



OPEN ACCESS

EDITED BY

Shruthi Viswanath,
National Centre for Biological Sciences, India

REVIEWED BY

Rajarshi Roy,
Purdue University, United States
Sree Ganesh Balasubramani,
University of California, San Francisco,
United States

*CORRESPONDENCE

Yuefan Deng,
✉ yuefan.deng@stonybrook.edu

RECEIVED 13 December 2024

ACCEPTED 20 January 2025

PUBLISHED 11 February 2025

CITATION

Niu Z, Kementzidis G, Rafailovich M, Simon M,
Papadopoulos E, Aktas BH and Deng Y (2025)
Simulations of pH and thermal effects on
SARS-CoV-2 spike glycoprotein.
Front. Mol. Biosci. 12:1545041.
doi: 10.3389/fmolb.2025.1545041

COPYRIGHT

© 2025 Niu, Kementzidis, Rafailovich, Simon,
Papadopoulos, Aktas and Deng. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with
these terms.

Simulations of pH and thermal effects on SARS-CoV-2 spike glycoprotein

Ziyuan Niu¹, Georgios Kementzidis¹, Miriam Rafailovich²,
Marcia Simon³, Evangelos Papadopoulos⁴, Bertal H. Aktas⁴ and
Yuefan Deng^{1*}

¹Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, NY, United States, ²Department of Materials Science and Chemical Engineering, Stony Brook University, Stony Brook, NY, United States, ³Department of Oral Biology and Pathology, Stony Brook University, Stony Brook, NY, United States, ⁴Division of Hematology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, United States

We performed triplicate and long-time all-atom molecular dynamics simulations to investigate the structures and dynamics of the SARS-CoV-2 spike glycoprotein (S-protein) for a broad range of pH = 1 through 11 and temperatures of 3°C through 75°C. This study elucidates the complex interplay between pH and thermal effects on S-protein structures, with implications for its behavior under diverse conditions, and identifies the RBD as a primary region of the structural deviations. We found: 1) Structural deviations in the S-protein backbone at pH = 1 are 210% greater than those at pH = 7 at 75°C, with most of the deviations appearing in the receptor-binding domain (RBD). Smaller structural changes are observed at pH = 3 and 11. 2) The pH and thermal conditions impact on the protein structures: substantial acidic and basic conditions expand the protein's solvent exposure, while high heat contracts. This effect is primarily pH-driven at extreme acidity and thermo-driven at moderate pH. 3) The Gibbs free energy landscape reveals that pH as the main driver of structural changes. 4) The parametrized methods enable the predictions of the S-protein properties at any reasonable pH and thermal conditions without explicit MD simulations.

KEYWORDS

SARS-CoV-2, spike protein, AAMD, pH solvent, thermal conditions

1 Introduction

Environmental factors such as pH (Talley and Alexov, 2010) and thermal conditions (Tilton et al., 1992) play a crucial role in the stability of virus and protein structures. Strong acids and bases are frequently used for antiviral sterilization (Rowan et al., 2021), while research from various *in vitro* studies (Sturman et al., 1990; Darnell et al., 2004; Chu et al., 2006) indicates that viruses can maintain their infectivity across a broad spectrum of pH conditions. These findings underscore the need to examine the structural stability of viral proteins, like the SARS-CoV-2 spike protein, under extreme pH conditions. Moreover, the pH within early endosomes starts at pH = 6.3 and gradually drops to pH < 5 within lysosomes (Chen and Geiger, 2020). The gastric acid and the gastrointestinal tract are approximately pH = 1.5 to 7.4 [6, 7]. By examining the S-protein at pH levels as low as 1, we aim to capture the protein's behavior under extreme conditions that mimic acidic environments.

Moreover, exposure to heat is a widely adopted strategy for family antiviral disinfection, even though certain viruses can withstand extreme temperatures, both high and low (Leclercq et al., 2014; Van Doremalen et al., 2020). Prior *in vitro* research (Lamarre and Talbot, 1989; Chan et al., 2011; Islam et al., 2020) has demonstrated significant couplings of the pH and thermal conditions on the coronavirus properties. Revealing the intersection relation of SARS-CoV-2 across various pH and thermal conditions is crucial for mitigating virus transmission and comprehending viral biological functions. Recent research indicates that SARS-CoV-2 maintains stability at room temperature from pH = 3 through pH = 10 (Chin et al., 2020). The SARS-CoV-2 retains its stability for up to 2 weeks at 4°C and 24 h at room temperature but becomes inactive in 5 min at 70°C (Chin et al., 2020). However, the detailed pH and thermal coupling effect on SARS-CoV-2 requires further investigation.

Numerous works (Kandzia et al., 2019; Ostrowska et al., 2019; Leonard et al., 2021; Zhang and Huang, 2021) have adopted computational simulations to enable detailed examination of complex systems safely and inexpensively. The *in silico* studies and conventional laboratory experiments enhance mutually in gaining temporal and spatial resolutions and deep scientific insights. For example, the widely adopted all atomic MD reveals the protein structure details at the atomic scale that laboratory experiments miss (Frances-Monerris et al., 2020; Niu et al., 2022). Thus, the development of accurate computational models is crucial for understanding key factors that influence viral activity. Limited by the prohibitive computational costs of simulating an entire virus using all atomic MD, we focus on the virus' S-protein that is, albeit much smaller, an essential virus virulence determinant. The S-protein, situated on the coronavirus envelope, attaches to the host cell angiotensin-converting enzyme 2 (ACE2) during infection (He et al., 2020).

