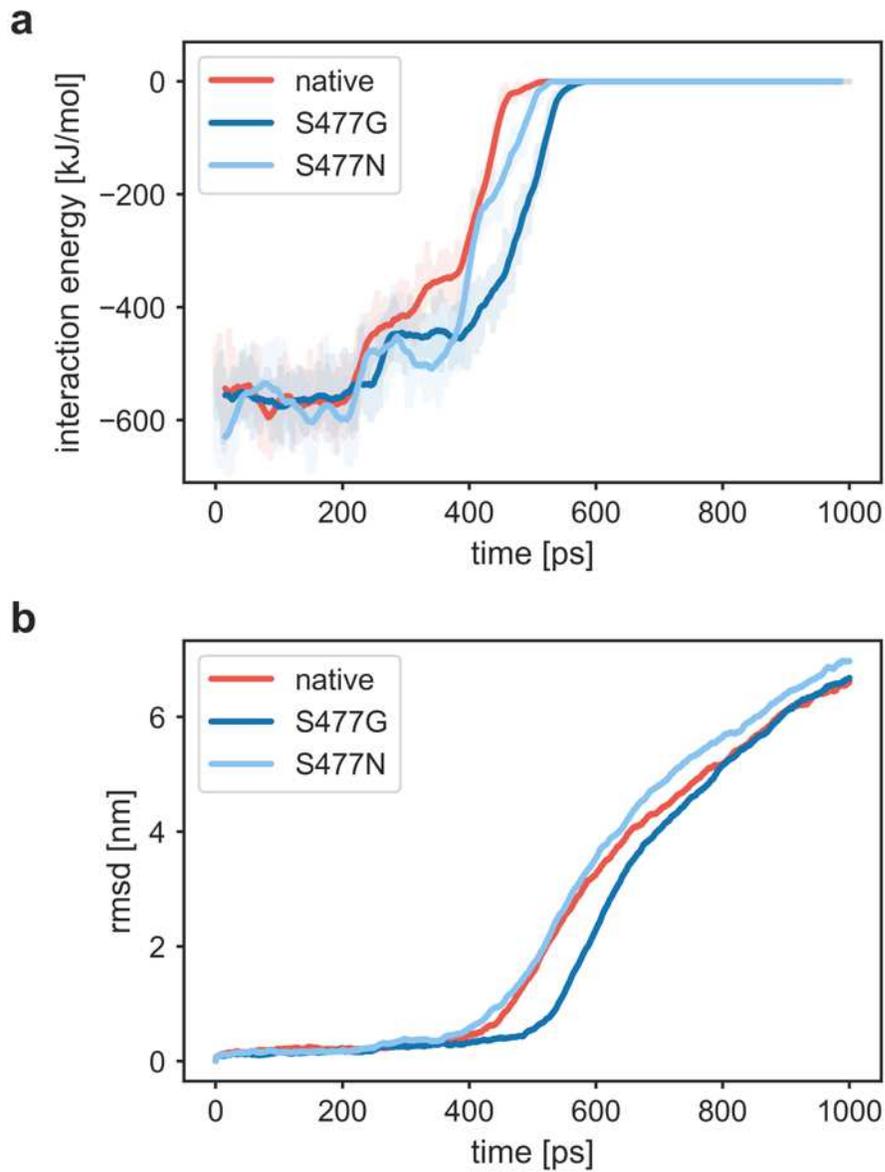
**Figure 3. Volumetric map analysis** of 100 ns trajectory showing the hACE2 residues within 5 Å of RBD, where RBD is shown in grey cartoon and residue 477 highlighted in yellow balls representation.



**Figure 4. Comparison of residue wise root mean square fluctuations**, each graph is averaged over three independent 10 ns simulations, standard error is shown in transparent thick boundaries along the lines. For clarity the first 10 residues of N and C terminal of RBD not shown. a) For unbound RBD and its S477 mutants b) RBD and its S477 mutants complexed with hACE2

