

Energy decomposition analysis of the SARS-CoV-2 ACE2–RBD interaction

Luis Carlos Ospina, Joseph Nosa and Nicolás Costa

December 30, 2025

Abstract

Protein–protein interactions are important in many biological processes, and small sequence changes can strongly affect the stability of these complexes. In this work, we study the interaction between the receptor binding domain (RBD) of the SARS-CoV-2 Spike protein and the human ACE2 receptor. The protein–protein interface was defined using two methods: a distance-based method and an energy-based method. The interaction energy was divided into electrostatic, van der Waals, and solvation contributions. The importance of individual interface residues was studied using both direct interaction energies and an *in silico* alanine scanning approach. In addition, known SARS-CoV-2 variants were analyzed and the results were compared with calculations performed using FoldX.

1 Introduction

Protein–protein interactions are important in many biological processes. Small changes in the structure or in the amino acid sequence of a protein can strongly affect how two proteins bind to each other. For this reason, studying the interaction energy between proteins can help us understand which residues are important for complex stability.

One well-known example is the interaction between the receptor-binding domain (RBD) of the SARS-CoV-2 Spike protein and the human ACE2 receptor. This interaction is a key step in the infection process, as it allows the virus to enter human cells. Because of this, many vaccines and therapeutic strategies aim to block or weaken the RBD–ACE2 binding.

In this work, the protein–protein interface is defined using a distance-based method. Then, the interaction energy is divided into electrostatic, van der Waals, and solvation contributions. This allows us to analyze how different residues contribute to the overall binding energy.

We also compare the distance-based interface with an alternative interface definition based on interaction energy. In addition, the importance of individual interface residues is studied using per-residue energy decomposition and an *in silico* alanine scanning approach. Finally, the effect of selected SARS-CoV-2 RBD variants is analyzed and compared with the results obtained using FoldX.

Overall, this analysis provides a clearer view of the energetic factors that stabilize the RBD–ACE2 complex and highlights the residues that play the most important role in binding.

2 Methodology

2.1 Structural data and preparation

The structure 6M0J was downloaded from the Protein Data Bank (X-ray diffraction, 2.45 Å resolution). The structure was preprocessed using `biobb_structure_checking` in order to obtain a clean and consistent input for interface detection and energy calculations. Only chains A (ACE2) and E (RBD) were kept. All heteroatoms were removed, including ions (e.g., Zn²⁺, Cl⁻), glycans/ligands (NAG), and crystallographic water molecules.

Alternative atom locations were handled by selecting the highest-occupancy conformations. The structure checker also fixed amide assignments when unusual contacts were detected and added missing terminal atoms (OXT) when needed. Finally, hydrogen atoms were added automatically, and histidines were assigned using the default protonation selection, producing a final PDB without heteroatoms and with hydrogens and charges suitable for the next steps. All preprocessing steps were executed through custom scripts: structure cleaning and standardization were performed with `step1_preprocess_structure`, chain extraction (A and E) with `step1_split_chains`, and PDB-to-PDBQT conversion with `pdb_to_pdbqt`. The resulting PDBQT files were used in subsequent steps for charge assignment and interaction calculations.

2.2 Definition of the protein–protein interface

The interface between both proteins was defined using a distance-based criterion. A residue was considered part of the interface if at least one of its heavy atoms was within a given cutoff distance from any atom of the opposite chain. Several cutoff distances (6.0, 6.5, 7.0 and 8.0 Å) were tested in order to evaluate the robustness of the interface definition. The final cutoff of 7.0 Å was selected as it provided the best compromise between capturing the majority of the interaction energy and avoiding the inclusion of residues not directly involved in binding. This choice was further supported by visual inspection of the interface using PyMOL. Interface residues were extracted using the script `step1_interface_residues_6m0j`.

2.3 Interaction energy calculation

The interaction energy between ACE2 (chain A) and the RBD (chain E) was computed as the difference between the total energy of the bound state (the protein complex) and the unbound state, where both chains are considered isolated in solution. Under this definition, the interaction energy represents the energetic contribution arising from the formation of the protein–protein complex.

$$\Delta G^{A-B} = \Delta G_{\text{elect}}^{A-B} + \Delta G_{\text{vdw}}^{A-B} + \Delta G_{\text{solv}}^{A-B} - \Delta G_{\text{solv}}^A - \Delta G_{\text{solv}}^B \quad (1)$$

The total interaction energy was decomposed into three contributions: electrostatic interactions, van der Waals interactions, and solvation free energy. Electrostatic interaction energies were computed for all atom pairs across chains A and E using Coulomb’s law with a distance-dependent dielectric function (Mehler–Solmajer model). Atomic partial charges were taken from CMIP-prepared PDBQT files. Van der Waals interactions were calculated using a

Lennard–Jones potential with parameters derived from AutoDock atom types and combined using standard mixing rules. In both cases, interaction energies were summed over all inter-chain atom pairs.

Solvation free energies were approximated using an ASA-based model. Solvent accessible surface area values were obtained using NACCESS for the protein complex and for each isolated chain. For each atom, the solvation contribution was computed as the product of its ASA and an atom-type-specific solvation parameter. The solvation contribution to the interaction energy was calculated as:

$$\Delta G_{\text{solv}}^{A-B} = G_{\text{solv}}(\text{complex}) - G_{\text{solv}}(A) - G_{\text{solv}}(B) \quad (2)$$

Interaction energies were first computed considering all atoms of both chains. The same calculations were then repeated considering only atoms belonging to interface residues, as defined in the previous subsection. This comparison was used to evaluate whether the distance-based interface definition captured most of the interaction energy associated with complex formation. All energy terms were computed with a script with all energy reusable functions: `energy_utils` and a script which imports this functions and returns the corresponding outputs: `step2_compute_interaction_energy`. All using CMIP-prepared PDBQT charges and NACCESS ASA outputs.

In addition, interactions across the interface were identified and classified according to their chemical nature using geometric and chemical criteria applied to the PDBQT structure. Atom pairs within a maximum scan distance of 7.0 Å were considered and classified as hydrogen bonds, salt bridges, aromatic stacking interactions, or van der Waals/hydrophobic contacts using standard distance thresholds. The detected interactions were summarized both at the atom-pair level and at the residue level. This was computed with the `step2_detect_interactions` script.

2.4 Optional energy-based interface analysis

In addition to the distance-based definition of the protein–protein interface described previously, an optional analysis was performed to explore an alternative interface definition based on interaction energy rather than geometric distance. The objective of this analysis was to evaluate whether energetically relevant residues coincide with the residues identified using a distance cutoff and to assess the robustness of the interface definition.

First, interaction energies were computed on a per-residue basis for all residues in both ACE2 and the RBD. For each residue, the total interaction energy was obtained by summing its electrostatic, van der Waals, and solvation contributions arising from interactions with the opposite chain. This allowed the identification of residues that contribute significantly to the overall interaction energy, independently of their spatial distance from the binding partner. Based on these values, an energy-based interface was defined by selecting residues whose absolute interaction energy exceeded a given threshold.

Different energy thresholds were tested in order to analyze how the size and composition of the energy-based interface change as a function of the cutoff. For each threshold, the resulting energy-based interface was compared with the distance-based interface obtained previously. This comparison was performed using standard metrics such as precision and recall, treating the distance-based interface as a reference. In addition, the fraction of the total interaction energy

captured by the energy-based interface was computed to evaluate how well it represents the overall binding energy.