The structural alternations to the S-protein resulting from any conditions including pH or thermal effects may alter the infectivity of the hosting virus. Various *in silico* models suggest that the S-protein remains stable in the 0°C–30°C range, yet mixed opinions persist at higher temperatures, particularly, at 40°C–60°C (Rath and Kumar, 2020; Marti et al., 2021; Khan et al., 2022; Niu et al., 2022). The RBDs of the S-protein exhibit the greatest stability from pH = 6 to pH = 9 (Xie et al., 2022). Beyond their individual effects, the complex coupling effects at pH and thermal conditions are also evident in the structure and dynamics of S-protein. For instance, lowering the pH from 7.4 to 6.0 has been shown to reduce the thermal sensitivity of the S-protein structure (Edwards et al., 2021; Warwicker, 2021). Our study aims to uncover more intricate details in pivotal pH and thermal conditions.

The stability and structural conformations of the RBD are key factors in determining the transmissibility and infectivity of SARS-CoV-2. The RBD demonstrates a higher affinity for ACE2 binding compared to other coronaviruses (Costa et al., 2020; Shang et al., 2020b). The precise details of MD simulation allow existing studies to capture the contributions of single residues on the RBD to the interactions with ACE2 (Pipito et al., 2022). The S-protein has two main conformations: a closed form, with all its RBDs down, and an open form, where at least one RBD is up (Walls et al., 2020; Wrapp et al., 2020; Yan et al., 2021). Insights from recent studies (Ke et al., 2020; Turonova et al., 2020) reveal that the ratio of these

conformations varies with external conditions. Furthermore, the S-proteins elude antibody neutralization by fluctuating between the open and closed states (Gur et al., 2020; Zhou et al., 2020). The closed S-protein plays a crucial role in eluding the surveillance of the host immune system, thereby enhancing the spread of the virus (Frances-Monerris et al., 2020; Shang et al., 2020a). Both *in silico* (Cia et al., 2022) and *in vitro* (Ke et al., 2020; Juraszek et al., 2021) studies indicate that the S-protein trimer predominantly remains in a closed conformation before binding to ACE2 and initiating viral entry into host cells. Therefore, our study focuses on the closed states of the S-protein, 6VXX.pdb (Walls et al., 2020) which crucial for the viral immune evasion and persistence.

In consideration of their physiological and practical significance, we perform MD simulations for pH = 1 (extreme acid), 3 (acid kitchen cleaner), 5 (lysosomes), 7 (body pH), 9 (bar soap) and 11 (household ammonia). At each pH, temperatures were set at 3°C (cold supply chain), 20°C (room temperature), 37°C (body temperature), 56°C (critical temperature) and 75°C (high temperature). The simulated time for experiment reached 450 ns, a significant increase over the time scales attained in other studies with similar system sizes (Malaspina and Faraudo, 2020; Rath and Kumar, 2020; Marti et al., 2021; Coutinho et al., 2022; Khan et al., 2022; Sahihi and Faraudo, 2022). This extended duration brings the S-protein closer to equilibrium, offering a solid foundation for our analyses. A detailed comparative table summarizing all relevant information is provided in our previous work (Niu et al., 2024). We analyzed the of S-protein hierarchically (at protein, domain and residue bases) under 30 different conditions. The protein-based analysis reveals overall conformational changes, whereas examining them at the residue level offers more detailed insights into individual residue alterations, aiding in identifying factors that influence overall changes. Given the critical role of the RBD in the S-protein, we have also conducted separate analyses of RBD alterations across various environments which provides a deeper understanding of the S-protein's behavior. Furthermore, our parametrized root mean squared deviation (RMSD) method leverages the 30 sets of discrete MD simulation data to forecast changes in the outcome distribution across a broad and continuous spectrum of pH and thermal conditions.

2 Generation and analysis of trajectories

2.1 The *in silico* experiments

Based on recent *in vitro* and *in silico* studies (Chin et al., 2020; Xie et al., 2022), we design simulations for six pH levels and five temperatures, resulting at 30 unique conditions, with triplicates conducted for each. The choice of pH = 1 represents an extremely acidic environment, allowing us to examine potential structural changes or denaturation in the S-protein under highly acidic conditions. Examining the limits of S-protein stability at such extreme acidity provides valuable insights into its structural resilience and mechanisms of denaturation. By comparison, pH = 3 and 11, while less extreme, are included to assess whether critical points exist where notable structural changes occur. The mild pH = 5, 7, and 9 conditions permit evaluations of the S-protein

stability at conditions that resemble the human body physiological environment.

The 3°C should be tested as it represents the typical cold storage conditions for biological samples and vaccines. Understanding the structural stability of the SARS-CoV-2 spike protein at this temperature is crucial for ensuring the S-protein minimal conformational changes during storage and transport. At the room temperature of 20°C, the S-protein should be included to evaluate the protein's stability under standard environmental conditions, which are common in laboratory experiments or during short-term handling. At the normal human body temperature of 37°C, we could understand the functions during infection. The higher temperature of 56°C is considered a threshold for many viruses, SARS-CoV-2 included, to denature within 30 min. Therefore, it is essential to test this critical point to assess the potential for denature. Finally, the 75°C should be tested as it represents an extreme thermal condition, providing information on the denaturation dynamics.

The MD simulations adapt the open-source GROMACS (Abraham et al., 2015) that was coupled with the CHARMM36 force field (Brooks et al., 2009; Best et al., 2012). The initial structure of the S-protein is retrieved from the Protein Data Bank (6VXX.pdb) (Walls et al., 2020), with missing loops in the 6VXX structure reconstructed using Robetta (Baek et al., 2021). The S-protein, comprising 1273 residues per chain, is placed in the explicit solvent. We employ the SPC/E water model, known for its ability to represent solvent accurately, capturing the critical interactions between the protein and water (Mark and Nilsson, 2001; Takemura and Kitao, 2007), with an optimal balance between computational efficiency and accuracy. The simulation box, sized at $21 \times 21 \times 21 \text{ nm}^3$, is subject to periodic boundary conditions in all three Cartesian dimensions, and the density of the system is 1.01 g/cm^3 .