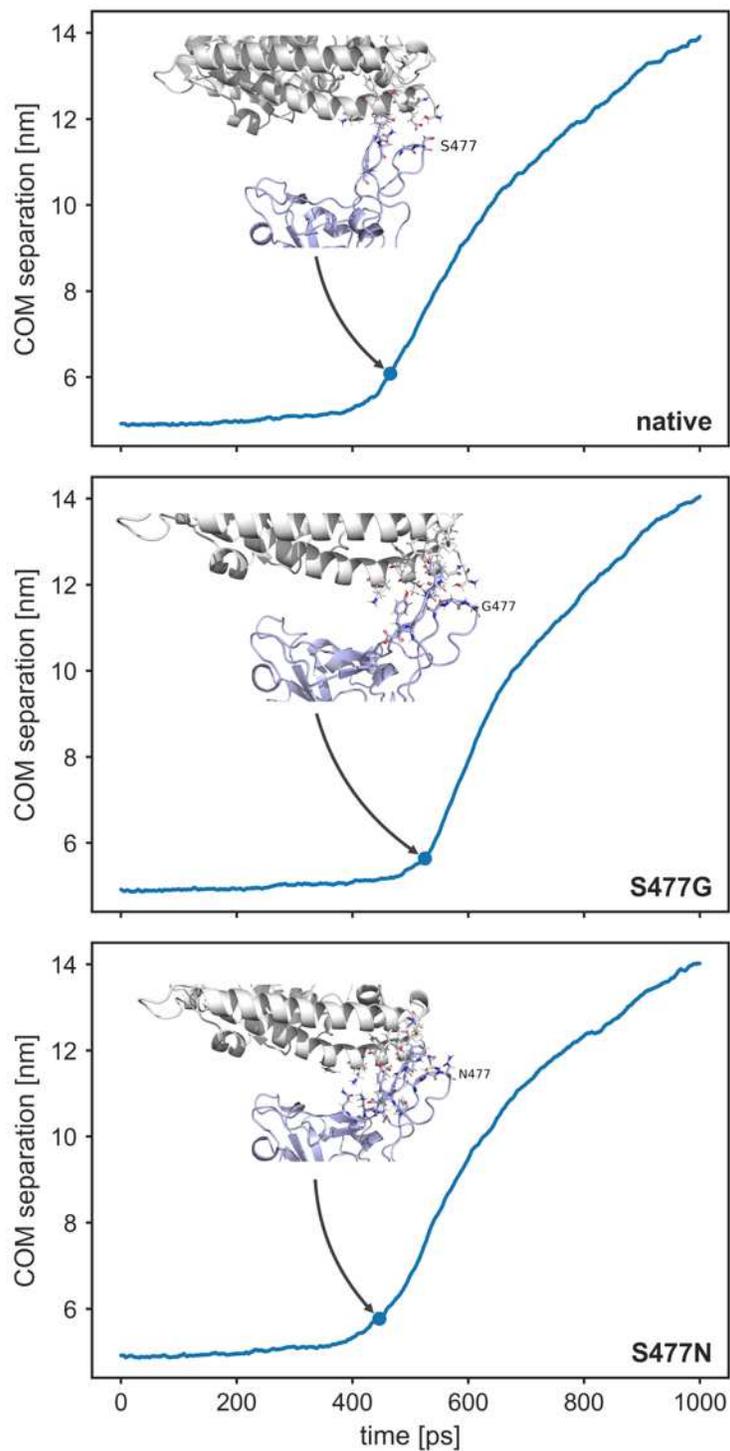


**Figure 5. Influence of the spring constant on the steering force of RBD: hACE2:Native RBD, hACE2:S477G and hACE2:S477N**

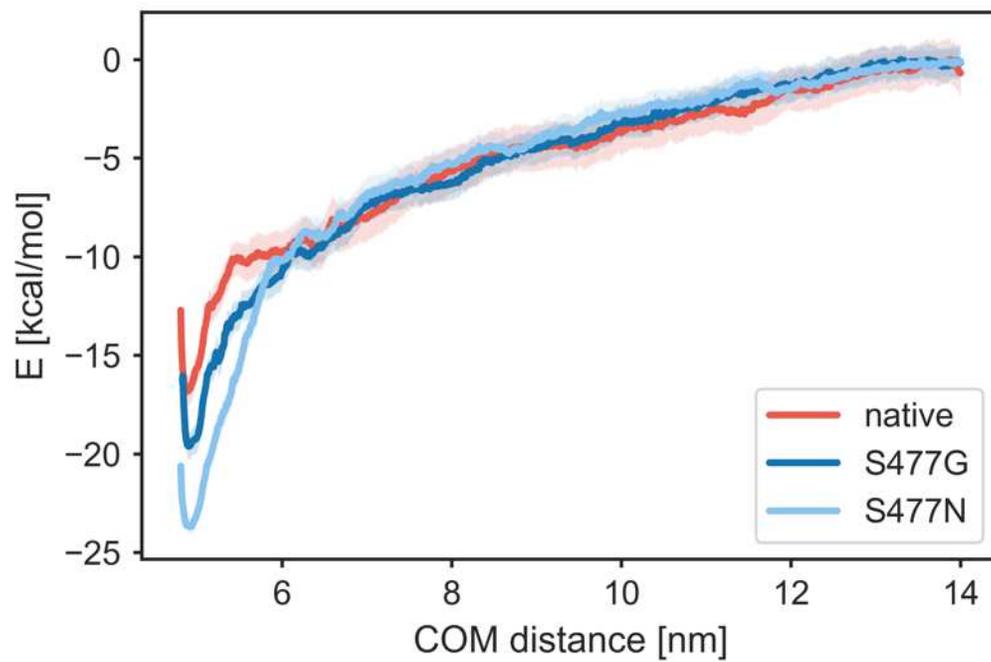


**Figure 6. Energy and conformational changes during SMD simulation at 250 kJ/mol/nm.**

Time dependence of the interaction energy between hACE2 chain and RBD chain of a) Native (b S477G c) S477N (d The comparison between Native RBD and its mutants e) Comparison of the RMSD of the carbon atoms in the protein backbone relative to the initial frame.

