

RESEARCH ARTICLE

Mechanistic insights into ligand dissociation from the SARS-CoV-2 spike glycoprotein

Timothy Hasse¹, Esra Mantei¹, Rezvan Shahoei¹, Shristi Pawnikar^{2,3}, Jinan Wang^{2,3}, Yinglong Miao^{2,3}, Yu-ming M. Huang^{1*}

1 Department of Physics and Astronomy, Wayne State University, Detroit, Michigan, United States of America, **2** Department of Molecular Biosciences, University of Kansas, Lawrence, Kansas, United States of America, **3** Center for Computational Biology, University of Kansas, Lawrence, Kansas, United States of America

✉ Current address: Department of Pharmacology and Computational Medicine Program, The University of North Carolina at Chapel Hill, North Carolina, United States of America

* ymhuang@wayne.edu

**OPEN ACCESS**

Citation: Hasse T, Mantei E, Shahoei R, Pawnikar S, Wang J, Miao Y, et al. (2024) Mechanistic insights into ligand dissociation from the SARS-CoV-2 spike glycoprotein. *PLoS Comput Biol* 20(3): e1011955. <https://doi.org/10.1371/journal.pcbi.1011955>

Editor: Alexander MacKerell, University of Maryland School of Pharmacy, UNITED STATES

Received: December 8, 2023

Accepted: February 29, 2024

Published: March 7, 2024

Copyright: © 2024 Hasse et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The simulation input files and eight successful ligand dissociation trajectories obtained from LiGaMD simulations are freely available at <https://doi.org/10.5281/zenodo.10702784>.

Funding: This work was supported by the Wayne State University Start-up fund to Y.M.H. The LiGaMD simulations were supported by the National Energy Research Scientific Computing (NERSC) Center and the Wayne State University High-Performance Computing Center. The funders

Abstract

The COVID-19 pandemic, driven by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has spurred an urgent need for effective therapeutic interventions. The spike glycoprotein of the SARS-CoV-2 is crucial for infiltrating host cells, rendering it a key candidate for drug development. By interacting with the human angiotensin-converting enzyme 2 (ACE2) receptor, the spike initiates the infection of SARS-CoV-2. Linoleate is known to bind the spike glycoprotein, subsequently reducing its interaction with ACE2. However, the detailed mechanisms underlying the protein-ligand interaction remain unclear. In this study, we characterized the pathways of ligand dissociation and the conformational changes associated with the spike glycoprotein by using ligand Gaussian accelerated molecular dynamics (LiGaMD). Our simulations resulted in eight complete ligand dissociation trajectories, unveiling two distinct ligand unbinding pathways. The preference between these two pathways depends on the gate distance between two α -helices in the receptor binding domain (RBD) and the position of the N-linked glycan at N343. Our study also highlights the essential contributions of K417, N121 glycan, and N165 glycan in ligand unbinding, which are equally crucial in enhancing spike-ACE2 binding. We suggest that the presence of the ligand influences the motions of these residues and glycans, consequently reducing accessibility for spike-ACE2 binding. These findings enhance our understanding of ligand dissociation from the spike glycoprotein and offer significant implications for drug design strategies in the battle against COVID-19.

Author summary

Our work explores the intricate process of ligand dissociation from the spike glycoprotein of SARS-CoV-2, the virus responsible for the ongoing COVID-19 pandemic. The spike glycoprotein plays a crucial role in the virus's ability to infect human cells by interacting with the ACE2 receptor. Our study investigates the molecular interactions involving the

had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

spike glycoprotein and linoleate, a small molecule with the potential to hinder viral entry by blocking ACE2 interaction. Using advanced computational simulations, we uncover the dynamic mechanisms governing linoleate dissociation from the spike protein, unveiling two distinct unbinding pathways. Our findings illuminate the structural changes of the spike glycoprotein during the dissociation process. Notably, we highlight the crucial roles of a gate formed by two α -helices and an N-linked glycan at position N343, which control the dissociation process and determine the unbinding pathway. These insights have significant implications for the development of novel antiviral therapies and improved COVID-19 vaccines. By unraveling the molecular details of this interaction, our work contributes to the global efforts aimed at combatting the COVID-19 pandemic.

Introduction

The ongoing COVID-19 pandemic, caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has had a significant global impact on health and economies, highlighting the urgent need for effective therapeutic interventions [1,2]. The spike glycoprotein of SARS-CoV-2 plays a critical role in viral entry into host cells, facilitated by its interaction with the human angiotensin-converting enzyme 2 (ACE2) receptor [3,4]. Consequently, the development of small molecules capable of disrupting the spike-ACE2 interaction emerges as a promising avenue for antiviral treatments [5,6].

The spike glycoprotein is composed of three identical chains (Fig 1A and 1B), each consisting of two functional subunits: the S1 subunit, which contains the receptor-binding domain (RBD), and the S2 subunit (S1A Fig), which facilitates membrane fusion [7–9]. The RBD binds to the ACE2 receptor on host cells, particularly through the receptor-binding motif (RBM), initiating the viral entry process [10]. The spike glycoprotein exists in two primary conformations: a "closed" state, characterized by all RBDs being in a "down" conformation, rendering them inaccessible for ACE2 binding; and an "open" state, where at least one RBD is in an "up" conformation, facilitating exposure and enabling binding to ACE2 [11,12].

Traditional assessment of protein-ligand binding affinities often relied on the equilibrium dissociation constant (K_D) or half-maximal inhibitory concentration (IC_{50}) [13,14], however a kinetic perspective, focusing on association (k_{on}) and dissociation (k_{off}) rate constants, has proven crucial for understanding their dynamic nature and impact on drug efficacy [15,16].

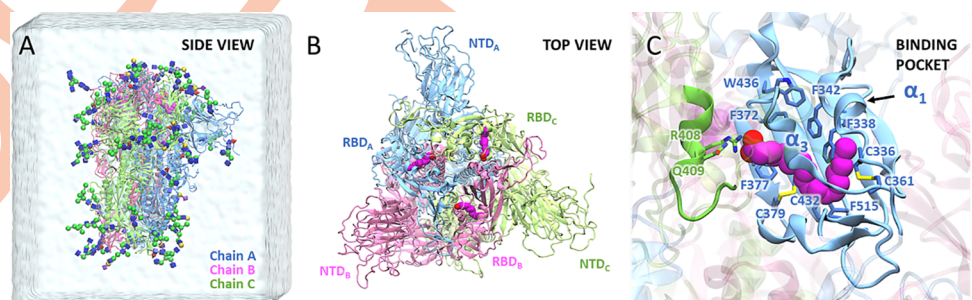


Fig 1. Structure of spike glycoprotein with bound linoleate (LA) molecules. (A) Side view of the spike trimer, comprising Chain A (blue), Chain B (pink), and Chain C (green). The structure is shown in the simulation water box with glycans in the 3D-SNFG cartoon representation. (B) Top view of the spike trimer, highlighting the receptor-binding domain (RBD) and N-terminal domain (NTD). Each RBD contains a bound LA molecule (magenta), with the corresponding subunit indicated as A, B, or C, noted as subscripts. (C) Close-up view of the free fatty acid (FFA) binding pocket within the RBD, illustrating the specific interaction of LA within the pocket.

<https://doi.org/10.1371/journal.pcbi.1011955.g001>

To design inhibitors that can disrupt the spike-ACE2 interactions, extensive investigations have explored the binding affinities of the SARS-CoV-2 spike glycoprotein with various compounds, such as antibodies [17], peptides [17,18], carbohydrates [19], and polyunsaturated fatty acids [20]. Among these spike binders, linoleate (LA), a free fatty acid (FFA) known for its crucial role in inflammation and immune modulation [21,22], has demonstrated remarkable potency in blocking the entry of SARS-CoV-2 [23]. The structure of the spike-LA complex has been revealed through cryo-electron microscopy, uncovering the binding of LA to a hydrophobic pocket within the RBD [23] (Fig 1C). The FFA binding pocket, enclosed by 6 α -helices and 13 β -sheets [24] (S1B Fig), positions at the interface of two RBDs from neighboring chains. The LA hydrophobic tail interacts with nonpolar residues of the α 1 and α 3 helices and the β -sheet on the primary RBD, while the LA hydrophilic head interacts with polar residues on the neighboring RBD (Fig 1C). Additionally, experimental data for K_D , as well as the k_{on} and k_{off} of LA, have been reported [23]. The availability of structural information along with thermodynamic and kinetic properties makes LA an excellent candidate for diffusion mechanism studies that can serve as a model for other potential spike binders.