In all MD simulations, the energy minimization process was conducted using the steepest descent method. The system equilibration was achieved through both canonical (NVT) and Parrinello-Rahman pressure coupling (NPT) with a 2 fs time step size, and the production runs were maintained in the NPT ensemble. The triplicate experiments are performed at each condition to ensure the accuracy and reproducibility of the MD simulations. Each triplicated simulation was initiated by sampling the initial velocities of all atoms in the system according to Maxwell-Boltzmann distribution. For all 90 simulations (3 replicates across 30 unique conditions), the simulation time is extended to 450 ns to ensure sufficient sampling of the S-protein's conformational dynamics under each condition. On the shared AiMOS supercomputer (Hanson, 2019) configured with IBM POWER9 processors and NVIDIA Volta V100 GPUs, we perform our experiments on a sub-partition of 4 nodes with a running speed of approximately 50 ns/day.

Conventional pH MD simulations may introduce potential biases because the charge states of titratable residues are kept constant. Although constant pH MD simulations can dynamically update these charge states, they are computationally intensive (Buslaev et al., 2022), especially when multiple titratable sites interact both electrostatically and dynamically. Achieving effective sampling without a loss in computational speed poses an additional challenge, particularly for large proteins with hundreds or thousands of titratable sites undergoing major conformational changes (de Oliveira et al., 2022; Lasham et al., 2024). Therefore, we choose

to calculate the pKa values for all titratable residues in the S-protein and perform MD simulations with fixed protonation states, making it feasible to study proteins of this scale efficiently.

Empirical computational tools like PROPKA3 (Olsson et al., 2011) can accurately predict the pKa values of ionizable groups within proteins, even when hundreds of residues are involved. We control the protonation and deprotonation states of titratable residues, as estimated by the pKa predictor PROPKA3 (Olsson et al., 2011). To achieve charge neutrality, Cl^- and Na^+ ions are added to acidic and basic solvents, enhancing the accuracy of the protonation states (Chen and Shen, 2014). Detailed information on protein characteristics, including the number of atoms in the protein, the number of atoms in water, and the protein's net charge, is available in a previous publication (Niu et al., 2024). We maintain these parameters consistent with the previous study to ensure direct comparability and control across results.

2.2 Measurements and analysis of protein properties

The S-protein is organized as a homotrimer, with each of its three chains consisting of 1273 amino acids intertwined. Each chain includes three main regions: a signal peptide (residues 1–13), an S1 domain (S1D, residues 14–685) and an S2 domain (S2D, residues 686–1273). The S1 and S2 domains of the S-protein play a crucial role in infecting a host cell, with S1D binding to the host cell receptor (He et al., 2020) and largely composed of beta-sheets (Walls et al., 2016), includes the N-terminal domain (NTD) and the receptor-binding domain (RBD) (Xia et al., 2020). The S2D fusing with the virus (He et al., 2020) and primarily containing α -helices that span the membrane (Walls et al., 2016).

Data analysis reflects averages from three independent experiments. To reveal the structural changes in the S-protein, we conduct hierarchical analyses at the protein, domain and residue levels. For the full-length protein, we evaluate backbone RMSD, radius of gyration (R_g), solvent accessible surface area (SASA), interface area between chains, and the number of hydrogen bonds (H bonds) formed between the protein and water molecules (P-W) as well as inter-mainchain H bonds, collectively reflecting the overall structural changes in the S-protein. Additionally, we examine the Gibbs free energy (GFE) landscape and secondary structure alterations.

At the domain level, our analysis includes RMSD for specific domains such as the NTD, RBD and S2D. For residue level insights, we measure the root mean squared fluctuation (RMSF) for each individual residue.

2.2.1 Gibbs free energy landscape

The GFE landscape is calculated as

$$\Delta G(x_1, x_2) = -k_B T \ln P(x_1, x_2)$$

where k_B is the Boltzmann constant, T denotes temperature, and $P(x_1, x_2)$ is the normalized joint probability for the first two principal components (PC1 and PC2) which capture most structural variation in the carbon alpha trajectory. The covariance matrix for the carbon alpha coordinates is computed using trajectory

data from the last 200 ns, with alignment to the first frame to mitigate rotational and translational effects in the simulation. By diagonalizing the covariance matrix, the resulting eigenvalues and eigenvectors allowed for projection onto PC1 and PC2, enabling the creation of a two-dimensional free energy landscape plot of the conformational space (Maisuradze and Leitner, 2007).

2.2.2 Parametrized RMSD

Non-linear least square analysis is applied to fit the RMSD data for the full protein, NTD, RBD and S2D for a range of pH and thermal conditions. This allows us to fit these data to obtain a function of the RMSD's in two variables--pH and temperature--and several parameters. Many function forms can be candidates and, multiple trials and insight of chemistry lead us to the following fitting function for all four measures:

$$\text{RMSD}(T, p^*) = R_0 * (1 + \alpha_T T) * (e^{-\beta_a p^*} + e^{\beta_b p^*})$$

where T represents temperature and $p^* = \text{pH} - 7$. The coefficients R_0 , α_T , β_a and β_b are free parameters fitted to the data, capturing the effects of pH and thermal conditions on the RMSD. More specifically, the β_a represents the coefficient dominant under acidic conditions where $p^* < 0$, while β_b applies to basic conditions.

Thermal effects are modeled linearly because no extreme structural shifts were observed within the ns time scales for the tested temperature range. This model, allowing us to assess the individual effects, and their couplings, at the pH and thermal conditions, offers predictive insights for protein structures at the untested conditions.