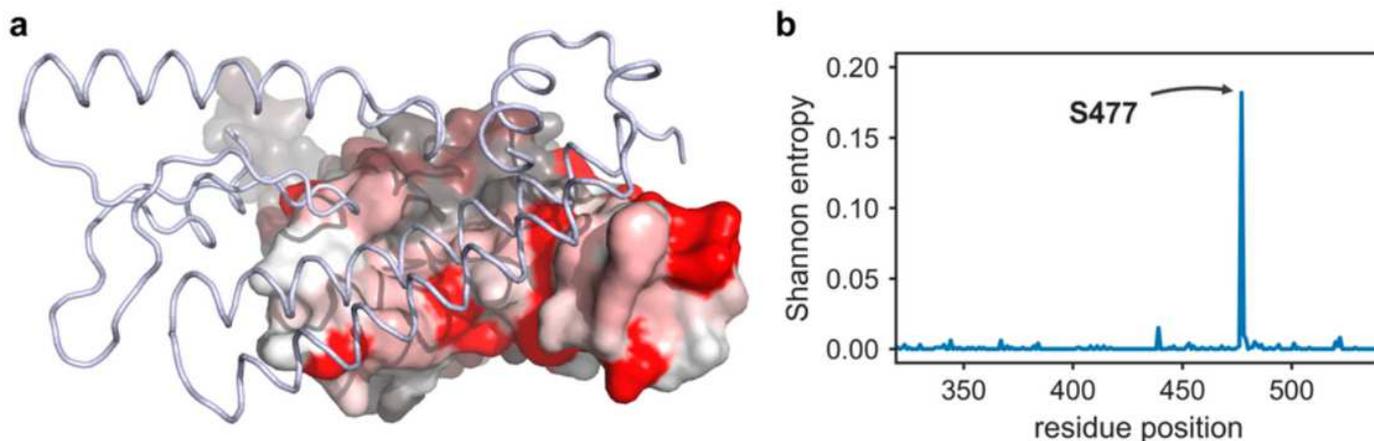


**Figure 7.** Figure showing the snapshot of hACE2:RBD interaction interface at the force rupturing event during SMD simulation. Where hACE2 shown in grey and RBD in light blue color. hACE2:Native RBD, hACE2:S477G and hACE2:S477N.



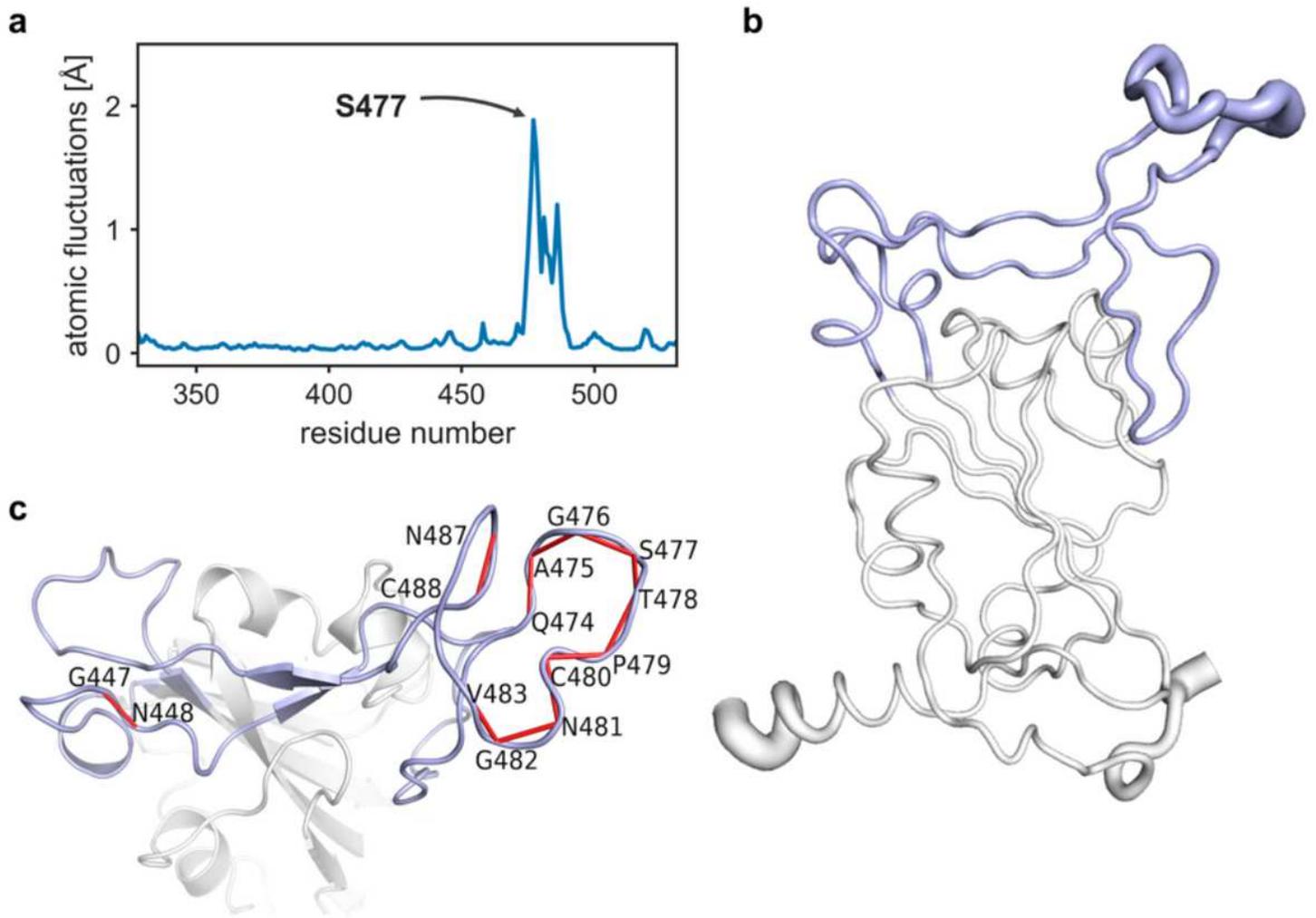
**Figure 8. Potential of mean force calculation from SMD simulation of RBD and its S477 mutants.** The uncertainty is shown by in the same coloured background on each plot.