Computational methods are crucial in spike glycoprotein studies, allowing for in-depth exploration of spike structure, dynamics, and interactions. These methods provide valuable insights that complement experimental data and could greatly contribute to the discovery of potential therapeutics and interventions. *Casalino et al.* [25] investigated the multifaceted roles of glycans in spike viral entry and immune recognition, while *Kapoor et al.* [26] highlighted the impact of posttranslational modifications on spike glycoprotein interactions with host cell receptors. *Wang et al.* [5] and *Piplani et al.* [27] utilized *in silico* techniques to identify small molecules targeting the spike's conserved FFA binding pocket as a starting point for broad-spectrum COVID-19 treatments. *Shoemark et al.* [28] examined the effects of LA and other ligands when bound to the spike FFA binding pocket, revealing insights into conformational stability. Additionally, *Oliveira et al.* [29] demonstrated how the FFA binding pocket influences distant, functionally important regions of the spike, providing valuable perspective into allosteric communication. Despite advancements in spike glycoprotein studies, a detailed understanding of ligand binding and unbinding mechanisms remains limited. To address the knowledge gap, we employ ligand Gaussian accelerated molecular dynamics (LiGaMD) [30] to investigate the kinetics of ligand dissociation from the spike glycoprotein. LiGaMD achieves enhanced sampling of ligand unbinding by selectively adding a harmonic boost potential to the ligand nonbonded interaction energy, enabling accelerated observations of the spike-ligand dissociation process without predefined reaction coordinates [30].

In this study, we conducted all-atom LiGaMD simulations on spike glycoprotein systems with bound LA molecules to characterize the pathways of ligand dissociation and the conformational changes associated with the spike glycoprotein. By analyzing the critical interactions and key residues present in each intermediate state, as well as studying the transitions occurring along these pathways, we identified distinct pathways for LA unbinding. Given that similar intermediates may occur in both ligand binding and unbinding processes [31], our study is dedicated to unraveling the complex process of ligand dissociation. Through this investigation, we aim to provide valuable insights into the kinetics of ligand diffusion and the intrinsic dynamical behavior observed during ligand unbinding processes. We also investigated the corresponding protein conformational changes and examined the influence of glycans on ligand dissociation. Our study sheds light on the kinetic process of ligand unbinding, further enhancing the understanding of the underlying mechanisms that govern ligand interactions, facilitating optimized strategies for drug design.

Methods

Model systems

The SARS-CoV-2 spike protein complex, consisting of three bound LA molecules, was obtained from the Protein Data Bank (PDB) under the code 6ZB5 [23]. This protein complex exhibited a closed conformation with C3 symmetry. The missing loops were built based on the structure reported by *Casalino et al.* [25]. We constructed spike protein systems with varying ligand configurations, including those with one, two, or three ligands present, achieved by manually removing LA molecules from the FFA binding pocket. The protonation states of protein sidechains were determined using the PROPKA program through the PDB2PQR server [32]. Additionally, glycan attachments onto the protein were built using the CHARMM-GUI online server [33,34], following the glycan positioning established in prior research by *Casalino et al.* [25] (S1 Table). The entire system, including the spike protein complex, bound LA molecules, water molecules, and ions, comprised approximately 750,000 atoms.

Simulation protocol

The AMBER20 simulation package [35] was employed for system preparation and enhanced sampling. The AMBER ff19SB force field [36] was utilized for the protein, while GLYCAM_06j [37] and GAFF2 [38] were used for carbohydrates and LAs, respectively. The systems underwent a comprehensive three-stage minimization process. This process involved 5,000 steps of hydrogen minimization, followed by 25,000 steps of sidechain minimization. The final step included 25,000 steps of whole system minimization. Subsequently, the systems were solvated using the TIP3P water model [39], extending to 20 Å from the protein surface. The systems were neutralized and Na⁺ and Cl⁻ ions were introduced at a concentration of 0.15 M to match extracellular NaCl concentration. A secondary round of minimization was conducted, with 5,000 steps dedicated to minimization of water molecules, followed by 25,000 steps of whole system minimization. Following this, the system underwent equilibration through a series of steps. Initial solvent equilibration was performed in the NPT ensemble with isotropic position scaling of the Berendsen barostat, at 310 K for 200 ps while restraining the protein. Subsequent equilibrations were performed in the NVT ensemble at 50, 100, 150, 200, 250, and 310 K. We performed 20 ps at each temperature, except 200 ps at 310 K. A final conventional MD simulation of 20 ns at 310 K ensured the system convergence to the appropriate thermodynamic internal energy before commencing LiGaMD simulations.

LiGaMD is an enhanced sampling technique designed to efficiently explore ligand unbinding pathways by adding a selective boost potential energy to the nonbonded interactions between the ligand and receptor [30]. The specifics of LiGaMD are elaborated in S1 Text. Our LiGaMD simulations comprised an 84 ns preparation phase, followed by a production run of 500 ns or until successful ligand dissociation, whichever happened first. The preparation run commenced with a 4 ns conventional MD stage at 310 K, used to gather potential energy statistics necessary for calculating the harmonic force constant and threshold energy parameters essential for the boost potential evaluation. Next, an 80 ns LiGaMD simulation was performed, in which the boost potential was first applied without parameter updates for 1 ns, and subsequently, the boost parameters were iteratively updated every 1 ns to determine new boost potentials. During the production run, the boost potential remained constant across the entire simulation, applied to both the ligand nonbonded potential energy and the potential energy of the remaining system components for optimal acceleration. The simulations adhered to an upper bound energy limit, maintaining the upper limit of the standard deviation of the ligand nonbonded boost potential at 6.0 kcal/mol. The upper limit for the boost potential of other

energy terms was set in the range of 60–150 kcal/mol, an order of magnitude greater than that of the ligand nonbonded boost potential, as it includes most atoms from protein and water. The details of LiGaMD parameters can be found in [S2 Text](#). All LiGaMD simulations were conducted in the NVT ensemble. These simulations employed Langevin dynamics [40] with a collision frequency of 5 ps^{-1} . A 9 \AA cutoff was applied, and the particle mesh Ewald summation [41] was enabled. Hydrogen-containing bonds were restrained using the SHAKE algorithm [42], and the simulation time step was set to 2 fs. Trajectories were recorded every 10 ps for subsequent analysis. In total, our study included over 100 simulations across systems with one, two, or three bound ligands. Each simulation selectively boosted a single ligand. This collective effort resulted in a simulation time exceeding 50 μs . For detailed descriptions of the simulation lengths for each protocol step, please refer to [S2 Table](#).

Simulation analysis

The analysis of simulation trajectories, encompassing heavy-atom root-mean-square deviation (RMSD), root-mean-square fluctuation (RMSF), distance measurement, and dihedral calculations, was executed using VMD [43] and CPPTRAJ [44] tools. The RMSD of LA relative to the initial equilibrated structure was computed to elucidate the ligand movement throughout the dissociation trajectory. Additional RMSD calculations were carried out for the complete spike protein as well as for the residues within the RBDs. The analysis also involved the calculation of the RMSF to examine the flexibility and rigidity of individual spike protein residues. Two-dimensional potential of mean force (PMF) profiles were derived from the LiGaMD trajectories without energetic reweighting, utilizing the PyReweight toolkit [45]. The PMF, which serves as an expression of free energy along reaction coordinates, was computed using the equation: $F(A) = -k_B T \ln P(A)$, where $F(A)$ represents the free energy along the reaction coordinate A , k_B is the Boltzmann constant, T is temperature, and $P(A)$ denotes the probability distribution of A . The bin size for both ligand RMSD and atomic distance was set at 0.5 \AA .

Results

In this study, we aim to unravel the mechanism of LA unbinding from the spike-LA complex structure by conducting LiGaMD simulations. The complex structure initially contains three LA molecules bound to the spike glycoprotein. To explore the intermediate interactions and different binding pockets during metastates, we performed LiGaMD simulations on seven holo spike systems, each containing varying numbers of LA molecules (one, two, or three). Since the current version of LiGaMD allows selective boosting of only one bound ligand in the simulation, we designed twelve simulation systems to encompass all possible scenarios. These systems included LA bound to different RBD subunits and were denoted as LA-RBD_A^A , LA-RBD_B^B , LA-RBD_C^C , LA-RBD_{AB}^A , LA-RBD_{AB}^B , LA-RBD_{BC}^B , LA-RBD_{BC}^C , LA-RBD_{AC}^A , LA-RBD_{AC}^C , LA-RBD_{ABC}^A , LA-RBD_{ABC}^B , and LA-RBD_{ABC}^C . The subscripts indicate the RBD subunit(s) that have LA bound molecules at the start of the simulations, while the superscripts represent the RBD subunit that the boosted ligand was bound to. To achieve successful ligand dissociations, we gradually increased the boost potential, starting from a low value and incrementally raising it until the ligand dissociation event was observed. Notably, although the spike protein exhibits trimeric symmetry with identical protomers, our models incorporated specific glycan types and positions, which were found to differ between each chain based on previous work [25,46]. This distinction resulted in unique environments for the ligands bound to each chain.

Table 1. List of successful spike-LA dissociation simulations. The subscript and superscript of RBD indicate the domains that contain bound and boosted LAs, respectively. $\sigma_{L, nb}$ and σ_D represent the user-defined upper limit of the standard deviation of the boost potential added to the nonbonded ligand potential terms and to the rest of the potential terms, respectively (details are in S1 Text).