To further characterize the interface, per-residue interaction energies were also recomputed considering only the residues belonging to the distance-based interface. This analysis allowed a clearer identification of stabilizing and destabilizing residues at the interface and facilitated the interpretation of residue-level energetic contributions. The most stabilizing and most destabilizing residues were highlighted to identify potential energetic hotspots.

Finally, several plots were generated to summarize the results of this optional analysis. These include plots showing the distribution of interaction energies across interface residues, comparisons between full-chain and interface-only interaction energies, and visual comparisons between distance-based and energy-based interface definitions. Together, these analyses provide complementary insight into the relationship between geometric proximity and energetic relevance at the RBD–ACE2 interface and support the validity of the distance-based approach used in the main analysis.

This optional analysis was fully automated using a set of custom Python scripts. Per-residue interaction energies for the full complex were computed using `step2_optional_compute_per_residue_full_energies.py`, while interaction energies restricted to interface residues were obtained with `step2_optional_per_residue_interface_energy.py`.

Energy-based interfaces for different thresholds were generated and compared to the distance-based interface using `step2_optional_comparison_energy_thresholds.py`. The corresponding metrics and visualizations were produced with `step2_optional_plot_thresholds_and_comparison.py` and `step2_optional_plots_interface_energies.py`.

Assumptions

Note that this computation was done under some assumptions. Firstly, it was assumed that no conformational changes occur in the protein structures between the bound and unbound states. Also, only non-bonded interactions between atoms belonging to different chains were considered, as no covalent bonds exist between the two proteins. Third, heteroatoms not directly related to the protein–protein interface were assumed to have a negligible effect and were not included in the calculations. Finally, solvation effects were approximated using ASA-based models for all atom types, and entropic contributions were not explicitly considered. These assumptions allow a simplified but consistent estimation of relative interaction energies and residue contributions.

2.5 Per-residue energy decomposition

To identify the energetic contribution of individual residues at the protein–protein interface, the total interaction energy was further decomposed into per-residue contributions. For each interface residue, the interaction energy was computed by summing all inter-chain interactions involving atoms of that residue and atoms of the opposite chain.

This per-residue decomposition allowed the identification of stabilizing and destabilizing residues at the interface and provided a residue-level view of the interaction energy distribution. The resulting per-residue energies were used as a reference for the alanine scanning analysis described in the following section.

2.6 Alanine scanning

Each interface residue was individually mutated to alanine by removing all side-chain atoms beyond the C_β atom, while keeping the backbone atoms (N, C_α , C, O) and the C_β . Glycine residues were excluded from this analysis, as they do not contain a side chain beyond the backbone.

For each mutant structure, the interaction energy between ACE2 (chain A) and the RBD (chain E) was recomputed using the same energy model and parameters as for the wild-type complex. The interaction free energy was calculated as:

$$\Delta G^{A-B} = \Delta G_{\text{elect}}^{A-B} + \Delta G_{\text{vdw}}^{A-B} + \Delta G_{\text{solv}}^{A-B} - \Delta G_{\text{solv}}^A - \Delta G_{\text{solv}}^B \quad (3)$$

Electrostatic interactions were computed using Coulomb's law with a distance-dependent dielectric function, van der Waals interactions were modeled using a Lennard–Jones potential with AutoDock-derived parameters, and solvation free energies were estimated using a surface-area-based model. Electrostatic and van der Waals contributions were summed over all inter-chain atom pairs, while solvation energies were computed as the difference between the solvated complex and the isolated chains.

The wild-type interaction energy (ΔG_{WT}) was computed once using interface atoms. For each alanine mutant, the mutated interaction energy (ΔG_{mut}) was computed using the modified atom list. The effect of each mutation was quantified as:

$$\Delta\Delta G = \Delta G_{\text{mut}} - \Delta G_{\text{WT}} \quad (4)$$

Positive $\Delta\Delta G$ values indicate destabilizing mutations, identifying residues that contribute favorably to binding in the wild-type complex, while negative values indicate stabilizing or energetically neutral substitutions. The results were stored in a CSV file and visualized using bar plots to highlight the most important energetic hotspot residues at the interface.

The Alanine scanning process was done with the `step3_ala_scan_interface` script.

2.7 Variant and FoldX analysis

2.7.1 Variant analysis

To evaluate the effect of clinically relevant SARS-CoV-2 mutations on the stability of the ACE2–RBD complex, we selected four well-known RBD variants: E484K, L452R, N501Y and T478K.

These mutations involve substitutions that modify charge, size, or hydrogen-bonding capability at the receptor-binding interface. E484K replaces a negatively charged glutamate with a positively charged lysine, and this alters local electrostatic interactions and contributes to immune escape. L452R introduces a charged arginine in place of a hydrophobic leucine, strengthening electrostatic contacts and increasing receptor affinity. N501Y replaces asparagine with a larger aromatic tyrosine, enhancing packing interactions and hydrogen bonding at a central hotspot of the interface. Finally, T478K introduces a positively charged lysine near the binding surface, which may modulate both receptor binding and antibody recognition. These substitutions have been selected during viral evolution due to their combined effects on ACE2 binding efficiency

and immune evasion.

Each mutation was introduced by replacing the corresponding side chain in the RBD and keeping the backbone fixed. Then, the interaction energy was recomputed using the same protocol applied to the wild-type complex.