# Figures



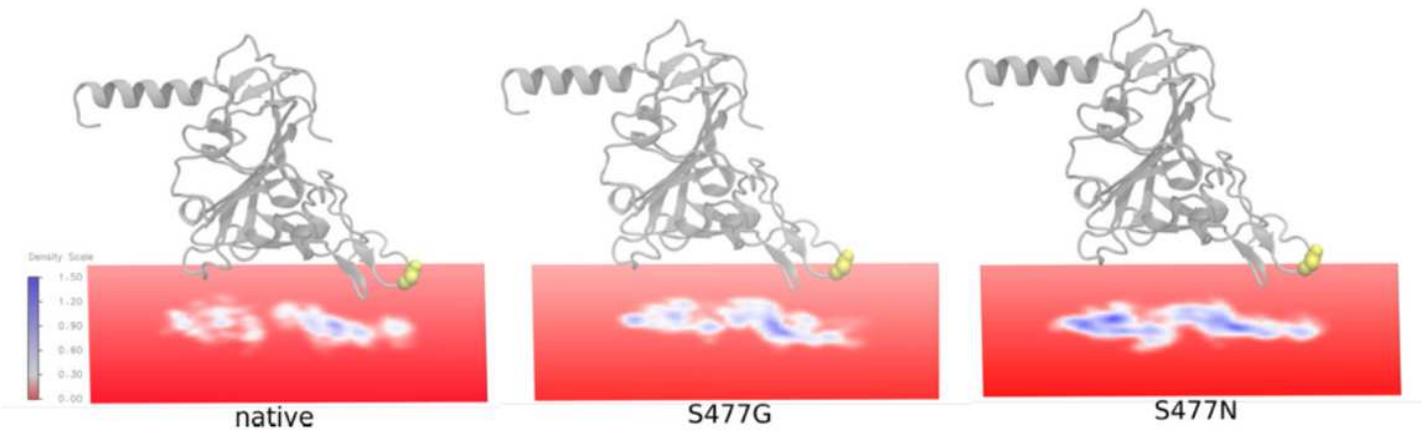
**Figure 1**

[See PDF file for accurate caption] Sequence analysis for reported mutations in RBD. a) Structural annotation of the reported mutations via conservation score of the residue position. Intensities of red and white colour is representing variable and conserved region, respectively, in RBD represented as surface. Whereas, hACE2 shown in light blue ribbon. b) The reported mutations in RBD shown as position vs Shannon entropy scores. The Shannon entropy is given by  $H = -\sum_{i=1}^M p_i \log_2 p_i$ . Where M is the total number of residue positions and  $p_i$  corresponds to probability of residue at position i. Higher the entropy value the lesser conserved a position is.