System	Bound Ligands	Boosted Ligand	$\sigma_{L, nb}$ (kcal/mol)	σ_D (kcal/mol)	Simulation Time	Dissociation Time	Dissociation Pathway
LA-RBD _A ^A -r1	RBD _A	RBD ^A	6.0	60.0	160.0 ns	143.4 ns	B
LA-RBD _A ^A -r2	RBD _A	RBD ^A	6.0	150.0	200.0 ns	184.4 ns	B
LA-RBD _A ^A -r3	RBD _A	RBD ^A	6.0	150.0	140.0 ns	119.7 ns	A
LA-RBD _A ^A -r4	RBD _A	RBD ^A	6.0	150.0	200.0 ns	178.2 ns	B
LA-RBD _B ^B	RBD ^B	RBD ^B	6.0	150.0	360.0 ns	348.7 ns	A
LA-RBD _{ABC} ^B	RBD _{ABC}	RBD ^B	6.0	120.0	420.0 ns	392.0 ns	A
LA-RBD _{AC} ^C	RBD _{AC}	RBD ^C	6.0	150.0	300.0 ns	287.0 ns	B
LA-RBD _{ABC} ^C	RBD _{ABC}	RBD ^C	6.0	60.0	360.0 ns	343.1 ns	A

<https://doi.org/10.1371/journal.pcbi.1011955.t001>

Two distinct LA dissociation pathways captured by LiGaMD simulations

The ligand movement was evaluated by calculating the heavy-atom RMSD of the LA molecule relative to its initial equilibrated structure. Complete dissociation of the ligand from the binding pocket was defined as when the ligand RMSD values exceeded 30 Å. Our LiGaMD simulations yielded a total of eight successful dissociation events, with dissociation times ranging from approximately 100 to 400 ns (Table 1 and Fig 2A). Among these eight dissociation events, four originated from the system with a single LA bound initially to RBD_A. Two dissociation events occurred from LA bound to RBD_B, involving one in a system with a single bound ligand and the other with three bound ligands. The remaining two dissociation events involved the unbinding of LA from RBD_C in systems with two and three bound ligands (Table 1).

During these dissociations, two distinct unbinding pathways, referred to Path A (S1 Movie) and Path B (S2 Movie) (Fig 2B), were identified. Path A entails the movement of LA between the α_1 and α_3 helices, whereas Path B involves LA exiting the pocket through the region between the α_1 helix and the β -sheet. Notably, our findings did not indicate any evident correlation between the unbinding pathway and the initial binding pocket, the initial number of bound ligands in the system, or the applied boost potential.

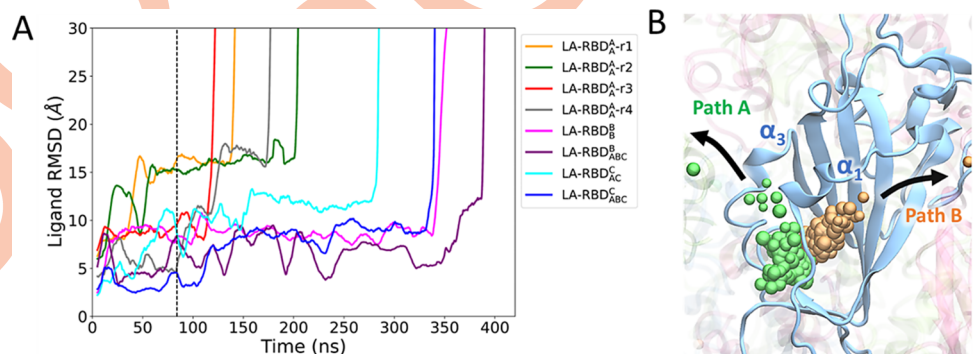


Fig 2. Linoleate (LA) dissociation pathways. (A) The RMSD of LA relative to the initial equilibrated structure for each successful dissociation simulation. A successful dissociation is defined by a ligand RMSD greater than 30 Å. The simulation time before and after the dashed line indicates the GaMD preparation and production run, respectively. (B) Representative trajectories depicting two distinct dissociation pathways: Path A (green) and Path B (orange). The ligand trajectory is shown by individual beads, representing the center of mass coordinates of LA over the simulation time.

<https://doi.org/10.1371/journal.pcbi.1011955.g002>

Our simulations started with the LA molecule fully bound in the FFA binding pocket. The polar head of LA predominantly interacts with polar residues R408, Q409, and K417 on the neighboring RBD loop, forming a polar anchor, while the hydrophobic tail is positioned between the $\alpha 1$ and $\alpha 3$ helices as well as the β -sheet of the primary RBD, effectively locking LA in place (S2A Fig). Following the initial bound state, LA undergoes a flipping motion, reorienting itself within the binding pocket. This motion causes the polar head group of LA to face away from the neighboring RBD and eventually exit the binding pocket, following either Path A or Path B. (S2B Fig). The ligand flipping motion typically occurs within the initial 20 ns, and the subsequent dissociation of the polar interaction transpires within the first 60 ns (S2B Fig). During LA dissociation, its polar head extends out of the binding pocket and positions the ligand between $\alpha 1$ and $\alpha 3$ for Path A, or between $\alpha 1$ and the β -sheet for Path B. Concurrently, the hydrocarbon tail gradually moves away from the surrounding hydrophobic residues (S2C and S2D Fig). LA may rebind to the pocket due to the interplay between solvent attraction to the LA polar head and the interaction between the fatty acid tail and hydrophobic residues of the binding pocket. This delicate balance between solvent and hydrophobic interactions plays a crucial role in determining whether LA remains in the bound state or dissociates from the pocket. From S3 Table, we found that after LA begins to dissociate, the number of water molecules around LA increases, indicating that solvent attraction contributes to LA dissociation from the binding pocket. Ultimately, regardless of whether LA follows Path A or Path B, both the polar head and the nonpolar tail of LA become fully exposed to the solvent (S2E and S2F Fig). It is worth noting that all unboosted ligands show minimal fluctuations and consistently remain fully bound throughout the simulations, resembling their initial conformation. This highlights their stability within the binding pocket during the simulations.

Conformational flexibility of the spike glycoprotein during LA dissociation

During the dissociation of LA from the complex, both the ligand and the spike glycoprotein experience significant conformational changes and dynamical fluctuations. We conducted an analysis to assess the conformational flexibility of the spike glycoprotein and the regions involving the RBDs during the ligand dissociation process. To examine the overall motion of the spike glycoprotein, we calculated the RMSD of the entire spike throughout the simulations (S3 Fig). The RMSD values generally increased as the ligand dissociation progressed. Initially, the change of RMSD was around 4 Å, but it typically rose above 6 Å and occasionally approached 8 Å, indicating deviations from the initial state and highlighting the impact of ligand dissociation on the overall protein structure. Next, we focused on the RMSD of the RBDs (S4 Fig), as these regions directly participate in ligand unbinding. Throughout the ligand dissociation trajectories, the RMSD of each RBD exhibited fluctuations. Specifically, when LA transitioned from the initial bound state to either Path A or Path B, the RMSD of the corresponding RBD increased, indicating conformational changes associated with the ligand dissociation. Interestingly, in cases where the LA polar head rebounded to the polar anchor residues R408, Q409, and K417, the RMSD of the RBD decreased rapidly within a short period of time. This dynamic behavior of the RBDs suggests structural changes accompanying the ligand dissociation process, significantly influenced by the ligand interaction with nearby residues.

We conducted RMSE analysis to explore the dynamic fluctuations of individual residues within the RBDs during LA dissociation (S5 Fig). Our investigation focused on the RMSE profiles of the RBDs containing a boost ligand in three distinct states: 1) the fully bound state, 2) during traversal along either Path A or Path B, and 3) following complete dissociation of LA from the RBD. S5 Fig illustrates that throughout the dissociation trajectory, both the $\alpha 1$ and $\alpha 3$ helices exhibit significant fluctuations, while the β -strands shown in S2 Fig generally display