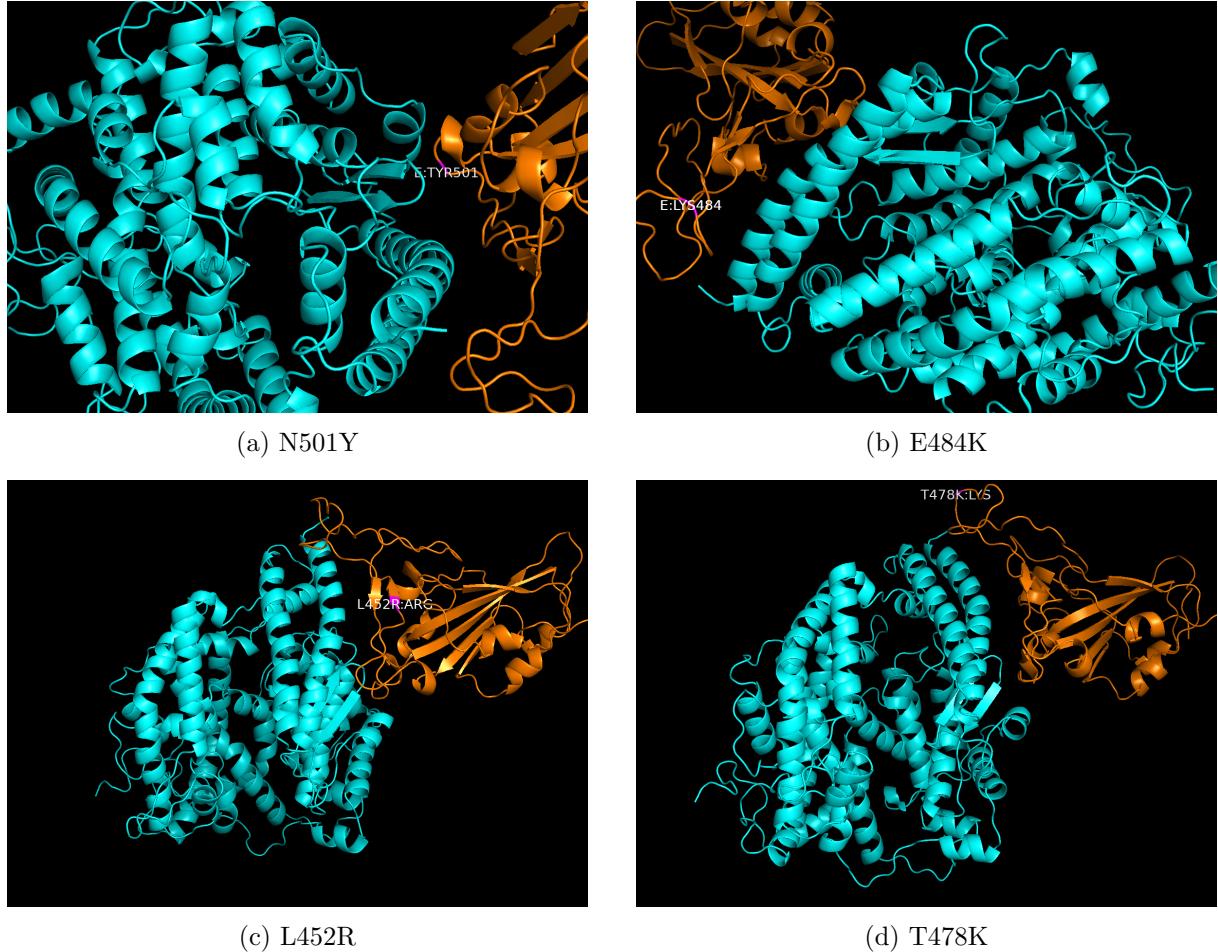


Figure 1: Structural localization of selected RBD mutations on the ACE2–RBD complex. ACE2 (chain A) and the RBD (chain E) are shown in the same orientation across panels, and the mutated residue is highlighted for each variant.

2.7.2 FoldX analysis

FoldX was used as an independent method to estimate the effect of RBD mutations on the stability of the ACE2–RBD complex. The calculations were performed using FoldX version 5.1, which uses an energy function that includes van der Waals interactions, hydrogen bonds, electrostatics, solvation effects and entropic contributions.

Before running the energy calculations, the wild-type structure was processed using the `RepairPDB` command to optimize side-chain conformations and remove local steric clashes. Each mutant structure (E484K, L452R, N501Y and T478K) were then generated starting from the repaired wild-type complex. The interaction energy between ACE2 (chain A) and the RBD (chain E) was computed using the `AnalyseComplex` command.

The objective of using this tool was to compare the effects of the mutations with those obtained from the custom energy decomposition approach.

3 Results

3.1 Interaction energy distribution and interface validation

The contribution of individual residues to the RBD–ACE2 interaction energy was analyzed by combining per-residue energy decomposition with a comparison between distance-based and energy-based interface definitions.

3.1.1 Fraction of interaction energy explained by energetic thresholds

The fraction of the total interaction energy captured by residues exceeding different absolute per-residue energy thresholds is shown in Figure 2. Using a low threshold (0.5–0.75 kcal/mol) captures more than 80% of the total interaction energy. As the threshold increases, the fraction of explained interaction energy decreases progressively, dropping to approximately 60% at 1.25 kcal/mol and below 40% at 3 kcal/mol.

This result indicates that the interaction energy is not concentrated exclusively in a small number of residues. Instead, while some residues act as strong energetic hotspots, a substantial fraction of the binding energy is distributed among residues with moderate contributions. Therefore, applying strict energy cutoffs leads to a significant loss of relevant energetic information.

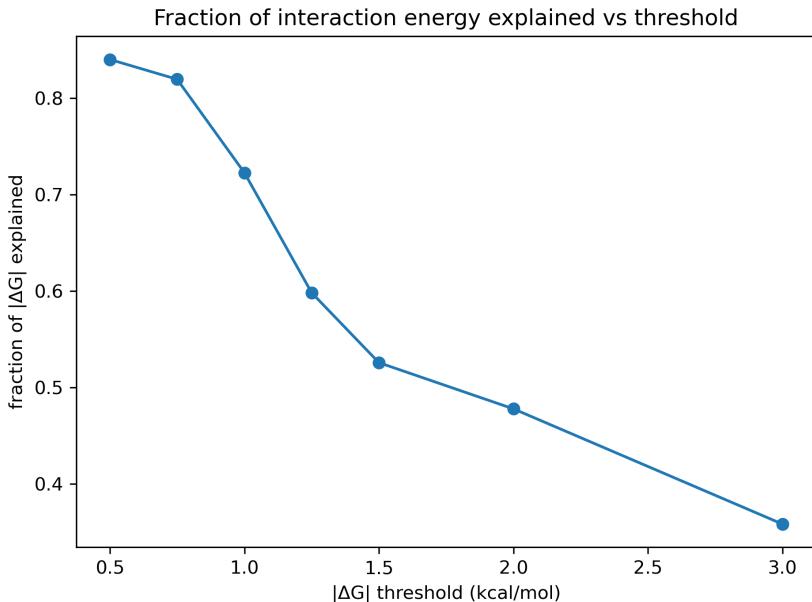


Figure 2: Fraction of total interaction energy explained as a function of the absolute per-residue interaction energy threshold.

3.1.2 Energy-based versus distance-based interface definition

Energy-based interfaces obtained using different interaction energy thresholds were compared with the distance-based interface defined using a 7 Å cutoff. Precision, recall, and F1-score were computed using the distance-based interface as a reference (Figure 3).

At low energy thresholds, the energy-based interface shows moderate precision and recall,

indicating partial overlap with the distance-based definition. As the threshold increases, precision rises sharply and approaches 1.0, meaning that residues selected by high energy thresholds almost always belong to the distance-based interface. However, recall decreases substantially, showing that many geometrically defined interface residues contribute weakly to the interaction energy and are not selected by strict energy-based criteria.

This analysis highlights a trade-off between precision and recall. Energy-based definitions with high thresholds identify a small number of highly stabilizing residues, whereas distance-based definitions capture a broader and more complete set of interface residues.