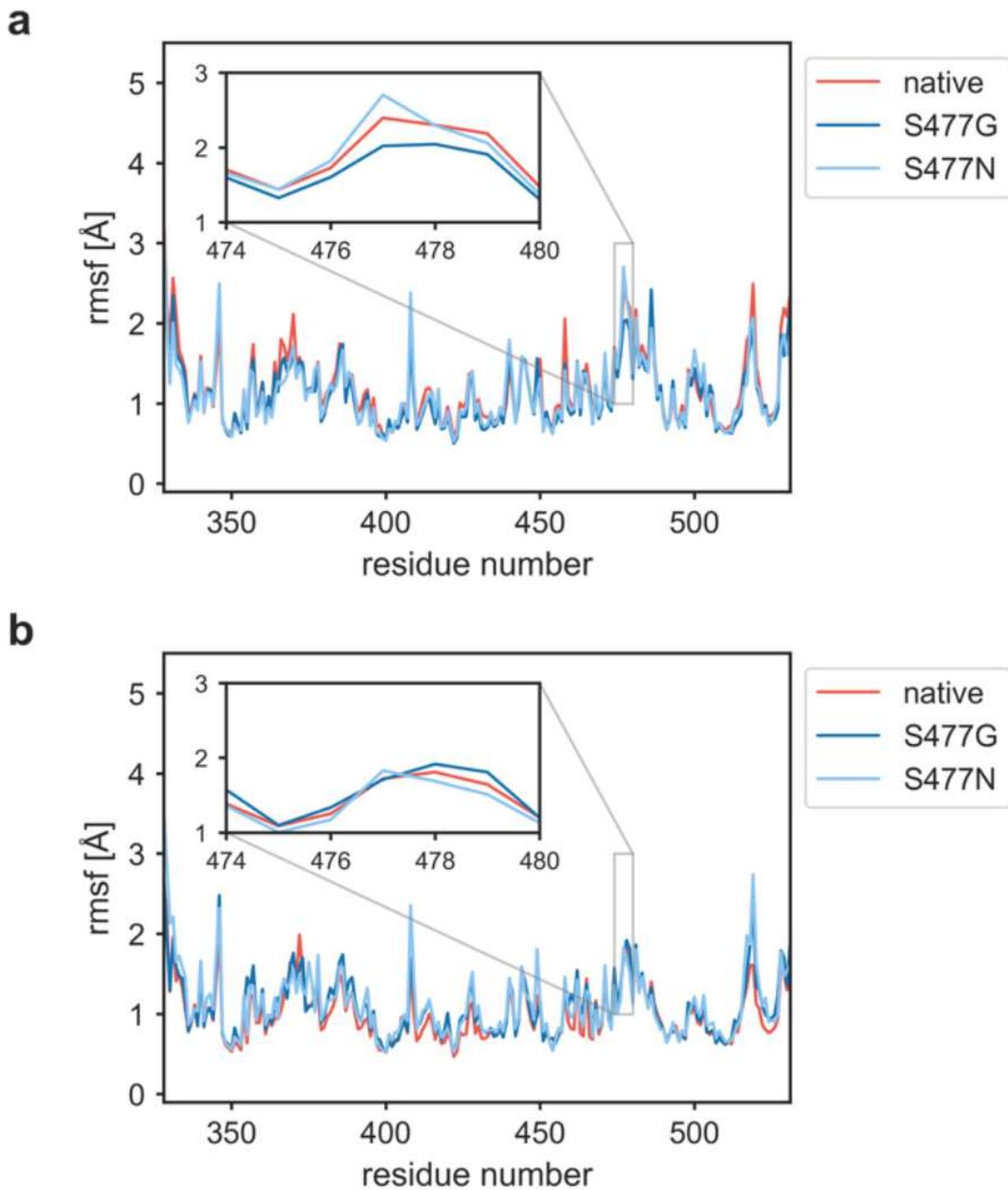
**Figure 2**

DynaMut flexibility analysis performed over the first 10 non-trivial modes of the molecule a) showing the residue wise atomic fluctuations in RBD, whereas the inset figure highlighting the residue wise atomic fluctuations in RBM b) Annotation of atomic fluctuations in RBD shown by the magnitude of the fluctuation represented by thin to thick tube, where blue colour corresponds to RBM residues c) Annotation of dynamical cross correlational among residues of RBD (dimmed in grey color), whereas RBM is shown in light blue cartoon. Only the highly correlated pairs of residues with a correlation coefficient of ( $\geq 0.9$ ) are annotated and connected by red lines.