The pH dependence is modeled by two exponential functions to depict the asymmetrical pH effects of the acidic and basic conditions on the protein. This form with the two fitting parameters β_a and β_b is flexible and sufficient to express such changes. From a chemical perspective, the relationship between pH and RMSD builds on the hydrogen ion concentration and its influence on protein conformation. RMSD measures structural deviations, which we relate to pH via $p^* = \text{pH} - 7$. Here, pH reflects the activity of hydrogen ions, expressed as:

$$\text{pH} \approx -\log_{10}[H^+] = -\frac{\ln[H^+]}{\ln 10}$$

To bridge the conceptual gap between pH and structural deviations we incorporate p^* and expand our chosen function:

$$\begin{aligned} e^{-\beta_a p^*} + e^{\beta_b p^*} &\approx e^{-\beta_a \left(-\frac{\ln[H^+]}{\ln 10} - 7 \right)} + e^{\beta_b \left(-\frac{\ln[H^+]}{\ln 10} - 7 \right)} \\ &= e^{7\beta_a} [H^+]^{\frac{\beta_a}{\ln 10}} + e^{-7\beta_b} \left(\frac{1}{[H^+]} \right)^{-\frac{\beta_b}{\ln 10}} \end{aligned}$$

The fitting parameters β_a and β_b express the effect of the concentrations of hydrogen ions on the molecule's configuration, going deeper into the causes of the changes in the structures of the S-protein.

3 Results

During MD simulations, the raw coordinates for all atoms were collected at 0.1 ns intervals and aligned by the protein backbone

at the center of the box. We evaluate the conformational stability of the S-protein across these 30 conditions, aiming to establish a foundational understanding of its structural dynamics at varying pH and thermal conditions.

3.1 The whole protein

Our protein-based analyses focus on residues 1–1162, which cover the bulbous head and part of the stalk region. This selection excludes the distal portion of the stalk to reduce the impact of its considerable fluctuations on our assessment of the S-protein's stability. The mean values from 250 ns to 450 ns, along with the standard errors from the triplicate simulations, are provided in [Supplementary Table S1](#) for the protein analyses. The time series data for the entire simulation period are shown in [Supplementary Figure S1](#).

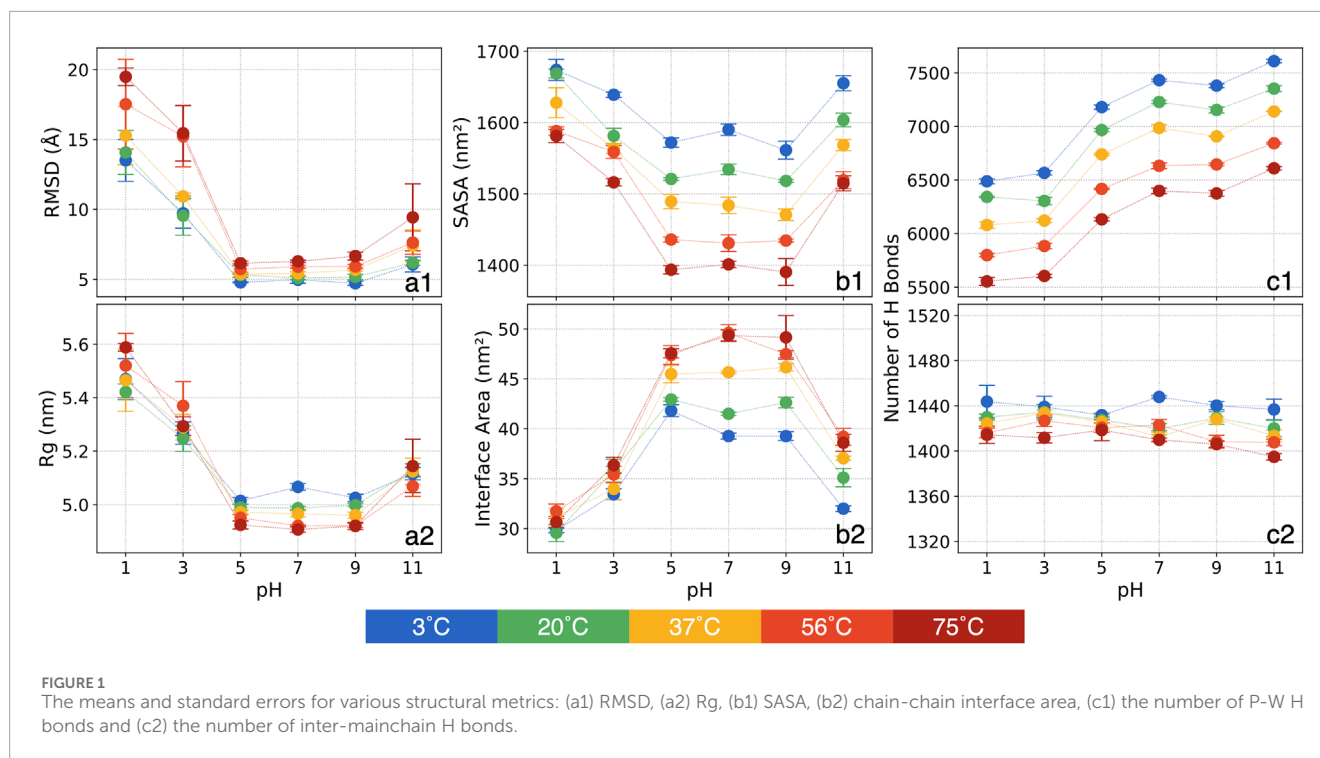
The RMSDs: We calculate RMSD by averaging the backbone atom coordinates at each timestep relative to their positions in the initial structure. The mean of RMSD values over the last 200 ns and the standard errors from the triplicate simulations are presented in [Figure 1a1](#).

In general, at a given temperature, the RMSD analysis reveals a high sensitivity at more acidic conditions of pH = 1 and 3. For pH = 1, in contrast to pH = 7 at the same temperature, RMSD increased by 170% at 3°C and by 210% at 75°C. At pH = 3, RMSD again separates into two temperature-dependent groups. For temperatures below 37°C, RMSD increases by roughly 90%–110%, while for temperatures above 56°C, it rises by approximately 150%. Conversely, at pH = 5 and 9, RMSD deviates minimally from that of pH = 7 regardless of temperatures, indicating relative stability. At pH = 11, RMSD shows less variation compared to acidic conditions, with the largest change occurring at 75°C, where RMSD increases by around 50%.