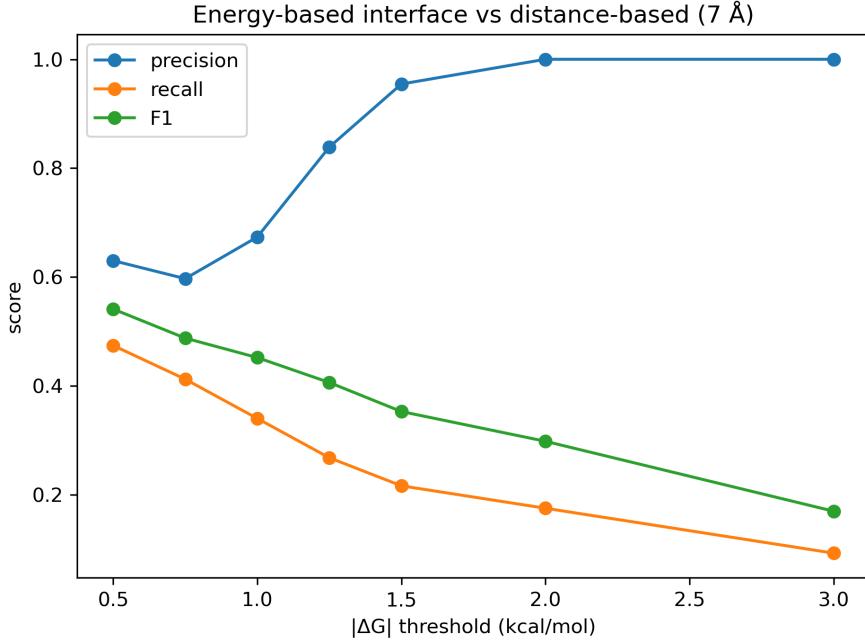


Figure 3: Precision, recall, and F1-score comparing energy-based interface definitions to the distance-based interface (7 Å) for different interaction energy thresholds.

3.1.3 Per-residue interaction energy at the interface

The total interaction energy per residue for interface residues is shown in Figure 4. The distribution is highly asymmetric, with most residues contributing weakly to the interaction energy and a limited number of residues acting as energetic hotspots.

Several residues exhibit strongly stabilizing contributions (negative ΔG), indicating a key role in complex formation. In contrast, a small number of residues contribute unfavorably, suggesting local steric clashes or desolvation penalties. This uneven distribution is characteristic of protein–protein interfaces, where binding affinity is dominated by a subset of residues rather than being evenly distributed across the interface.

3.1.4 Comparison of interface definitions

The overlap between the distance-based interface (7 Å cutoff) and the energy-based interface defined using a 0.5 kcal/mol threshold is illustrated in Figure 5. A substantial number of

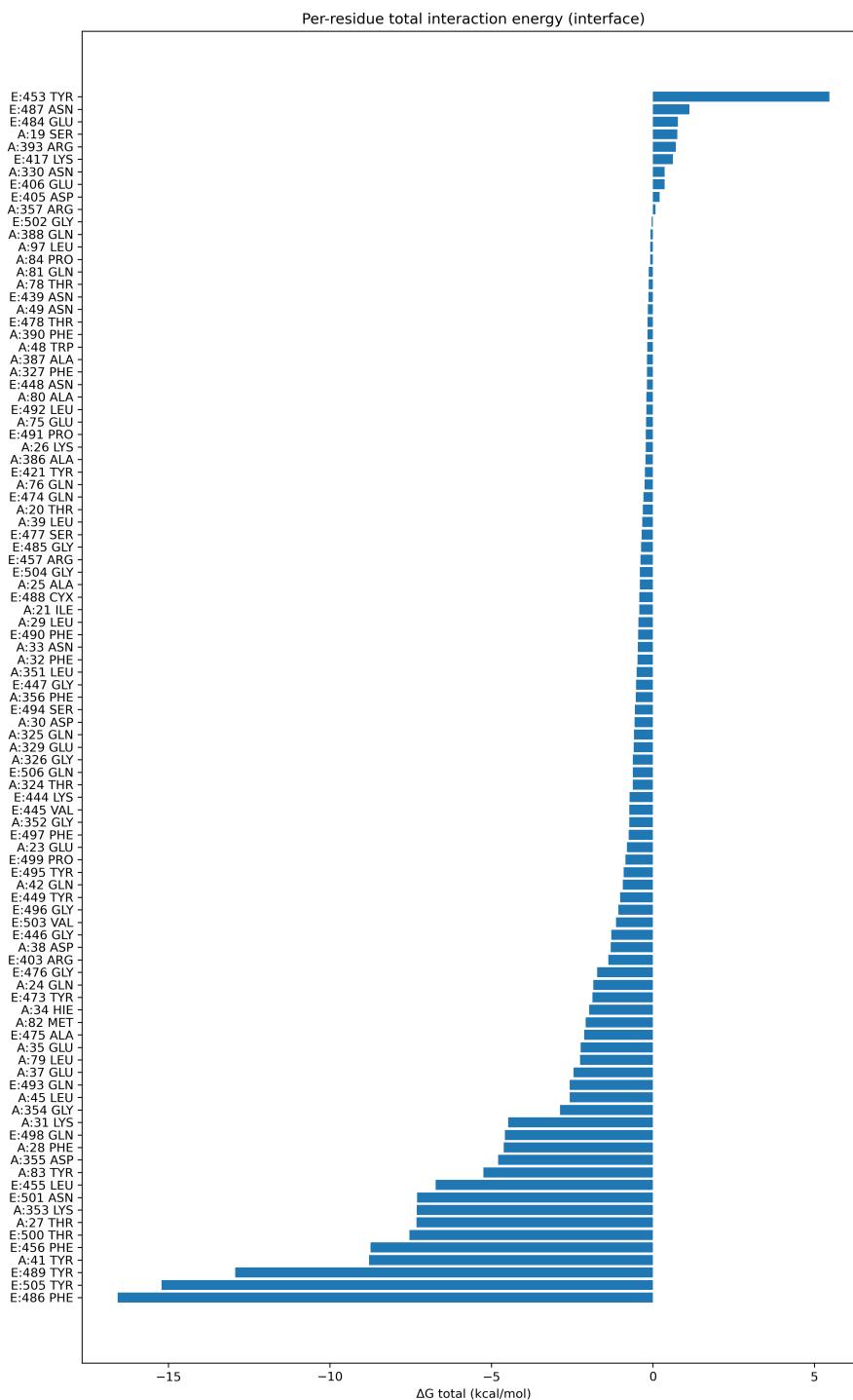


Figure 4: Total per-residue interaction energy for residues at the RBD–ACE2 interface.

residues are shared between both definitions, indicating that many geometrically close residues also contribute significantly to binding energy.

However, each method also identifies unique residues. Some residues are geometrically close but energetically weak, while others contribute energetically despite not being among the closest contacts. This result shows that geometric proximity alone does not fully capture energetic relevance and that energy-based analyses provide complementary insight.

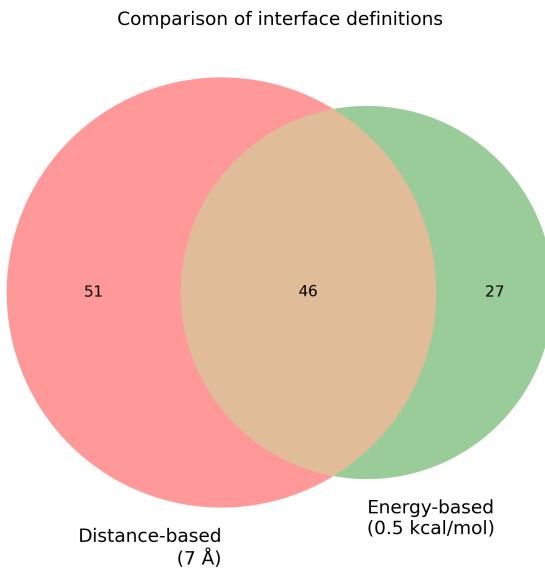


Figure 5: Overlap between distance-based (7 \AA) and energy-based (0.5 kcal/mol) interface definitions.