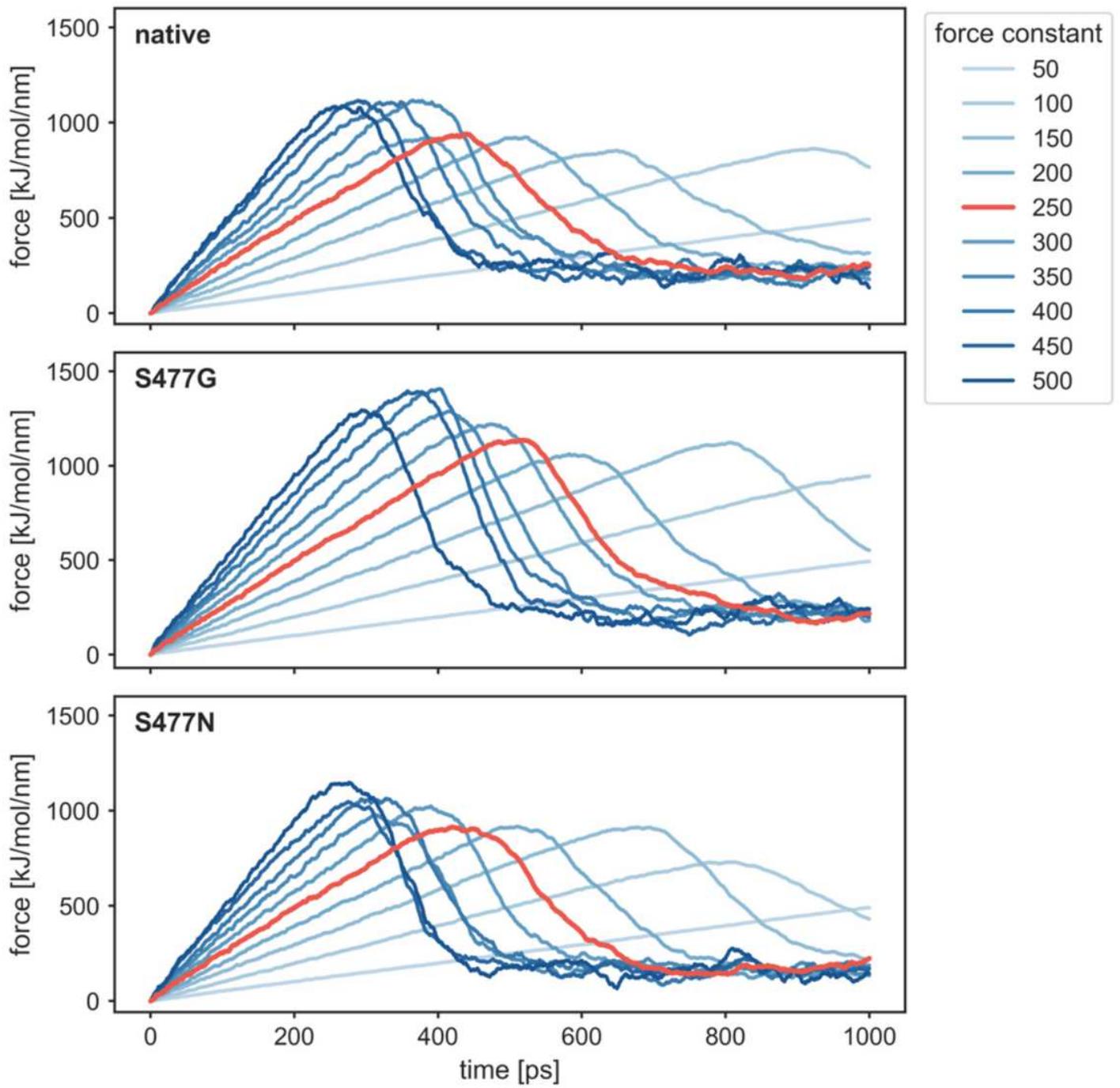
**Figure 3**

Volumetric map analysis of 100 ns trajectory showing the hACE2 residues within 5 Å of RBD, where RBD is shown in grey cartoon and residue 477 highlighted in yellow balls representation.