At a given pH condition, RMSD increases with temperature. This increase remains modest (within 2 Å) for pH = 5, 7, and 9. However, at pH = 1 and 3, the thermal effect on the S-protein intensifies considerably. Notably, at pH = 1, RMSD is 13.5 Å at 3°C and rises steadily with temperature, reaching 19.5 Å at 75°C. At pH = 3, RMSD exhibits a distinct separation into two groups with increasing temperature. For temperatures between 3°C and 37°C, RMSD is centered around 9.6 Å, but it escalates to 15.3 Å at 56°C and 75°C. At pH = 11, RMSD also rises with temperature but to a lesser extent than at acidic conditions, increasing from 6.1 Å to 9.4 Å between 3°C and 75°C.

The considerable RMSD variations at pH = 3 with rising temperatures highlight three distinct structural states for the 30 conditions: 1) RMSD as high as 13.5 Å is observed at pH = 1 and all temperatures and at pH = 3 for relatively higher temperatures, showing substantial structural deviations; 2) Intermediate RMSD at around 10 Å is observed at low temperatures for pH = 3 and at 75°C for pH = 11, indicating the S-protein undergoing a transitional phase; and 3) Low RMSD at around 5.5 Å is observed at pH = 5 through pH = 11 at all tested thermal conditions, revealing minimal structural deviations.

The Rg represents how atoms are distributed relative to the protein's center of mass, providing valuable insights into the protein's



overall size and structural configuration. A lower Rg value reflects a more compact protein structure, while a higher Rg indicates a more extended form. Shifts in Rg serve as important markers for alterations in the protein's tertiary structure. The mean of Rg values and standard errors over the last 200 ns of the triplicate simulations are shown in [Figure 1a2](#).

Our analysis reveals that, at a given temperature, Rg increases by approximately 10% at pH = 1 and 5% at pH = 3 compared to pH = 7, indicating a more expanded structure under highly acidic conditions. This expansion is likely due to the protonation of amino acid side chains, which disrupts the electrostatic interactions, leading to a looser protein structure. At basic conditions at pH = 11, Rg increases only slightly (around 2%), indicating a minor expansion.

Thermal effects on Rg, however, are less visible at any pH values. Overall, Rg decreases slightly (about 2%) with increasing temperature from 3°C to 75°C at each pH, suggesting a trend toward a more compact structure as temperature rises. This trend hints at a tendency of the S-protein to adopt a more condensed conformation at increasing temperatures, a result corroborated in subsequent analyses of SASA and interface area.

These observations suggest that the S-protein remains relatively compact at neutral and mildly basic conditions, while extreme acidic conditions promote expansion. Additionally, the protein tends to remain more extended at lower temperatures. Thus, our Rg analysis highlights distinct effects at pH and thermal conditions on the S-protein's conformation.

The SASA analysis offers more insights into the folding patterns and stability of the S-protein, particularly regarding its overall shape, whether more extended or contracted. The mean of SASA values over the final 200 ns, along with the standard errors from the triplicate simulations, are presented in [Figure 1b1](#).

At a constant temperature, SASA increases at extreme pH conditions. At pH = 1, SASA increases by approximately 85 nm² at 3°C, with this upward trend becoming more pronounced as temperature rises. By 75°C, SASA shows an increase of 180 nm² at pH = 1 compared to pH = 7. Similarly, pH = 3 and 11 display increased SASA relative to pH = 7, although the increase is less than at pH = 1. For pH = 5 and 9, SASA shows only minor deviations from that of pH = 7. The elevated SASA at pH = 1, 3 and 11 suggest greater structural deviation in specific domains of the S-protein, indicating an expansion relative to its structure at pH = 7. Further analyses will offer more detailed insights into the specific structural changes in these domains. When comparing different thermal conditions at a constant pH, SASA decreases as temperature increases. The SASA reduction is more pronounced, approximately 180 nm², within the moderate pH range, whereas it decreases by around 130 nm² at pH = 3 and 11. At pH = 1, SASA decreases by only 93 nm², indicating that thermal effects are impeded by the extreme pH conditions.

Overall, the analysis reveals two primary trends: SASA decreases with increasing temperature and tends to be higher under strongly acidic or basic conditions. This suggests that extreme pH promotes an extension of the S-protein, while high temperature promotes a more contracted state. Additionally, the results show couplings where the thermal effects diminish at more acidic conditions.

The chain-chain interface area represents the regions where the S-protein chains make contact and interact, playing a crucial role in stabilizing the quaternary structure of the trimeric S-proteins. The average interface areas between chains were calculated using the PDBEPIA [61] server and are displayed in [Figure 1b2](#).

In our analysis, the interface area decreases substantially under acidic and basic pH conditions at the same thermal conditions. While pH = 5 and 9 show relatively stable values close to those observed at pH = 7, pH = 11 exhibits a notable reduction of

approximately 8 nm² across all thermal conditions. A clear trend appears in acidic conditions: the interface area decreases more significantly at higher temperatures. Specifically, at pH = 1, the interface area averages around 30 nm² across all thermal conditions. Compared to pH = 7, the interface area at 3°C decreases by 10 nm², but this difference grows to 18 nm² at 75°C. A similar trend, though less pronounced, is observed at pH = 3.

Interestingly, at a constant pH, the interface area tends to increase as temperature rises, especially at pH > 5. For example, at pH = 7, the interface area increases by around 10.3 nm² from 3°C to 75°C, and at pH = 9 and 11, it increases by approximately 7.5 nm². At acidic conditions, however, this increase is more moderate: at pH = 5, it only increases by about 5.6 nm², and at pH = 1 and 3, the increase is less than 2 nm².