3.1.5 Energy decomposition by interaction type

The interaction energy was further decomposed into electrostatic, van der Waals, and solvation contributions (Figure 6). Van der Waals interactions constitute the dominant stabilizing contribution for most interface residues, reflecting the importance of packing and shape complementarity at the RBD–ACE2 interface.

Electrostatic interactions contribute significantly for a smaller number of residues, typically involving charged or polar amino acids engaged in salt bridges or hydrogen bonds. Solvation contributions are often unfavorable due to desolvation penalties upon complex formation, but these penalties are compensated by favorable inter-chain interactions.

3.1.6 Full-complex versus interface-only interaction energy

A comparison between the interaction energy computed for the full complex and that computed using only interface residues is shown in Figure 7. The interface-only calculation captures most of the total interaction energy, although it is slightly less favorable than the full-complex value.

This result confirms that the distance-based interface definition is appropriate and that residues outside the interface contribute only marginally to the overall interaction energy.

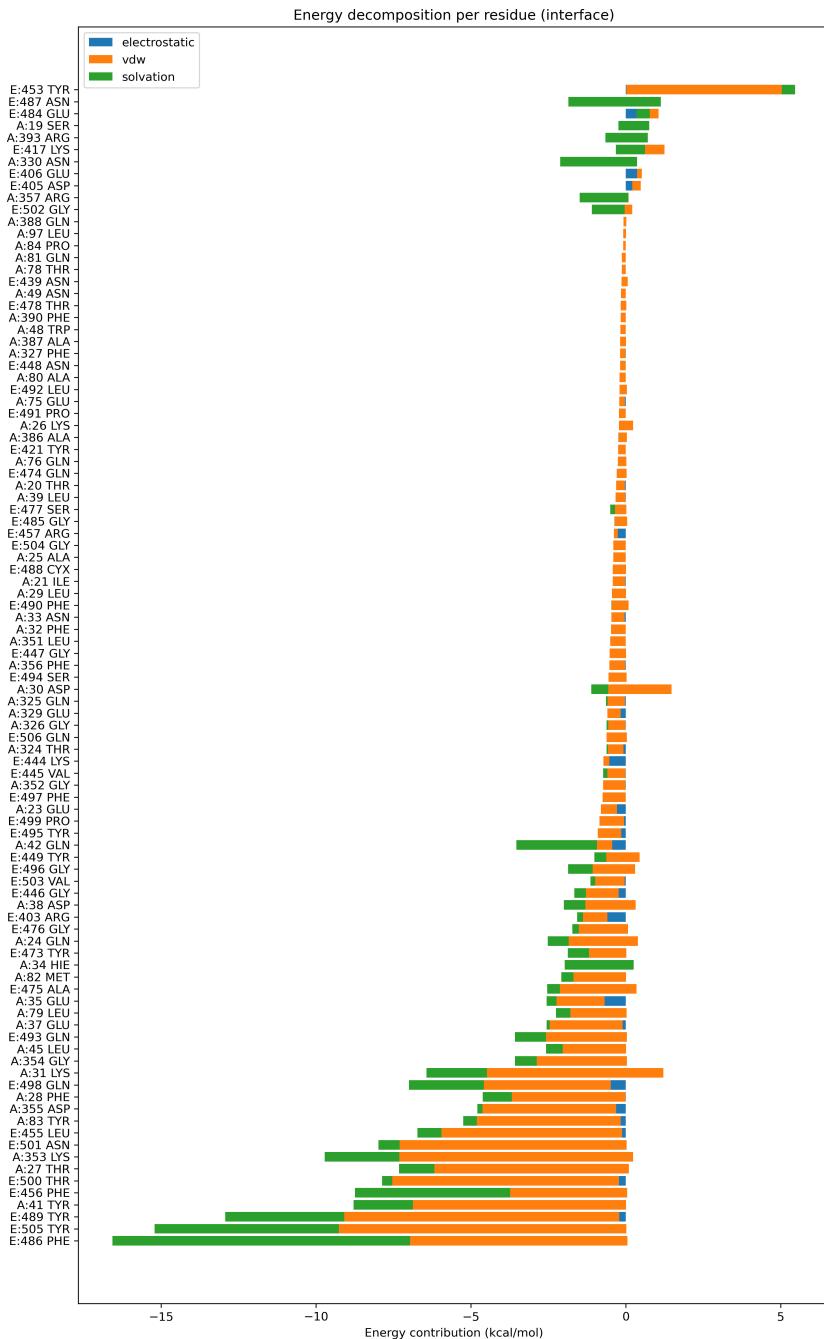


Figure 6: Decomposition of the total RBD–ACE2 interaction energy into electrostatic, van der Waals, and solvation contributions.

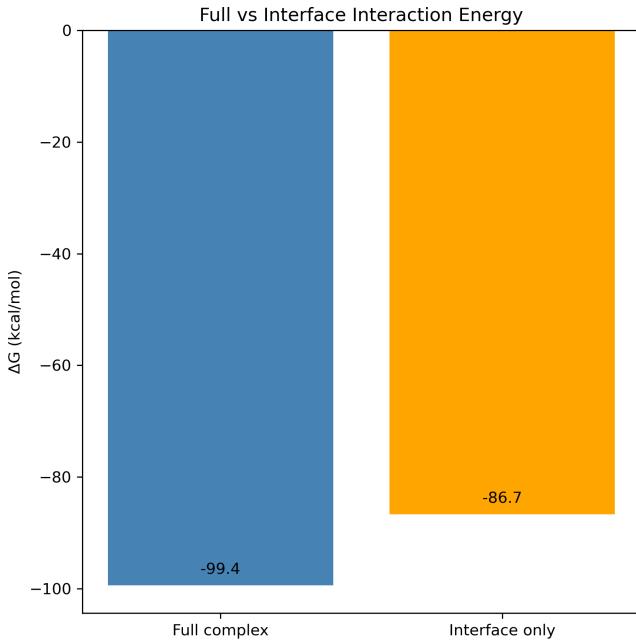


Figure 7: Comparison of interaction energy computed for the full complex and using only interface residues.

3.1.7 Interface contacts by interaction type

To characterize the nature of the RBD–ACE2 interface, detected inter-chain contacts were classified according to their interaction type. The total number of contacts of each type is shown in Figure 8.

The interface is dominated by van der Waals and hydrophobic contacts, which account for the vast majority of detected interactions. This observation is consistent with the strong contribution of van der Waals terms observed in the energy decomposition analysis (Figure 6) and highlights the importance of packing and shape complementarity in stabilizing the RBD–ACE2 complex.

In contrast, hydrogen bonds, salt bridges, and aromatic stacking interactions are much less frequent. Although these interactions are fewer in number, they often involve charged or polar residues and can contribute significantly to specificity and local stabilization of the interface. Together, these results indicate that binding is driven primarily by a dense network of weak hydrophobic contacts, complemented by a smaller number of directional and electrostatic interactions that fine-tune the interaction.