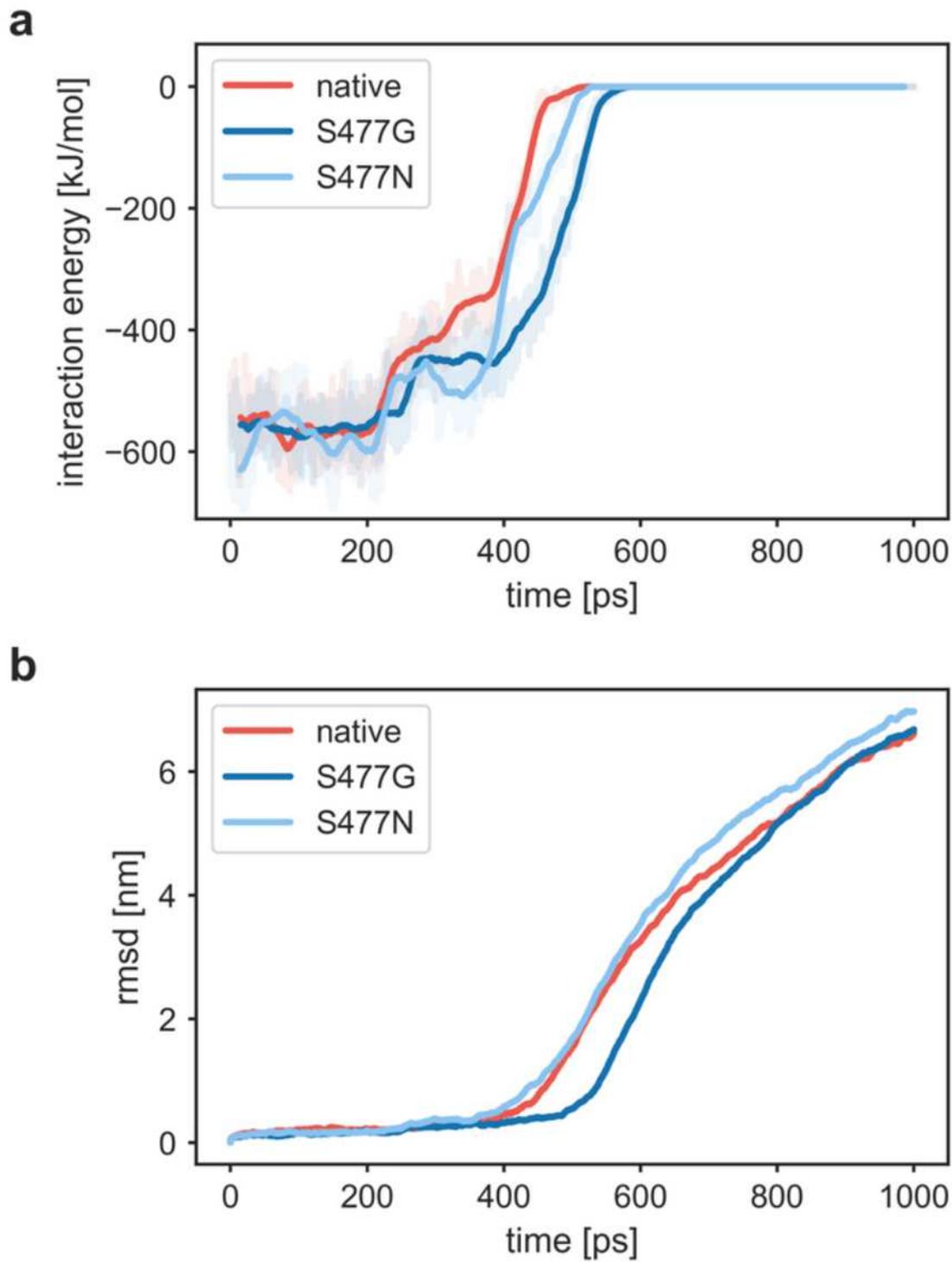
**Figure 4**

Comparison of residue wise root mean square fluctuations, each graph is averaged over three independent 10 ns simulations, standard error is shown in transparent thick boundaries along the lines. For clarity the first 10 residues of N and C terminal of RBD not shown. a) For unbound RBD and its S477 mutants b) RBD and its S477 mutants complexed with hACE2



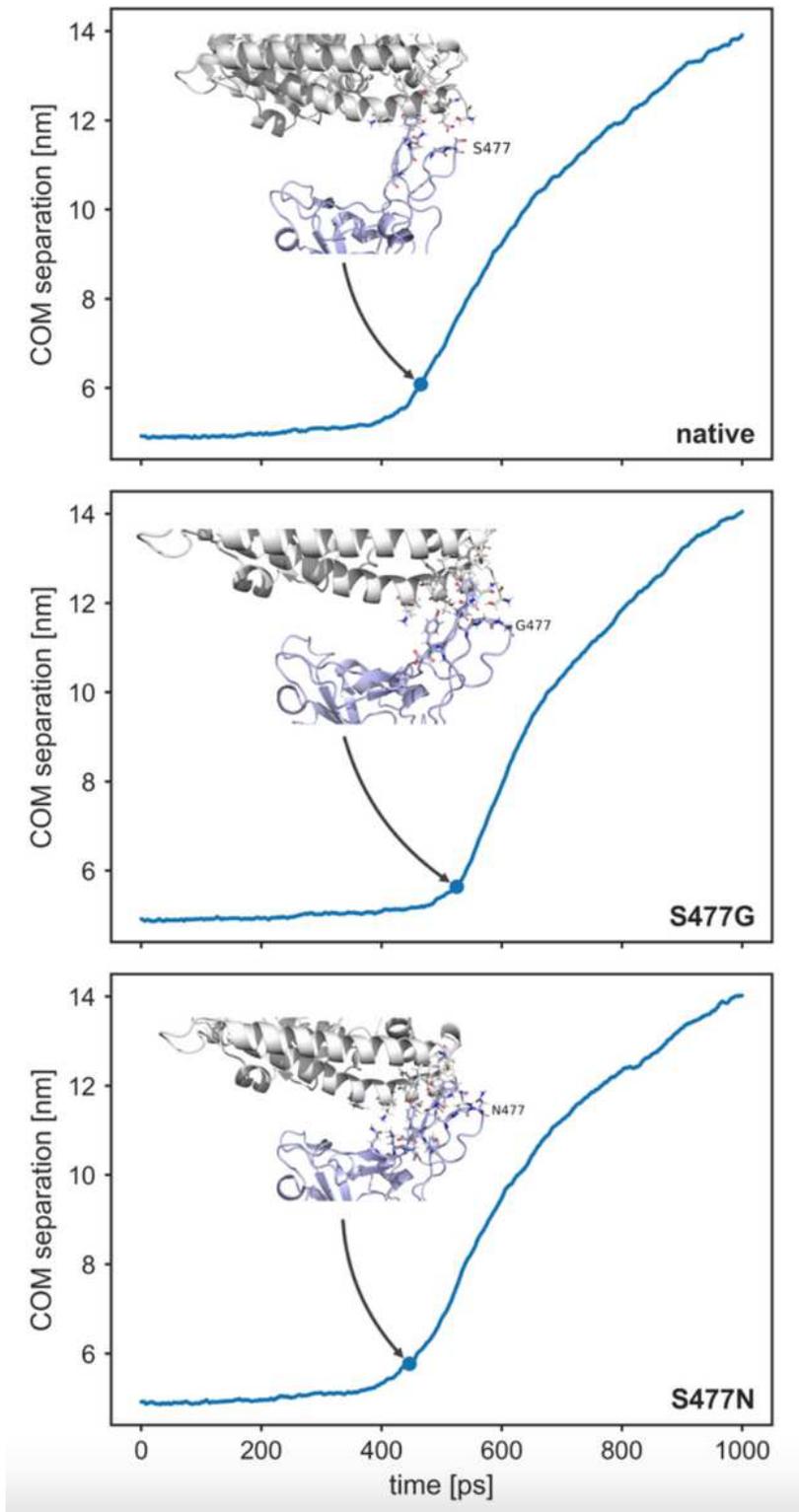
**Figure 5**

Influence of the spring constant on the steering force of RBD: hACE2:Native RBD, hACE2:S477G and hACE2:S477N



**Figure 6**

Energy and conformational changes during SMD simulation at 250 kJ/mol/nm. Time dependence of the interaction energy between hACE2 chain and RBD chain of a) Native (b S477G c) S477N (d The comparison between Native RBD and its mutants e) Comparison of the RMSD of the carbon atoms in the protein backbone relative to the initial frame.