These findings collectively suggest that interchain interactions weaken under strongly acidic and basic pH, particularly at pH = 1, 3, and 11, which may result in structural changes or partial separation of the trimer into individual chains. Additionally, heating tends to increase the interface area, indicating a compaction effect as confirmed by SASA measurements, with chains moving closer together. An interesting intersectional effect is observed: at extreme pH conditions, pH appears to dominate, causing a smaller increase in interface area even at higher temperatures. In contrast, at moderate pH conditions, thermal effects dominate the variations of the interface area.

The H bonds: Experimental studies indicate that H bonds play a critical role in protein stability (Pace et al., 2014), rooted in the significance of H bonds in the shaping of the tertiary structure. In our analysis, H bonds were counted based on cutoff values for the angle and distance between donor and acceptor atoms. The mean of H bond counts for the final 200 ns and the standard errors from the triplicate simulations are shown in Figures 1c1, c2 for P-W and inter-mainchain H bonds, respectively.

At a constant temperature, the number of P-W H bonds tends to decrease at acidic conditions and increase at basic conditions. Compared to pH = 7, at pH = 1 and 3, around 900 P-W H bonds are broken. At pH = 5, approximately 200 P-W H bonds break. At pH = 9, the number of P-W H bond shows no noticeable change, whereas at pH = 11, around 200 additional P-W H bonds formed. Interestingly, while pH = 1 and 3 show a significant decrease in P-W H bonds relative to pH = 7, there is an increase in SASA at low pH. This indicates that as the protein unfolds, hydrophobic residues typically buried in the core may become more exposed, leading to increased SASA but competing with water molecules in forming the P-W H bond. Conversely, at pH = 11, the formation of additional P-W H bonds aligns with increased SASA, suggesting that more hydrophilic residues are interacting with the solvent.

When comparing different thermal conditions at a constant pH, a clear linear trend can be found, approximately 1,000 P-W H bonds are broken as it heats up from 3°C to 75°C. Compared to SASA, both measures decrease with increasing temperature, indicating a distinct structural response likely due to high temperatures causing a contraction of the protein. An interesting result in the P-W H bond analysis is that, unlike with other metrics, pH and thermal conditions appear to influence protein structural variations independently.

Across all tested pH values, no significant changes are observed in the counts of inter-mainchain H bonds at any given thermal

condition. Within the same pH, as temperature increased from 3°C to 75°C, the inter-mainchain H bonds broken by less than 2%, a minimal reduction. This finding suggests that mainchain interactions remain highly stable, even at extreme pH and high temperatures.

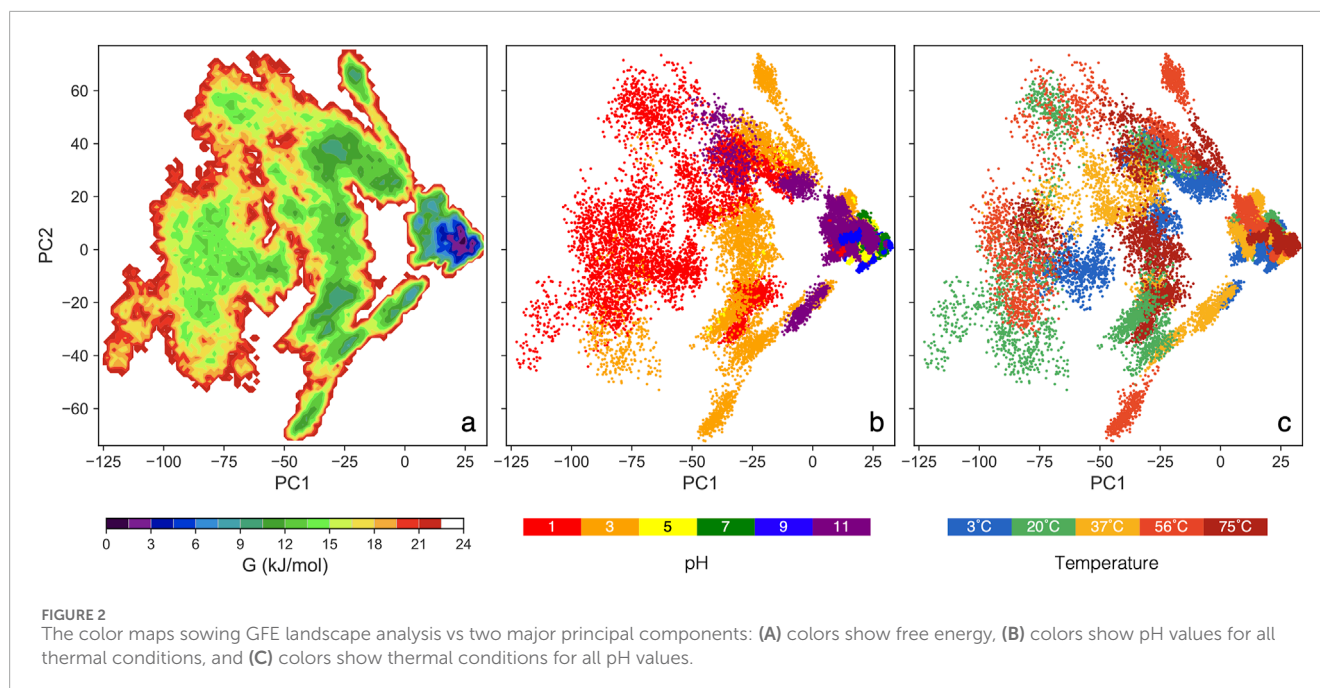
The GFE landscape: It is important to note that the GFE value is not energy measured from first principles but a statistical approximation based on the assumption that the equilibrium distribution of states follows the Boltzmann distribution (Papaleo et al., 2009). Each point on the plot represents a conformational state from the MD simulation, with smaller distances between points indicating structural similarity and denser regions suggesting stable protein conformations. The GFE landscape for our 90 simulations (spanning 30 conditions with triplicates) is presented in Figure 2. Each simulation includes data from 250 ns to 450 ns, concatenated to construct a single comprehensive covariance matrix. PC1 accounts for 42%, while PC2 accounts for 13% of the total variance across all conditions.