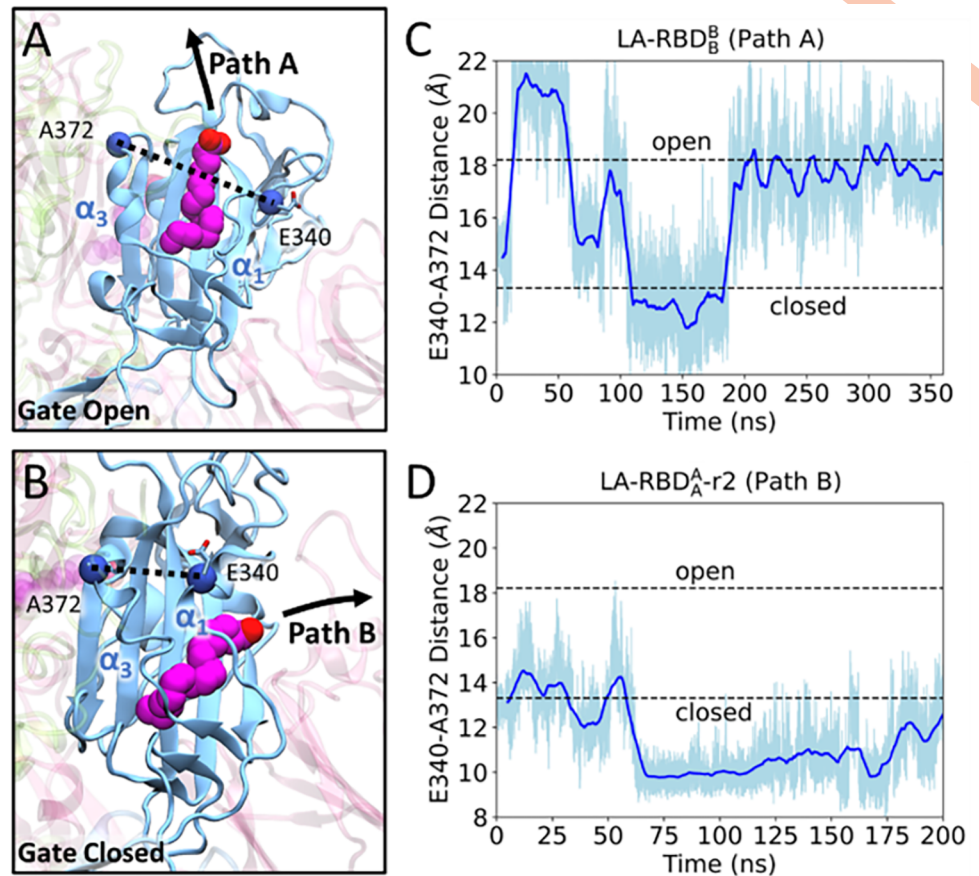


Fig 3. Dynamic opening and closing of the FFA binding pocket. (A) Ligand dissociation along Path A involves LA passing through the open gate formed by the $\alpha 1$ and $\alpha 3$ helices. The opening of the gate is measured by the distance between E340 on $\alpha 1$ and A372 on $\alpha 3$. (B) Ligand dissociation along Path B occurs when $\alpha 1$ and $\alpha 3$ move closer, resulting a closed gate and allowing the ligand to move between $\alpha 1$ and β -sheets. The distance between E340 and A372 for a Path A (C) and Path B (D) changes over the simulation time.

<https://doi.org/10.1371/journal.pcbi.1011955.g003>

more stability. As anticipated, the dissociation of the ligand along either Path A or Path B results in a significant enhancement in the fluctuation of all residues. After LA fully dissociated, the RMSF values reduce to the level that is similar to those observed in the fully bound state. Additionally, the RBM region, which houses crucial residues interacting with the ACE2 receptor during viral entry [10], was the most flexible region during LA dissociation.

Given the importance of $\alpha 1$ and $\alpha 3$ during ligand dissociation, we conducted further investigations to explore their significance in the process of LA unbinding. The $\alpha 1$ and $\alpha 3$ helices of the RBD function as a gate that undergoes conformational changes to accommodate the movement of the ligand unbinding from the pocket. In Path A, the $\alpha 1$ and $\alpha 3$ helices move apart from each other, creating a space that enables the ligand to exit the pocket between them. (Fig 3A). On the other hand, Path B requires $\alpha 1$ to move closer to $\alpha 3$, permitting the ligand to leave through the pocket between $\alpha 1$ and the β -sheet (Fig 3B). To quantitatively analyze these structural changes and the gate controlled by $\alpha 1$ and $\alpha 3$, we measured the distance between residues E340 of $\alpha 1$ and A372 of $\alpha 3$. We defined the gate open and closed states by distances of 18.21 Å (PDB ID 6VXX [47]) and 13.31 Å (PDB ID 6ZB5 [23]), respectively, according to the investigation of multiple spike structures from the PDB (S4 Table). Throughout the ligand unbinding trajectory, the gate distance fluctuates between the open and closed states.

Specifically, for ligands following Path A, the gate must open to a distance greater than 18.21 Å to enable the ligand to pass through (Figs 3C and S6). Conversely, Path B does not necessitate the gate to fully close to 13.31 Å, although the gate distance below 15.0 Å was required for the ligand to exit the binding pocket (Figs 3D and S6). These findings emphasize the dynamic nature of the gate and its role in facilitating the ligand dissociation process along both pathways.

Free energy profiles of LA unbinding from the spike glycoprotein

Utilizing the ligand RMSD and gate distance as reaction coordinates, we constructed a potential of mean force (PMF) diagram (Fig 4) to reveal the energetic change of LA dissociation. The PMF analysis revealed five local energy minimum states: the initial fully bound state and four intermediate states. During the ligand unbinding along Path A, the ligand transitions from the bound state to intermediate I-1A, then to intermediate I-2A, and ultimately becomes unbound. Conversely, for Path B, the ligand moves from bound to I-1B, then to I-2B, and finally becoming unbound.

Initially, in the ligand-bound state, LA is fully buried in the binding pocket formed by $\alpha 1$, $\alpha 3$, and β -sheet, while the gate distance is approximately 14.5 Å. As the ligand begins to dissociate, the polar interactions between LA and the residues of the RBD break, and the carboxyl headgroup of LA starts to interact with the surrounding water molecules outside the binding pocket, resulting in intermediate states I-1A and I-1B. During these states, three out of eight trajectories demonstrate that the ligand can transition between intermediate states I-1A and I-1B, as the gate swings open and closed, with the gate distance ranging from 12.5 Å to 17.5 Å. Moreover, the ligand can rebind to the binding pocket by forming hydrophobic interactions with the residues on the β -sheet close to $\alpha 1$ and $\alpha 3$. Subsequently, the ligand will continue to move farther away from the center of the pocket. Along Path A, the hydrophobic tail of LA is located between $\alpha 1$ and $\alpha 3$, forming nonpolar interactions with the α -helices, and the gate distance is around 17.5 Å, corresponding to intermediate state I-2A. On the other hand, along Path B, LA is positioned between $\alpha 1$ and the β -sheet, with the gate distance closing to

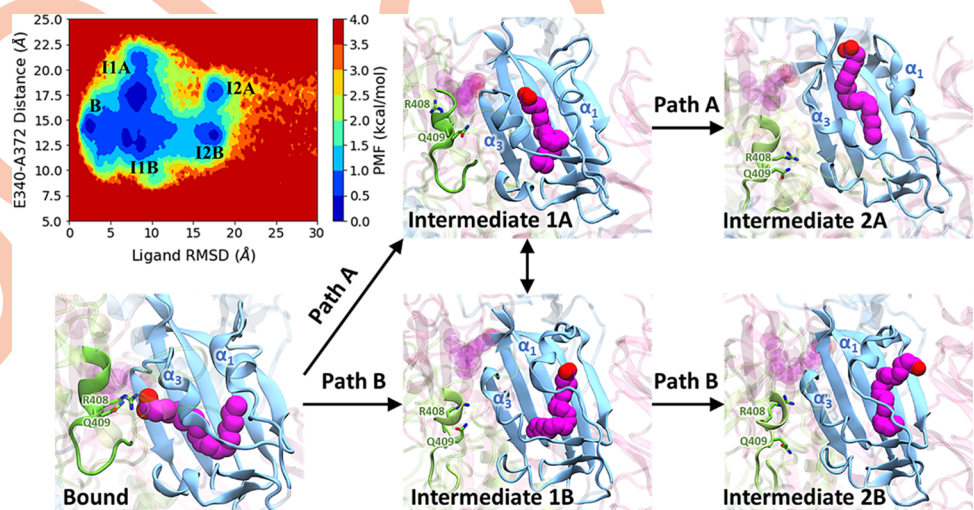


Fig 4. PMF profile of LA dissociation. The 2D PMF plot illustrates ligand RMSD and gate distance changes, indicating the intermediate conformations during LA unbinding processes. Arrows show the LA snapshots, starting in the fully bound state and moving along either Path A or Path B. A dual arrow indicates the transition of LA between dissociation paths at intermediate stages I-1A and I-1B.

<https://doi.org/10.1371/journal.pcbi.1011955.g004>

approximately 13.0 Å, denoted as intermediate state I-2B. Once LA reaches either state I-2A or I-2B, it eventually dissociates from the FFA binding pocket. After dissociation, the gate distance exhibits fluctuations within the range of 10–20 Å, while the ligand RMSD can reach up to 200 Å, consistently remaining greater than 30 Å (S7 Fig).