**Figure 7**

Figure showing the snapshot of hACE2:RBD interaction interface at the force rupturing event during SMD simulation. Where hACE2 shown in grey and RBD in light blue color. hACE2:Native RBD, hACE2:S477G and hACE2:S477N.

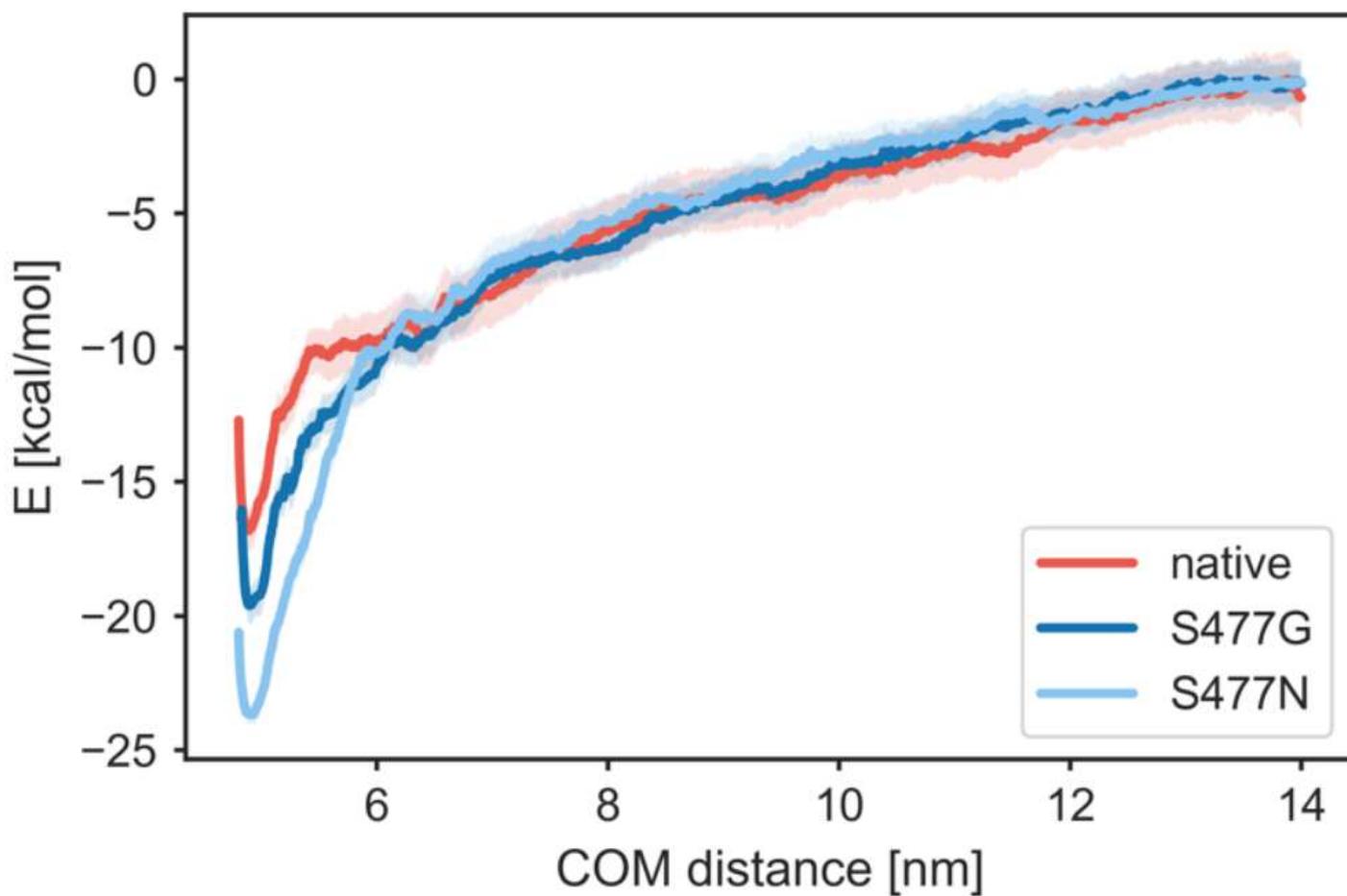


Figure 8

Potential of mean force calculation from SMD simulation of RBD and its S477 mutants. The uncertainty is shown by in the same coloured background on each plot.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementarymovie1.mp4](#)