In Figure 2B, a clear trend is pH = 5, 7, and 9 cluster closely, aligning with a deep valley in Figure 2A, which indicates consistent and stable protein conformations. In contrast, pH = 3 and 11 form broader, shallower ensembles than the moderate pH range, though still deeper than pH = 1. At pH = 1, the protein spans a large, flat region in Figure 2A, suggesting the most extensive structural variation. The deep energy valleys at moderate pH suggest that the S-protein favors stable conformations within this range, with thermal condition having little impact on structural stability (Chin et al., 2020; Xie et al., 2022). These GFE landscapes demonstrate notable pH-dependent variations, highlighting pH as a key factor influencing the conformational dynamics of the S-protein.

Spatial distribution within the landscape also provides insights into structural similarity, as the relative positions and distances between ensembles reflect the degree of geometric resemblance (Becker, 1997). Trajectories at pH = 5, 7, and 9 project onto similar PC1 and PC2 values, clustering closely across all thermal conditions, which suggests that the protein adopts similar conformations within this moderate pH range. Conversely, the trajectories at pH = 1 are significantly separated from the moderate pH cluster, covering a broader range of conformations. Notably, pH = 3 and 11 occupy an intermediate region, serving as a transition between the extreme pH = 1 and the moderate pH range. In Figure 2C, no distinct trend emerges, reinforcing that pH conditions, rather than thermal, are the primary driver of structural changes in the S-protein.

The secondary structures: Define Secondary Structure of Proteins (DSSP) (Kabsch and Sander, 1983; Touw et al., 2015) provides quantitative insights by determining the number of residues involved in various secondary structural elements, offering a perspective on the structural features that contribute to protein stability and folding. Supplementary Table S2 presents the DSSP analysis results across our 30 simulation conditions. The Secondary Structure (SS) column shows the combined percentages of structural elements such as α -helix, β -sheet, β -bridge, 3_{10} -helix, π -helix and turns. These results suggest that the secondary structure of the S-protein remains relatively stable, maintaining around 60% integrity across diverse pH and thermal conditions, including the extreme combination of pH = 1°C and 75°C.

This stability suggests that the secondary structure of the protein is well-maintained under a range of pH and thermal conditions, at



least for the initial 450 ns of the simulation within our conditions. Furthermore, the DSSP and inter-mainchain H bond analyses across the 30 conditions consistently show secondary structural stability. We hypothesize that, in longer simulations, the secondary structure may become destabilized under more extreme conditions, such as highly acidic or high-temperature environments, like pH = 1°C and 75°C.

3.2 The domain-based analysis

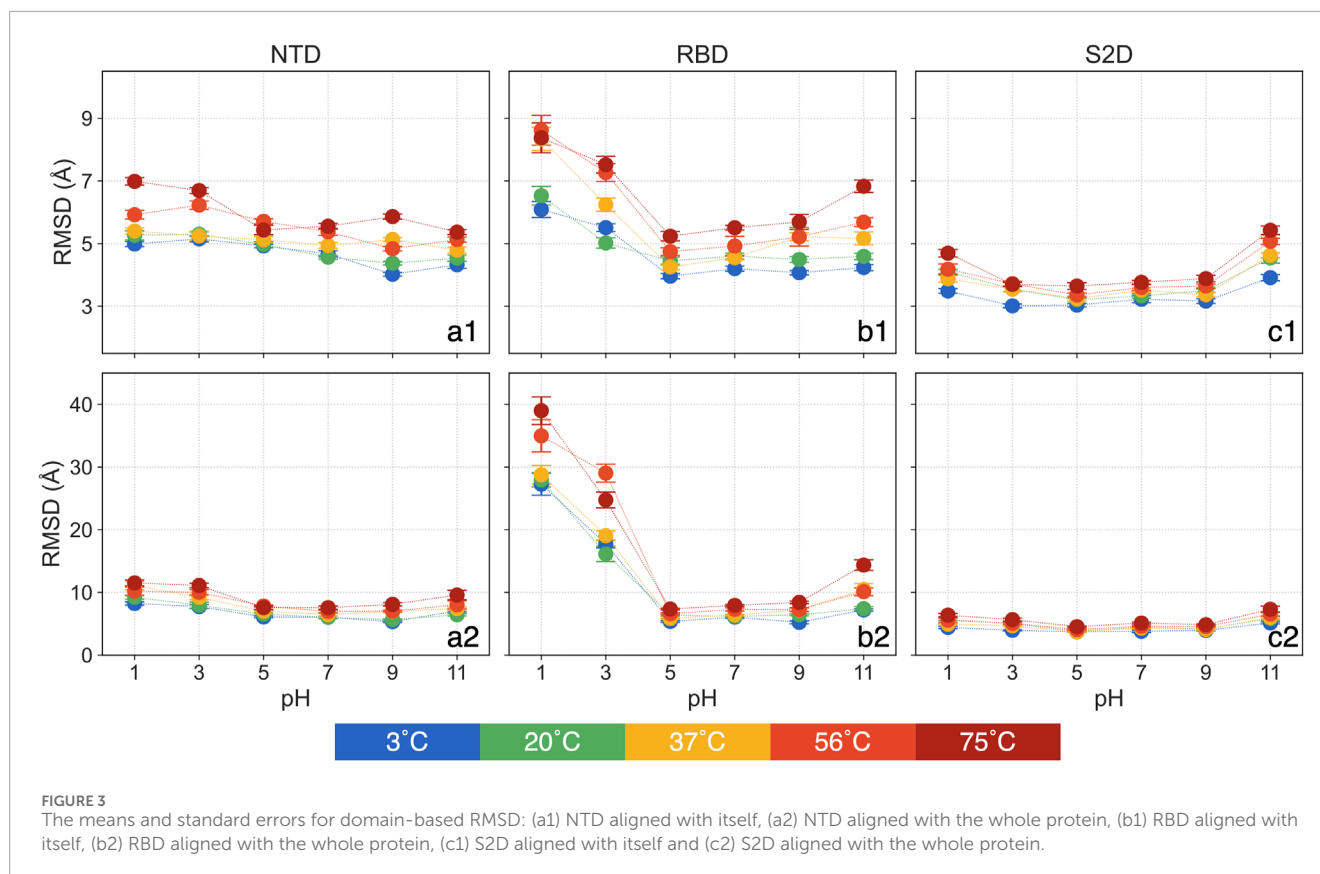
In examining domain-specific conformational changes, specifically in the NTD, RBD and S2D, we focus on the backbone RMSD to assess their relative stability and flexibility. The RMSD is calculated by aligning each domain both to itself and to the entire protein, allowing us to interpret two aspects of structural change: 1) The RMSD aligned to the domain itself reveals internal structural deviations, capturing changes independent of backbone translational and rotational movement in the simulation. 2) The RMSD aligned with the whole protein shows how each domain shifts relative to the entire structure, highlighting changes driven by domain translation and rotation. This targeted approach allows us to assess each domain's potential to act as a pH or temperature sensor, offering insights into the functional implications of structural adaptation. All backbone RMSD analyses in this section focus on the mean over the 250 ns–450 ns, with standard errors from the triplicate simulations shown in [Figure 3](#) and [Supplementary Table S3](#). In [Figure 3](#), the first row presents RMSD values aligned by each domain, while the second row shows RMSD values aligned by the whole protein. The time series data for the entire simulation period are shown in [Supplementary Figure S2](#).