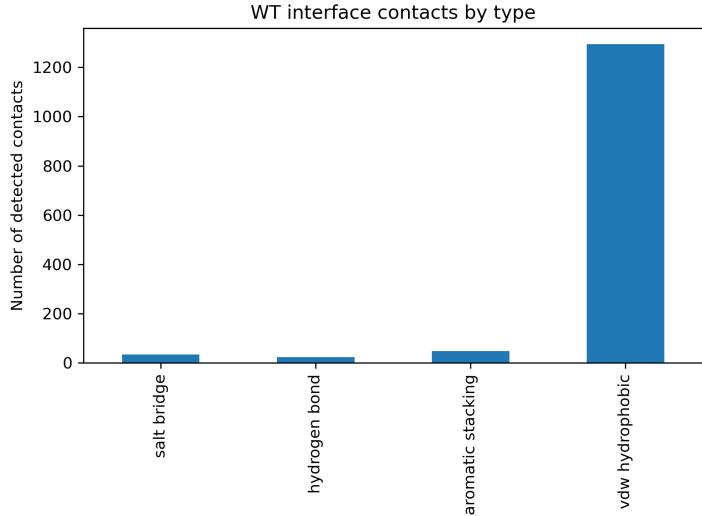


Figure 8: Number of detected inter-chain contacts at the RBD–ACE2 interface classified by interaction type.

3.1.8 Most stabilizing and destabilizing interface residues

To summarize the most relevant energetic contributions at the RBD–ACE2 interface, the results of the per-residue interaction energy analysis were ranked to identify the residues with the largest stabilizing and destabilizing effects.

Figure 9 shows the ten most destabilizing residues upon alanine mutation, i.e. residues with the largest positive $\Delta\Delta G$ values. These mutations significantly weaken the interaction, indicating that the corresponding wild-type residues play an important role in stabilizing the complex. In particular, residue E:453 Tyr shows a markedly larger destabilizing effect compared to the rest, highlighting it as a major energetic hotspot at the interface.

Conversely, Figure 10 displays the ten most stabilizing residues, which exhibit strongly negative interaction energies in the wild-type complex. These residues contribute most favorably to binding and dominate the interaction energy landscape of the interface. The strongest stabilizing contributions are observed for aromatic and hydrophobic residues, such as tyrosine and phenylalanine, consistent with the dominant role of van der Waals interactions observed in the energy decomposition analysis.

Together, these ranked plots provide a compact summary of the most relevant interface residues in energetic terms and serve as an effective alternative to a tabular representation of the strongest interaction contributions requested in Step 2.

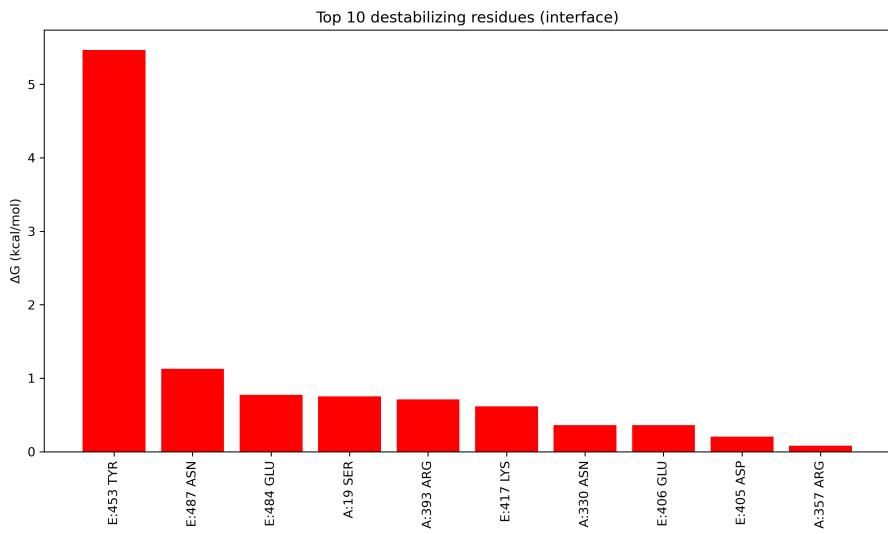


Figure 9: Top ten destabilizing interface residues identified by alanine scanning, ranked by $\Delta\Delta G$. Positive values indicate mutations that weaken binding.

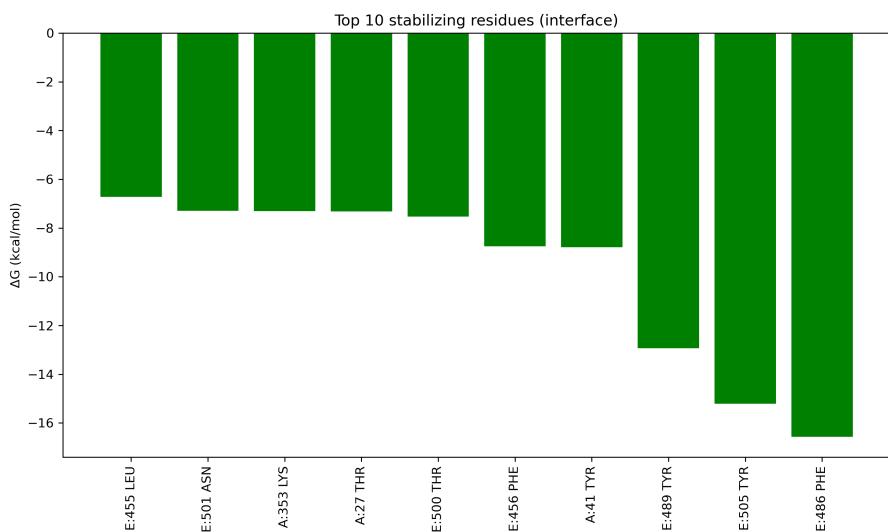


Figure 10: Top ten stabilizing interface residues in the wild-type complex, ranked by their contribution to the total interaction energy. Negative values indicate strong stabilizing effects.

3.2 Energy-based alanine scanning

To quantify the energetic importance of individual interface residues, we performed an *in silico* alanine scan over the set of residues classified as interfacial in Step 2. For each residue, the side chain was replaced by alanine (keeping the backbone fixed) and the change in binding free energy was computed as

$$\Delta\Delta G = \Delta G_{\text{mut}} - \Delta G_{\text{WT}}. \quad (5)$$

A positive $\Delta\Delta G$ indicates that the mutation destabilizes binding (i.e., the wild-type residue is a *hotspot* that contributes favourably to complex formation), whereas a negative $\Delta\Delta G$ indicates a stabilizing mutation.

Figure 11 reports the top 25 residues ranked by $\Delta\Delta G$. The strongest hotspots are located predominantly on the RBD side (e.g., E486, Y505, Y489), consistent with their central position in the binding epitope. On the ACE2 side, residues such as Y41, Y83 and Y27 also show large effects, reflecting their role as key anchors that stabilize the interface through complementary packing and polar contacts. Overall, the alanine scan suggests that a relatively small subset of residues dominates the binding energetics, which is consistent with the “hotspot” nature of many protein–protein interfaces.