N-Glycan at N343 modulates the LA unbinding pathways

The glycan at N343 located on the α_1 helix plays a crucial role in determining the pathways of ligand unbinding. Our findings reveal a synchronous motion between the N343 glycan and the α_1 helix. The dihedral angle shown in Fig 5C experiences significant changes as the ligand dissociates from the spike along different pathways. In Path A, the glycan rotates away from the pocket between α_1 and α_3 (Fig 5A), resulting in the fluctuation of dihedral angle between 200 and 280 degrees (Fig 5C). Conversely, during the dissociation of LA along Path B, the α_1 helix swings towards α_3 , effectively closing the gate, while the glycan also moves towards α_3 , thereby hindering the pocket between α_1 and α_3 and impeding path A (Fig 5B). Throughout this trajectory, the dihedral angle fluctuates between 100 and 180 degrees (Fig 5D).

Additionally, the orientation of the glycan at N343 also plays a role in influencing the dissociation of LA by engaging in interactions with other residues or glycans within the spike glycoprotein. Our analysis of LA unbinding trajectories revealed two predominant states of the N343 glycan. In one state, the glycan shifts toward the N-terminal domain (NTD), while in the

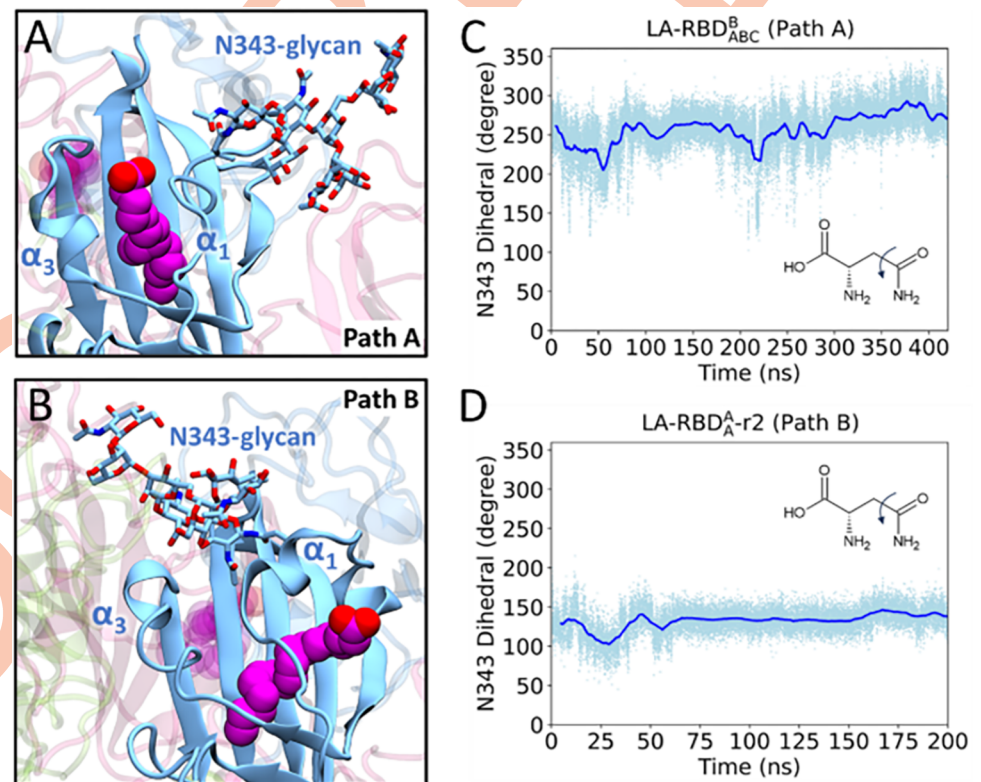


Fig 5. Dynamics of the glycan at N343. (A) In the case of dissociation along Path A, the N343 glycan extends into the solvent, moving away from the α_1 and α_3 helices. This allows LA to traverse through the open gate created by the α_1 and α_3 helices. (B) Conversely, during dissociation along Path B, the N343 glycan moves towards the α_1 and α_3 helices, restricting LA from progressing through Path A. Measurement of the dihedral angle of residue N343 is shown for Path A (C) and Path B (D), spanning the simulation duration.

<https://doi.org/10.1371/journal.pcbi.1011955.g005>

other state, it moves closer to the RBM. When the glycan adopts the former orientation, its outward positioning exposes the region between $\alpha 1$ and $\alpha 3$, which potentially facilitates the movement of LA along Path A. In this configuration, the glycan extends further into the solvent and establishes interactions with two glycans bound to residues N122 and N165 (S8A Fig). Conversely, in the latter state, the glycan interacts with residues within the RBM, including K417, Y453, L455, F486, N487, Y489, Q493, Q498, N501, and Y505 (S8B and S8D Fig). This interaction pattern constrains the adjacent RBD conformation and enhances the accessibility of LA dissociation along Path B.

Discussion

Understanding the intricacies of ligand dissociation processes remains paramount in the pursuit of designing novel inhibitors. Conventional MD simulations are often insufficient in exploring complete ligand dissociation events given that their typical sampling timescales fall within the range of nanoseconds to microseconds [48,49]. This timeframe is not able to fully capture unbinding processes, which usually extend from microseconds to seconds [50]. While several enhanced sampling methods, such as umbrella sampling [51,52], metadynamics [53,54], adaptive biasing force [55], and steered MD [56], are employed to investigate atomistic details of ligand unbinding, they commonly rely on predefined reaction coordinates. In this study, we applied LiGaMD [30] to investigate ligand dissociation processes without introducing biased results from preselected reaction coordinates. Specifically, the method achieves acceleration by enhancing the nonbonded interactions within the ligand-binding site, enabling the sampling of elusive LA unbinding events [30]. Our findings uncover two distinct LA dissociation pathways and their corresponding intermediate states. Moreover, our analysis highlights the significance of the $\alpha 1$ and $\alpha 3$ gate and the N343 glycan in regulating the preferred LA unbinding routes.

While LiGaMD enables the sampling of multiple ligand unbinding events from the spike glycoprotein, obtaining a reliable PMF profile through post-LiGaMD analysis remains challenging. In this study, we conducted over 100 simulations involving systems with one, two, or three bound ligands, accumulating more than 50 μ s of simulation time. However, only 8 complete ligand dissociation events were captured, resulting in a success rate of less than 8%. The small boost potential, such as 10–50 kcal/mol, did not enable us to sample the ligand dissociation within a 1 μ s simulation. Given the low rate of capturing ligand unbinding, we had to increase the boost potential to a level where dissociation events became observable. However, this adjustment necessitated a boost level that precluded accurate PMF reweighting. This trade-off was essential for capturing the dynamics of ligand unbinding while recognizing the limitations imposed by the need for enhanced sampling. Instead, we present the unweighted PMF of key reaction coordinates. While this unweighted PMF may not precisely depict the energetic changes of intermediates during ligand unbinding, it still represents the metastates within distinct unbinding pathways, as well as ligand-protein interactions at various stages.

Conformational selection and induced fit are two broadly recognized mechanisms in ligand binding and unbinding [57]. Conformational selection hypothesizes a protein with various conformational states, and ligand binding stabilizes one of these pre-existing states. Conversely, induced fit suggests a protein undergoes conformational changes upon ligand binding to optimize the interaction and create a complementary binding site. In our analysis of eight LA unbinding trajectories, we presented an even distribution of LA dissociation between Path A and Path B, which depended on two key factors: the gate distance between $\alpha 1$ and $\alpha 3$ and the position of the N343 glycan. Specifically, when the gate opens and the N343 glycan moves towards the RBD, LA dissociates along Path A; otherwise, it follows Path B. Regarding the gate

distance, we noted a wide range of gate distances across various available spike protein structures, as outlined in [S4 Table](#), indicating the diversity of pre-existing gate distances within the spike. Moreover, our simulations of a holo spike glycoprotein revealed significant N343 glycan fluctuations throughout the simulation. During these fluctuations, the glycan could either point towards the adjacent RBM or the NTD. Hence, our findings strongly indicate that LA dissociation predominantly aligns with conformational selection principles.

The role of K417 in the LA dissociation is critical, as it forms polar interactions with LA during ligand unbinding. K417 engages in the spike-ACE2 binding exclusively in the SARS-CoV-2 system and has been suggested as contributing to the increased binding affinity of the SARS-CoV-2 spike to ACE2 as compared to the SARS-CoV spike [10,58]. When the RBD is in the “up” conformation, K417 is exposed to the solvent and forms polar interactions with residue D30 of ACE2 receptor. However, in the presence of LA bound to the spike, the RBD adopts a “down” conformation, resulting in K417 orienting towards the FFA binding pocket. This interaction suggests that LA binding contributes to stabilizing the RBD in the down conformation, consequently impeding the spike-ACE2 binding.

The study of SARS-CoV-2 variants is crucial due to their potential impact on transmissibility, vaccine effectiveness, and the development of therapeutics. While the mutation study is beyond the scope of this research, we still highlight potential mutation sites that could influence ligand diffusion. The K417N mutation, present in Beta, Delta-plus, and Omicron variants, along with the R408S mutation specific to Omicron subvariants BA.2 and BA.4, may potentially destabilize the initial bound state and facilitate LA unbinding. These mutations affect residues involved in the polar anchor that initially binds the polar LA headgroup [59]. Additionally, mutations on the $\alpha 1$ or $\alpha 3$ helices, prevalent in Omicron variants, could influence the gate dynamics that is critical for the selection of Path A or Path B during LA dissociation. Mutations within the RBM of the Alpha, Delta, and Omicron variants may affect ACE2 binding affinities, indirectly changing LA dissociation through modifications in the RBD dynamics. The D614G mutation found in the Beta, Delta, and Omicron variants could affect the allosteric motion of the FFA binding pocket, potentially influencing the unbinding process of LA and similar ligands [59,29]. Further research will be necessary to clarify the specific effects of mutations from various variants on ligand diffusion.

N-glycans are recognized for their impact on the spike glycoprotein. For example, N165 and N234 glycans have been found to enhance accessibility for spike-ACE2 binding [60,25], while N282, N331, and N343 glycans provide shielding over the RBD in its down conformation [60,61,47]. Additionally, N616 and N1134 glycans have been identified as contributors to reduced spike glycoprotein stability [60,62], while N717, N801, and N1074 glycans are associated with decreased viral infectivity [60,63]. Our study further elucidates the critical roles played by N122, N165, and N343 glycans in the process of ligand dissociation. When LA dissociates through Path A, interactions between N343 glycan and N122 and N165 glycans were found. When LA dissociates through Path B, N343 glycan interacts with a set of residues as depicted in [S8B and S8D Fig](#), which are notably involved in the spike-ACE2 binding process. MD simulations performed by *Toelzer et al.* [24] also revealed hotspots of LA interactions on the RBD surface, including the RBM and the glycosylated residue N343 located near the pocket entrance. The presence of LA restrains the motion of the N165 glycan and the crucial residues that potentially form interactions with ACE2, suggesting that developing ligands with the capability to modulate glycan motions is a promising therapeutic strategy for addressing COVID-19 infection.

In our study, we found different LA dissociation behaviors in single-bound systems depending on the specific binding site. LAs bound to RBD_A exhibited rapid dissociation, requiring less applied boost in the simulation. Conversely, LA unbinding from RBD_B and

RBD_C took longer and required higher boost levels. Moreover, compared to systems with a single bound ligand, those with multiple bound ligands demonstrated prolonged dissociation times and complex trajectories. These trajectories often involved multiple ligand flips and rebinding events throughout the unbinding process. Due to the limited number of available LA unbinding trajectories for analysis, the influence of the number of bound ligands on dissociation time and pathways remains unclear. Future efforts will focus on extensive sampling to collect enough successful unbinding events, enabling robust statistical analysis of ligand dissociation rates and a thorough exploration of how different ligand environments impact the dissociation process. Additionally, although the spike protein comprises three identical chains, glycosylation patterns vary across each chain. Investigating the influence of distinct glycan attachments on each chain will be crucial in understanding how different glycans contribute to preferred ligand dissociation pathways. Given current computational limitations, our simulations only encompassed the spike head. Future studies of the complete spike system, including the viral membrane and molecular crowding, will provide a biologically realistic environment for studying ligand diffusion. Expanding our investigation to include other ligands alongside LA will further enhance our understanding of ligand diffusion dynamics of the spike glycoprotein.

In summary, we applied LiGaMD simulations to reveal the dynamics of ligand dissociation from the SARS-CoV-2 spike glycoprotein. Through extensive simulations, we identified two distinct ligand dissociation pathways, underscoring the critical role of the gate formed by $\alpha 1$ and $\alpha 3$ helices in the RBD, as well as the glycan attached to N343, in determining the preferred dissociation routes. These crucial insights significantly enhance our understanding of the ligand unbinding process from the spike glycoprotein, with implications for drug discovery in the battle against COVID-19. The fundamental knowledge learned from this system can also be applied to the treatment of influenza or other viruses.

Supporting information

S1 Movie. Ligand dissociation along Path A. The spike trimer, comprising Chain A (blue), Chain B (pink), and Chain C (green). Initially, the structure is shown in the simulation water box with glycans in the 3D-SNFG cartoon representation. The LA molecule is shown in magenta, the glycan is represented in bond form. (MP4)

S2 Movie. Ligand dissociation along Path B. The spike trimer, comprising Chain A (blue), Chain B (pink), and Chain C (green). Initially, the structure is shown in the simulation water box with glycans in the 3D-SNFG cartoon representation. The LA molecule is shown in magenta, the glycan is represented in bond form. (MP4)

S1 Text. Ligand Gaussian accelerated molecular dynamics (LiGaMD). (DOCX)

S2 Text. LiGaMD simulation input parameters. (DOCX)

S1 Table. Glycan attachments on the spike glycoprotein. (DOCX)

S2 Table. Summary of simulation protocols and system setup. This table outlines the detailed steps involved in system preparation and LiGaMD simulation for spike glycoprotein systems. The minimization processes were carried out in stages including the minimization of

protein hydrogen atoms, protein backbone, protein, ligand, and glycans. Followed by the addition of a water box, minimization of water molecules, and then minimization of the entire system. After executing water equilibration, the systems were gradually equilibrated at 50, 100, 150, 200, 250 and 310 K. LiGaMD preparation was performed by first performing conventional MD to collect the boost parameters, then applying a fixed boost potential, and then adjusting the boost potential while updating the boost parameters. The LiGaMD production simulation is run with the final updated boost parameters for either 500 ns or until successful ligand dissociation.

(DOCX)

S3 Table. The number of water molecules within distances of 5, 6, 7, and 8 Å around LA at various stages along dissociation pathways A and B for model systems LA–RBD_{AC}^C and LA–RBD_{ABC}^C. Please refer to S2 Fig for each dissociation state.

(DOCX)

S4 Table. Structure list of receptor-binding domains (RBDs) of the spike protein obtained from Protein Data Bank (PDB). The C_α distance between E340 and A372 was measured. The RBDs that cannot be determined either up or down conformation are noted as N/A.

(DOCX)

S1 Fig. The structures and sequence of the SARS-CoV-2 spike glycoprotein. (A) Schematic of the spike protein primary structure for one chain: N-terminal domain (NTD, 16–291), receptor-binding domain (RBD, 319–541), receptor-binding motif (RBM, 438–506), furin cleavage site (S1/S2), fusion peptide (FP, 817–834), central helix (CH, 987–1034), connecting domain (CD, 1080–1140). Representative icons indicate N-glycans (blue and green) at N17, N61, N122, N165, N234, N282, N331, N343, N616, N709, N717, N801, N1074, N1098, and N1134. (B) Sequence and secondary structures of the spike RBD. Blue and green indicate α helices and β sheets, respectively.

(PDF)

S2 Fig. Movement of LA along two dissociation pathways. (A) LA is tightly bound in the FFA binding pocket. (B) LA undergoes a flipping motion, initiating the dissociation process. (C) LA moves along Path A. (D) LA moves along Path B. (E) LA fully dissociates from the binding pocket along Path A. (F) LA fully dissociates from the binding pocket along Path B.

(PDF)

S3 Fig. RMSD analysis of the full spike protein from eight dissociation trajectories. Green and orange indicate trajectories dissociating along Path A and B, respectively.

(PDF)

S4 Fig. RMSD analysis of the spike RBD in complex with a boosted LA for trajectories, LA–RBD_{ABC}^B (A) and LA–RBD_{AC}^A (B). The color highlights different LA states: fully bound (blue), traveling along Path A (green), and moving along Path B (orange).

(PDF)

S5 Fig. RMSF analysis of the spike RBD in complex with LA from eight dissociation simulations. Line color represents RMSF profiles derived from trajectories during fully bound state (blue), traversal of Path A (green), traversal of Path B (orange), and after dissociation of the ligand (black). The highlighted colors correspond to different secondary structures: blue represents α1 (residues 337–343) and α3 (residues 364–369) helices, green corresponds to each β-strand (β2 residues 354–358, β3 residues 376–379, β5 residues 395–403, β6 residues 431–437, and β11 residues 507–514), and pink denotes the receptor-binding motif (RBM, residues 438

to 506).
(PDF)

S6 Fig. The distance between E340 and A372 from eight dissociation simulations. It includes the open and closed distances as horizontal dashed lines, with the dissociation time indicated by a vertical dashed line.

(PDF)

S7 Fig. PMF profile of LA dissociation. The plot depicts the PMF changes with the ligand RMSD and the gate distance between E340 and A372, including the states after LA fully dissociates with the ligand RMSD exceeds 30 Å.

(PDF)

S8 Fig. Movement of N343-glycan during LA dissociation. As LA dissociates from Chain A (blue) along Path A, the N343-glycan interacts with the glycans on N122 and N165 on the NTD in Chain B (pink) (A). Conversely, during LA dissociation along Path B, the glycan interacts with the residues of the RBM in Chain C (green), as illustrated in (B), (C), and (D).