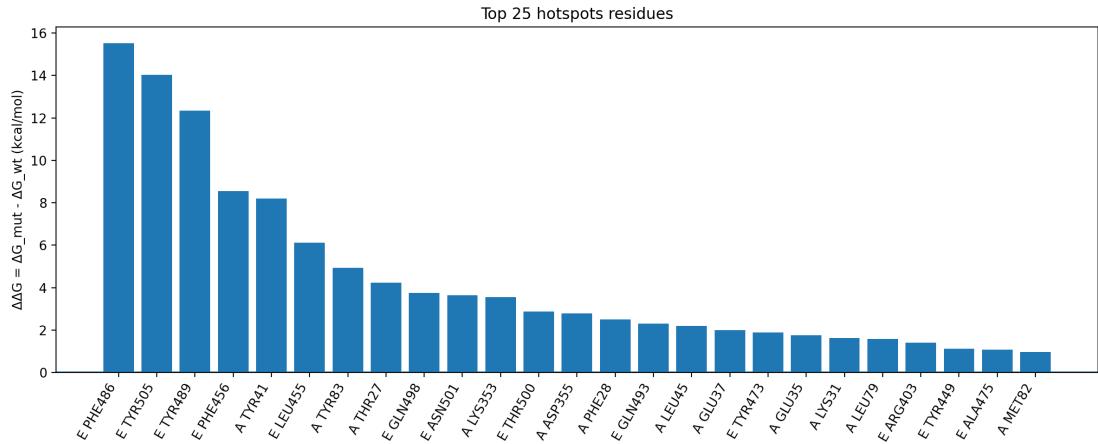


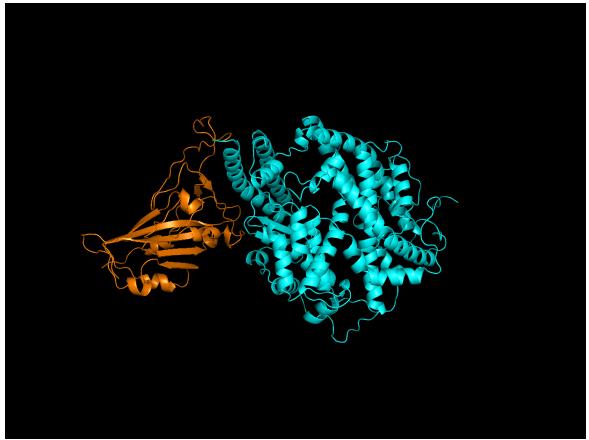
Figure 11: Top 25 interface residues ranked by alanine scanning $\Delta\Delta G$. Positive values indicate destabilizing mutations (hotspots).

3.3 Structural visualization of energetic hotspots and variants

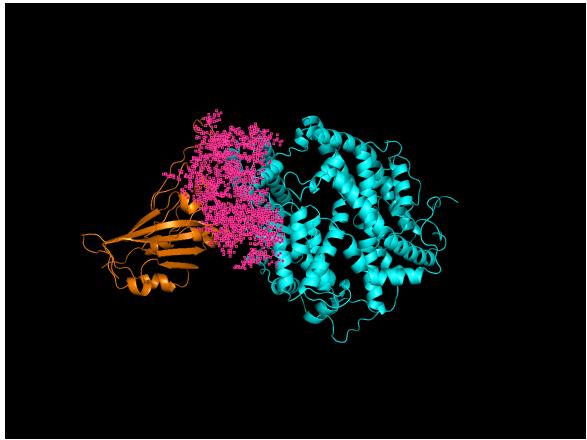
To connect the energetic signatures to structural context, we visualized the ACE2–RBD complex and highlighted (1) the physical interface, (2) energy-based interface residues, (3) the strongest stabilizing hotspots, and (4) clinically relevant RBD mutations.

Figure 12a shows the overall complex architecture (ACE2 in cyan; RBD in orange), while Figure 12b highlights the interfacial region. Using the energy-based interface definition from Step 2 ($|\Delta G| > 0.5$ kcal/mol), residues with strong energetic contributions cluster tightly at the binding surface (Figure 12c), providing structural support for the energy-threshold approach.

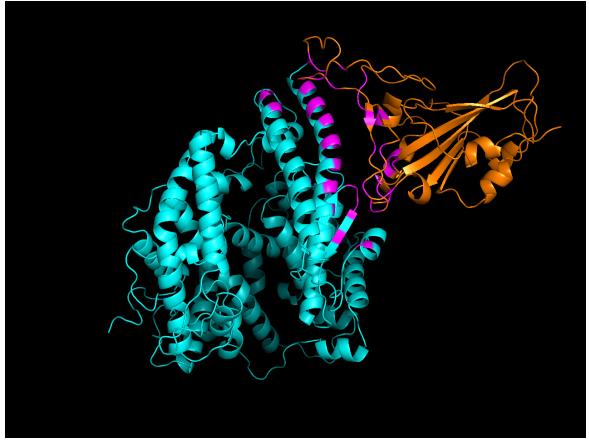
Among these, the most stabilizing hotspots (Figure 12d) are concentrated in the central contact patch and are consistent with the large-magnitude contributions observed in the per-residue energy decomposition and alanine scan.



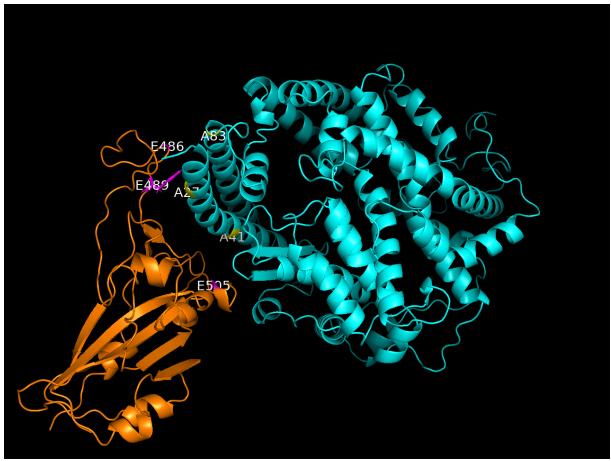
(a) Overall ACE2–RBD complex.



(b) ACE2–RBD interface region.



(c) Energy-based interface residues ($|\Delta G| > 0.5$ kcal/mol).



(d) Top stabilizing interface hotspots.

Figure 12: Structural views of the ACE2–RBD complex highlighting the overall architecture, the binding interface, the energy-based interface residues, and the main stabilizing hotspots.

3.4 Variant analysis and FoldX comparison

3.4.1 Variant analysis: Effect of RBD mutations on interaction energy

Table 1 summarizes the interaction energies of the ACE2–RBD complex. ΔG_{total} corresponds to the interaction energy computed using all atoms of the full protein complex, whereas $\Delta G_{\text{interface}}$ represents the interaction energy contributed by residues at the binding interface. The interface fraction indicates the proportion of the total interaction energy explained by interface residues.

System	Full interaction energy				Interface interaction energy				Interface fraction
	ΔG_{elec}	ΔG_{vdW}	ΔG_{solv}	ΔG_{total}	ΔG_{elec}	ΔG_{vdW}	ΔG_{solv}	$\Delta G_{\text{interface}}$	
WT	-3.50	-89.49	-6.40	-99.38	+2.50	-82.76	-6.40	-86.65	0.87
E484K	-8.59	-87.58	-3.66	-99.83	+2.13	-80.81	-3.66	-82.35	0.82
L452R	-4.84	-87.87	-2.67	-95.38	+3.78	-81.11	-2.67	-80.00	0.84
T478K	-4.71	-87.85	-2.67	-95.23	+3.47	-81.10	-2.67	-80.30	0.84
N501Y	-2.14	-85.44	-3.59	-91.17	+3.91	-78.70	-3.59	-78.38	0.86

Table 1: Energy decomposition of the ACE2–RBD interaction for the wild-type complex and selected RBD variants. The interaction energy is decomposed into electrostatic, van der Waals and solvation contributions, reported both for the full complex (chains A–E) and for the interface region only. Total interaction energies correspond to the sum of all contributions. The interface fraction indicates the proportion of the total interaction energy explained by interface residues. All energies are reported in kcal/mol.