(PDF)

Author Contributions

Conceptualization: Jinan Wang, Yinglong Miao, Yu-ming M. Huang.

Data curation: Timothy Hasse, Esra Mantei, Rezvan Shahoei, Shristi Pawnikar.

Formal analysis: Timothy Hasse, Jinan Wang, Yinglong Miao, Yu-ming M. Huang.

Funding acquisition: Yu-ming M. Huang.

Project administration: Yu-ming M. Huang.

Supervision: Yinglong Miao, Yu-ming M. Huang.

Visualization: Timothy Hasse.

Writing – original draft: Timothy Hasse.

Writing – review & editing: Timothy Hasse, Yu-ming M. Huang.

References

1. Chan JFW, Yuan S, Kok KH, To KKW, Chu H, Yang J, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: A study of a family cluster. *The Lancet*. 2020 Feb; 395(10223):514–23. [https://doi.org/10.1016/S0140-6736\(20\)30154-9](https://doi.org/10.1016/S0140-6736(20)30154-9) PMID: 31986261
2. Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *The Lancet*. 2020 Feb; 395(10224):565–74. [https://doi.org/10.1016/S0140-6736\(20\)30251-8](https://doi.org/10.1016/S0140-6736(20)30251-8) PMID: 32007145
3. Jackson CB, Farzan M, Chen B, Choe H. Mechanisms of SARS-CoV-2 entry into cells. *Nat Rev Mol Cell Biol*. 2022 Jan; 23(1):3–20. <https://doi.org/10.1038/s41580-021-00418-x> PMID: 34611326
4. Ni W, Yang X, Yang D, Bao J, Li R, Xiao Y, et al. Role of angiotensin-converting enzyme 2 (ACE2) in COVID-19. *Crit Care*. 2020 Dec; 24(1):422. <https://doi.org/10.1186/s13054-020-03120-0> PMID: 32660650
5. Wang Q, Meng F, Xie Y, Wang W, Meng Y, Li L, et al. In silico discovery of small molecule modulators targeting the achilles' heel of SARS-CoV-2 spike protein. *ACS Cent Sci*. 2023 Feb 8; 9(2):252–65. <https://doi.org/10.1021/acscentsci.2c01190> PMID: 36844485
6. Bojadzic D, Alcazar O, Chen J, Chuang ST, Condor Capcha JM, Shehadeh LA, et al. Small-molecule inhibitors of the coronavirus spike: ACE2 protein–protein interaction as blockers of viral attachment and

- entry for SARS-CoV-2. *ACS Infect Dis*. 2021 Jun 11; 7(6):1519–34. <https://doi.org/10.1021/acsinfectdis.1c00070> PMID: 33979123
7. Zhang J, Xiao T, Cai Y, Chen B. Structure of SARS-CoV-2 spike protein. *Current Opinion in Virology*. 2021 Oct; 50:173–82. <https://doi.org/10.1016/j.coviro.2021.08.010> PMID: 34534731
 8. Huang Y, Yang C, Xu X feng, Xu W, Liu S wen. Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19. *Acta Pharmacol Sin*. 2020 Sep; 41(9):1141–9. <https://doi.org/10.1038/s41401-020-0485-4> PMID: 32747721
 9. Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh CL, Abiona O, et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*. 2020 Mar; 367:1260–3. <https://doi.org/10.1126/science.abb2507> PMID: 32075877
 10. Lan J, Ge J, Yu J, Shan S, Zhou H, Fan S, et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature*. 2020 May; 581(7807):215–20. <https://doi.org/10.1038/s41586-020-2180-5> PMID: 32225176
 11. Costello SM, Shoemaker SR, Hobbs HT, Nguyen AW, Hsieh CL, Maynard JA, et al. The SARS-CoV-2 spike reversibly samples an open-trimer conformation exposing novel epitopes. *Nat Struct Mol Biol*. 2022 Mar; 29(3):229–38. <https://doi.org/10.1038/s41594-022-00735-5> PMID: 35236990
 12. Ke Z, Oton J, Qu K, Cortese M, Zila V, McKeane L, et al. Structures and distributions of SARS-CoV-2 spike proteins on intact virions. *Nature*. 2020 Dec; 588(7838):498–502. <https://doi.org/10.1038/s41586-020-2665-2> PMID: 32805734
 13. Zhang Q, Zhao N, Meng X, Yu F, Yao X, Liu H. The prediction of protein–ligand unbinding for modern drug discovery. *Expert Opin Drug Discov*. 2022 Feb; 17(2):191–205. <https://doi.org/10.1080/17460441.2022.2002298> PMID: 34731059
 14. Salahudeen MS, Nishtala PS. An overview of pharmacodynamic modelling, ligand-binding approach and its application in clinical practice. *Saudi Pharm J*. 2017 Feb; 25(2):165–75. <https://doi.org/10.1016/j.jsps.2016.07.002> PMID: 28344466
 15. Borisov DV, Veselovsky AV. Ligand–receptor binding kinetics in drug design. *Biochem Moscow Suppl Ser B*. 2020 Jul; 14(3):228–40.
 16. IJzerman AP, Guo D. Drug–target association kinetics in drug discovery. *Trends in Biochem Sci*. 2019 Oct; 44(10):861–71. <https://doi.org/10.1016/j.tibs.2019.04.004> PMID: 31101454
 17. Costa CFS, Barbosa AJM, Dias AMGC, Roque ACA. Native, engineered and de novo designed ligands targeting the SARS-CoV-2 spike protein. *Biotechnology Advances*. 2022 Oct; 59:107986.
 18. Norman A, Franck C, Christie M, Hawkins PME, Patel K, Ashhurst AS, et al. Discovery of cyclic peptide ligands to the SARS-CoV-2 spike protein using mRNA display. *ACS Cent Sci*. 2021 Jun 23; 7(6):1001–8. <https://doi.org/10.1021/acscentsci.0c01708> PMID: 34230894
 19. Lee YK, Chang WC, Prakash E, Peng YJ, Tu ZJ, Lin CH, et al. Carbohydrate ligands for COVID-19 spike proteins. *Viruses*. 2022 Feb; 14(2):330. <https://doi.org/10.3390/v14020330> PMID: 35215921
 20. Goc A, Niedzwiecki A, Rath M. Polyunsaturated ω -3 fatty acids inhibit ACE2-controlled SARS-CoV-2 binding and cellular entry. *Sci Rep*. 2021 Mar; 11(1):5207.
 21. Hidalgo MA, Carretta MD, Burgos RA. Long chain fatty acids as modulators of immune cells function: Contribution of FFA1 and FFA4 receptors. *Front Physiol*. 2021 Jul; 12:668330. <https://doi.org/10.3389/fphys.2021.668330> PMID: 34276398
 22. Alvarez-Curto E, Milligan G. Metabolism meets immunity: The role of free fatty acid receptors in the immune system. *Biochem Pharmacol*. 2016 Mar; 114:3–13. <https://doi.org/10.1016/j.bcp.2016.03.017> PMID: 27002183
 23. Toelzer C, Gupta K, Yadav SKN, Borucu U, Davidson AD, Kavanagh Williamson M, et al. Free fatty acid binding pocket in the locked structure of SARS-CoV-2 spike protein. *Science*. 2020 Nov; 370(6517):725–30. <https://doi.org/10.1126/science.abd3255> PMID: 32958580
 24. Toelzer C, Gupta K, Yadav SKN, Hodgson L, Williamson MK, Buzas D, et al. The free fatty acid–binding pocket is a conserved hallmark in pathogenic β -coronavirus spike proteins from SARS-CoV to Omicron. *Sci Adv*. 2022 Nov; 8(47):eadc9179.
 25. Casalino L, Gaieb Z, Goldsmith JA, Hjorth CK, Dommer AC, Harbison AM, et al. Beyond shielding: The roles of glycans in the SARS-CoV-2 spike protein. *ACS Cent Sci*. 2020 Oct; 6(10):1722–34. <https://doi.org/10.1021/acscentsci.0c01056> PMID: 33140034
 26. Kapoor K, Chen T, Tajkhorshid E. Posttranslational modifications optimize the ability of SARS-CoV-2 spike for effective interaction with host cell receptors. *Proc Natl Acad Sci USA*. 2022 Jul; 119(28):e2119761119. <https://doi.org/10.1073/pnas.2119761119> PMID: 35737823
 27. Piplani S, Singh P, Petrovsky N, Winkler DA. Identifying SARS-CoV-2 drugs binding to the spike fatty acid binding pocket using in silico docking and molecular dynamics. *Int J Mol Sci*. 2023 Feb; 24(4):4192. <https://doi.org/10.3390/ijms24044192> PMID: 36835602