All four mutations modify the interaction energy relative to the wild-type complex, although their effects differ in magnitude. Also, for all variants a large fraction of the total interaction energy is explained by the binding interface, therefore indicating that the mutations primarily affect local interactions rather than inducing global energetic changes.

The E484K mutation leads to the strongest stabilization of the ACE2–RBD interaction in our energy model, being slightly more favorable in total interaction energy terms than the wild-type complex. This effect is mainly driven by changes in electrostatic interactions at the interface, what is consistent with the introduction of a positively charged lysine at position 484.

In contrast, the N501Y mutation shows the largest destabilization among the analyzed variants, with both total and interface interaction energies becoming less favorable compared to the wild-type complex. Despite N501 being located within a known energetic hotspot, the replacement of asparagine by tyrosine alters the local interaction network in a way that is not fully compensated in this energy framework.

The L452R and T478K mutations show intermediate effects, what moderately reduces binding stability.

3.4.2 FoldX comparison

Table 2 summarizes the interaction energies computed with FoldX for the wild-type ACE2–RBD complex and the selected RBD variants. Here, ΔG_{int} corresponds to the interaction energy between chains A (ACE2) and E (RBD) as computed by the `AnalyseComplex` command. For each mutant, the reported value represents the mean interaction energy across the repaired models generated by FoldX, and $\Delta\Delta G$ values are calculated relative to the wild-type complex.

System	ΔG_{int} (kcal/mol)	n	$\Delta\Delta G$ vs WT
WT	-15.69	1	0.00
E484K	-16.27	5	-0.58
L452R	-15.83	5	-0.14
T478K	-15.69	5	0.00
N501Y	-9.03	5	+6.66

Table 2: FoldX interaction energies (ΔG_{int}) between ACE2 (chain A) and RBD (chain E) computed using the `AnalyseComplex` command. For each mutant, the reported value corresponds to the mean interaction energy across the repaired models generated by FoldX (n), with $\Delta\Delta G$ calculated relative to the wild-type complex. All energies are in kcal/mol.

Overall, the FoldX interaction energies show clear differences across the different variants.

The E484K mutation leads to a slightly more favorable interaction energy than the wild-type complex, resulting in a small negative $\Delta\Delta G$ value. This suggests a small stabilization of the ACE2–RBD interaction due to the introduction of a positively charged lysine at the interface.

The L452R and T478K mutations show very limited effects on binding affinity, with interaction energies close to that of the wild-type complex. These results indicate that, within the FoldX energy framework, both mutations have a neutral impact on overall complex stability.

In contrast, the N501Y mutation produces a destabilization of the ACE2–RBD interaction, with a less favorable interaction energy and a positive $\Delta\Delta G$. These results show that, despite being located in a key interaction region, the N501Y substitution alters local contacts in a way that is not fully compensated in the FoldX energy model.

Taken together, FoldX reproduces the main trends observed in the custom energy analysis, with E484K showing stabilization, N501Y causing destabilization, and L452R and T478K having moderate effects.

4 Discussion

In this project, we used a simple energy decomposition approach to study the ACE2–RBD interaction. We also tested different ways to define the interface and evaluated the effect of relevant RBD mutations, comparing our results with FoldX.

First, the distance-based interface definition with a 7.0 Å cutoff captures most of the total interaction energy. For WT and all variants, the interface fraction stays high (0.82–0.87), meaning that most of the binding energy comes from residues located at the contact surface, and that residues outside the interface have a much smaller energetic contribution.

The energy-based interface analysis shows a clear trade-off. Low energy thresholds include many residues and recover a large fraction of the binding energy, while high thresholds select only a few strong residues. With high thresholds, precision is high (most selected residues are truly at the geometric interface), but recall is low (many geometric interface residues are missed). This indicates that binding is dominated by a few strong hotspots, but a large number of residues contribute moderately.

Contact classification shows that most detected contacts are van der Waals/hydrophobic, which matches the strong vdW contribution observed in the energy decomposition. Hydrogen

bonds and salt bridges are less frequent, but they can still be important for specificity and for local stabilization in key regions.

For the variants, changes in binding energy are mainly explained by local effects at the interface. E484K produces the strongest stabilization in our custom energy model, mainly due to electrostatic changes at the interface. In contrast, N501Y shows the strongest destabilization in our framework, suggesting that the mutation disrupts the local interaction network and the effect is not compensated by other terms. L452R and T478K show smaller and more intermediate effects.

FoldX reproduces the main qualitative trends: E484K is slightly stabilizing and N501Y is destabilizing, while L452R and T478K are close to neutral. Differences in absolute values are expected because FoldX uses a different energy function and includes additional terms (e.g., empirical corrections and entropy). Also, our method assumes rigid structures and uses approximate solvation, so results should be interpreted as relative trends within a simplified model rather than exact binding free energies.

Overall, our results support a consistent picture of ACE2–RBD binding, where most of the interaction energy comes from the interface, with a few strong hotspots and many weaker vdW contacts that collectively stabilize the complex.

5 Conclusions

In this work, we applied a simple and interpretable energy decomposition framework to analyze the ACE2–RBD interaction. Despite the approximations involved, the approach allows a clear identification of the main energetic determinants of binding and provides consistent results across different analyses.

The results show that the ACE2–RBD interaction is largely driven by residues located at the binding interface, with a limited number of residues acting as energetic hotspots and a broader set of residues contributing through weaker interactions. In particular, residues such as F486, Y505 and Y489 on the RBD, and Y41, Y83 and Y27 on ACE2, are key contributors to complex stability.

The analysis of SARS-CoV-2 variants indicates that mutations mainly affect binding through local changes at the interface, and FoldX supports the observed trends. While the absolute interaction energies should not be interpreted as exact binding free energies, the model is well suited for relative comparisons between residues and variants.

Overall, this study shows that simplified energy-based analyses can capture the essential features of protein–protein interactions and provide useful insight into the energetic organization of the ACE2–RBD complex.

Author Contributions

Luis Carlos Ospina performed the structural preparation and interface definition (Step 1), implemented the optional energy-based interface analysis in Step 2, and contributed to the overall structure and organization of the report.

Joseph Nosa carried out the alanine scanning analysis (Step 3) and the structural visualization of interface residues and interactions using PyMOL (Step 4).

Nicolás Costa implemented the interaction energy calculations and interface validation (Step 2), and performed the variant analysis and FoldX calculations (Steps 5 and FoldX).

All authors discussed the results and approved the final version of the manuscript.

References

- [1] Lan, J. et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature*, 2020.
- [2] Mehler, E. L., Solmajer, T. Electrostatic effects in proteins. *Protein Engineering*, 1991.
- [3] Schymkowitz, J. et al. The FoldX web server. *Nucleic Acids Research*, 2005.
- [4] Hubbard, S. J., Thornton, J. M. NACCESS, Computer Program. Department of Biochemistry and Molecular Biology, University College London, 1993.