28. Shoemark DK, Colenso CK, Toelzer C, Gupta K, Sessions RB, Davidson AD, et al. Molecular simulations suggest vitamins, retinoids and steroids as ligands of the free fatty acid pocket of the SARS-CoV-2 spike protein. *Angew Chem Int Ed*. 2021 Jan; 60(13):7098–110.
29. Oliveira ASF, Shoemark DK, Avila Ibarra A, Davidson AD, Berger I, Schaffitzel C, et al. The fatty acid site is coupled to functional motifs in the SARS-CoV-2 spike protein and modulates spike allosteric behaviour. *Comput Struct Biotechnol J*. 2022 Jan; 20:139–47. <https://doi.org/10.1016/j.csbj.2021.12.011> PMID: 34934478
30. Miao Y, Bhattarai A, Wang J. Ligand gaussian accelerated molecular dynamics (LiGaMD): Characterization of ligand binding thermodynamics and kinetics. *J Chem Theory Comput*. 2020 Sep; 16(9):5526–47. <https://doi.org/10.1021/acs.jctc.0c00395> PMID: 32692556
31. Huang Y ming M. Multiscale computational study of ligand binding pathways: Case of p38 MAP kinase and its inhibitors. *Biophysical Journal*. 2021 Sep; 120(18):3881–92. <https://doi.org/10.1016/j.bpj.2021.08.026> PMID: 34453922
32. Dolinsky TJ, Czodrowski P, Li H, Nielsen JE, Jensen JH, Klebe G, et al. PDB2PQR: Expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Research*. 2007 May; 35(Web Server):W522–5. <https://doi.org/10.1093/nar/gkm276> PMID: 17488841
33. Park SJ, Lee J, Qi Y, Kern NR, Lee HS, Jo S, et al. CHARMM-GUI glycan modeler for modeling and simulation of carbohydrates and glycoconjugates. *Glycobiol*. 2019 Apr; 29(4):320–31. <https://doi.org/10.1093/glycob/cwz003> PMID: 30689864
34. Park SJ, Lee J, Patel DS, Ma H, Lee HS, Jo S, et al. Glycan Reader is improved to recognize most sugar types and chemical modifications in the Protein Data Bank. Valencia A, editor. *Bioinform*. 2017 Oct; 33(19):3051–7. <https://doi.org/10.1093/bioinformatics/btx358> PMID: 28582506
35. Salomon-Ferrer R, Case DA, Walker RC. An overview of the AMBER biomolecular simulation package: AMBER biomolecular simulation package. *WIREs Comput Mol Sci*. 2013 Mar; 3(2):198–210.
36. Tian C, Kasavajhala K, Belfon KAA, Raguette L, Huang H, Miguels AN, et al. ff19SB: Amino-acid-specific protein backbone parameters trained against quantum mechanics energy surfaces in solution. *J Chem Theory Comput*. 2020 Jan; 16(1):528–52. <https://doi.org/10.1021/acs.jctc.9b00591> PMID: 31714766
37. Kirschner KN, Yongye AB, Tschampel SM, González-Outeiriño J, Daniels CR, Foley BL, et al. GLY-CAM06: A generalizable biomolecular force field. *J Comput Chem*. 2008 Mar; 29(4):622–55.
38. He X, Man VH, Yang W, Lee TS, Wang J. A fast and high-quality charge model for the next generation general AMBER force field. *J Chem Phys*. 2020 Sep; 153(11):114502. <https://doi.org/10.1063/5.0019056> PMID: 32962378
39. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B*. 1998 Apr; 102(18):3586–616. <https://doi.org/10.1021/jp973084f> PMID: 24889800
40. Liu J, Li D, Liu X. A simple and accurate algorithm for path integral molecular dynamics with the langevin thermostat. *J Chem Phys*. 2016 Jul; 145(2):024103. <https://doi.org/10.1063/1.4954990> PMID: 27421393
41. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh ewald method. *J Chem Phys*. 1995 Nov; 103(19):8577–93.
42. Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys*. 1977 Mar; 23(3):327–41.
43. Humphrey W, Dalke A, Schulten K. VMD: Visual molecular dynamics. *J Mol Graph*. 1996 Feb; 14(1):33–8. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5) PMID: 8744570
44. Roe DR, Cheatham TE. PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. *J Chem Theory Comput*. 2013 Jul; 9(7):3084–95. <https://doi.org/10.1021/ct400341p> PMID: 26583988
45. Miao Y, Sinko W, Pierce L, Bucher D, Walker RC, McCammon JA. Improved Reweighting of Accelerated Molecular Dynamics Simulations for Free Energy Calculation. *J Chem Theory Comput*. 2014 Jul; 10(7):2677–89. <https://doi.org/10.1021/ct500090q> PMID: 25061441
46. Watanabe Y, Allen JD, Wrapp D, McLellan JS, Crispin M. Site-specific glycan analysis of the SARS-CoV-2 spike. *Science*. 2020 Jul; 369(6501):330–3. <https://doi.org/10.1126/science.abb9983> PMID: 32366695
47. Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, Veesele D. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell*. 2020 Apr; 181(2):281–92. <https://doi.org/10.1016/j.cell.2020.02.058> PMID: 32155444
48. Hollingsworth SA, Dror RO. Molecular dynamics simulation for all. *Neuron*. 2018 Sep; 99(6):1129–43. <https://doi.org/10.1016/j.neuron.2018.08.011> PMID: 30236283

49. Klepeis JL, Lindorff-Larsen K, Dror RO, Shaw DE. Long-timescale molecular dynamics simulations of protein structure and function. *Curr Opin Struct Biol*. 2009 Apr; 19(2):120–7. <https://doi.org/10.1016/j.sbi.2009.03.004> PMID: 19361980
50. Tiwary P, Limongelli V, Salvalaglio M, Parrinello M. Kinetics of protein–ligand unbinding: Predicting pathways, rates, and rate-limiting steps. *Proc Natl Acad Sci USA*. 2015 Feb; 112(5):E386–91. <https://doi.org/10.1073/pnas.1424461112> PMID: 25605901
51. Kästner J. Umbrella sampling. *WIREs Comput Mol Sci*. 2011 Nov; 1(6):932–42.
52. Torrie GM, Valleau JP. Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. *J Comput Phys*. 1997 Feb; 23(2):187–99.
53. Bussi G, Laio A. Exploring complex free-energy landscapes by metadynamics. *Nat Rev Phys*. 2020 Jul; 2(4):200–12.
54. Laio A, Gervasio FL. Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. *Rep Prog Phys*. 2008 Dec; 71(12):126601–22.
55. Comer J, Gumbart JC, Hénin J, Lelièvre T, Pohorille A, Chipot C. The adaptive biasing force method: everything you always wanted to know but were afraid to ask. *J Phys Chem B*. 2015 Jan; 119(3):1129–51. <https://doi.org/10.1021/jp506633n> PMID: 25247823
56. Do PC, Lee EH, Le L. Steered molecular dynamics simulation in rational drug design. *J Chem Inf Model*. 2018 Jul; 58(8):1473–82. <https://doi.org/10.1021/acs.jcim.8b00261> PMID: 29975531
57. Vogt AD, Di Cera E. Conformational selection or induced fit? A critical appraisal of the kinetic mechanism. *Biochem*. 2012 Jul; 51(30):5894–902.
58. Han P, Li L, Liu S, Wang Q, Zhang D, Xu Z, et al. Receptor binding and complex structures of human ACE2 to spike RBD from omicron and delta SARS-CoV-2. *Cell*. 2022 Feb; 185(4):630–640. <https://doi.org/10.1016/j.cell.2022.01.001> PMID: 35093192
59. Oliveira ASF, Shoemark DK, Davidson AD, Berger I, Schaffitzel C, Mulholland AJ. SARS-CoV-2 spike variants differ in their allosteric responses to linoleic acid. Fu H, editor. *J Mol Cell Biol*. 2023 Mar; 15(3):mjad021.
60. Gong Y, Qin S, Dai L, Tian Z. The glycosylation in SARS-CoV-2 and its receptor ACE2. *Sig Transduct Target Ther*. 2021 Nov; 6(1):396. <https://doi.org/10.1038/s41392-021-00809-8> PMID: 34782609
61. Sztain T, Ahn SH, Bogetti AT, Casalino L, Goldsmith JA, Seitz E, et al. A glycan gate controls opening of the SARS-CoV-2 spike protein. *Nat Chem*. 2021 Oct; 13(10):963–8. <https://doi.org/10.1038/s41557-021-00758-3> PMID: 34413500
62. Teng S, Sobitan A, Rhoades R, Liu D, Tang Q. Systemic effects of missense mutations on SARS-CoV-2 spike glycoprotein stability and receptor-binding affinity. *Brief Bioinform*. 2021 Mar; 22(2):1239–53. <https://doi.org/10.1093/bib/bbaa233> PMID: 33006605
63. Wang L, Wang L, Zhuang H. Profiling and characterization of SARS-CoV-2 mutants' infectivity and antigenicity. *Sig Transduct Target Ther*. 2020 Sep; 5(1):185. <https://doi.org/10.1038/s41392-020-00302-8> PMID: 32883949