The RMSD of the NTD, aligned with itself ([Figure 3a1](#)), does not display any clear trend across different pH values. Only at 75°C, the NTD RMSD show approximately a 25% increase for pH = 1 and 3

compared to pH = 7. At same pH, the NTD RMSD showing a slightly increase as it heats up from 3°C to 75°C, for example, it increases less than 17% (0.8 Å) at pH = 7. The NTD RMSD, at pH = 5 and 9 and aligned with the entire protein ([Figure 3a2](#)), shows no significant deviations from that of pH = 7, regardless of temperatures. However, at both acidic and basic conditions, NTD RMSD increases, with greater sensitivity observed in acidic environments. For instance, at pH = 1 and 3, RMSD is as high as 11.5 Å at 75°C, while at pH = 11, it is 9.6 Å compared to 7.6 Å at pH = 7. At a constant pH, the NTD RMSD slightly increases with rising temperature. In summary, the NTD remains relatively stable at the tested pH and thermal conditions with minor changes at higher temperatures. However, at acidic conditions, the NTD is more prone to rotation and translation.

The RMSD of the RBD, aligned with itself ([Figure 3b1](#)), reveals a clear intersection between pH and thermal effects. At low temperatures, only acidic conditions significantly affect the RBD, but as temperature increases, basic conditions (pH = 11) also begin to impact. Specifically, at 3°C and 20°C, RBD RMSD at pH = 11 shows no significant difference from pH = 7. However, at 37°C and 56°C, RBD RMSD at pH = 11 rises by about 15% compared to pH = 7, and by 24% at 75°C. The RBD RMSD remains stable for pH = 5 and 9 across all thermal conditions. At pH = 1, RBD RMSD increases by approximately 43% to around 6.3 Å compared to around 4.4 Å at pH = 7 at 3°C and 20°C, reaching 8.3 Å at 37°C but showing no additional change at higher temperatures.

For the RBD RMSD aligned with the whole protein ([Figure 3b2](#)), pH = 5 and 9 again show no significant deviations from pH = 7 across thermal conditions. The RBD RMSD, aligned with the protein, follows a similar trend to the self-aligned data, clustering into three groups: 7.3 Å at 3°C and 20°C, 10.2 Å at 37°C and 56°C, and 14.4 Å at 75°C. At extremely acidic conditions (pH = 1 and 3), RMSD rises dramatically. At pH = 1, RMSD increases to around 28 Å at 3°C to 37°C and further to 35 Å and 39 Å at 56°C and 75°C, respectively, with a similar but slightly lower pattern at pH = 3.



These results highlight three key observations: 1) Individual structural deviations in RBD are most pronounced at pH = 1 and 3 at any temperatures, showing effects only at high temperatures for pH = 11; 2) Structural changes at pH = 1 peaks at 56°C. 3) The RBD experiences considerable translation and rotation in highly acidic conditions (pH = 1 and 3), a result with possible implications on the binding to the ACE2 receptor and hence the infectivity of the virus at these conditions. This observation aligns with prior findings (Xie et al., 2022).

The RMSD of the S2D, whether aligned with itself (Figure 3c1) or with the entire protein (Figure 3c2), shows a distinct trend compared to the NTD and RBD. At pH = 3, 5, and 9, there are no significant deviations from pH = 7 across all thermal conditions, with only pH = 1 and 11 exhibiting an increase in RMSD relative to pH = 7. Interestingly, this increase is more pronounced at pH = 11, suggesting that the S2D demonstrates a high tolerance to acidic environments.

The parametrized RMSD: Our coefficients are listed in Table 1, along with the coefficient of determination (R^2), and the results are presented in Figure 4. In this analysis, β_a represents the dominant coefficient at acidic conditions, while β_b applies under basic conditions. We observe that for both overall protein RMSD and RBD RMSD, $\beta_a > \beta_b$, indicating that the structural changes of the S-protein at acidic pH (especially pH = 1 and 3) are more skewed. Conversely, the S2D shows the opposite trend, with $\beta_b > \beta_a$, consistent with the domain data for pH = 11, which shows greater changes than at acidic conditions like pH = 1. The NTD domain displays much smaller β_a and β_b values compared to other

TABLE 1 The coefficients for the protein and its individual domains.

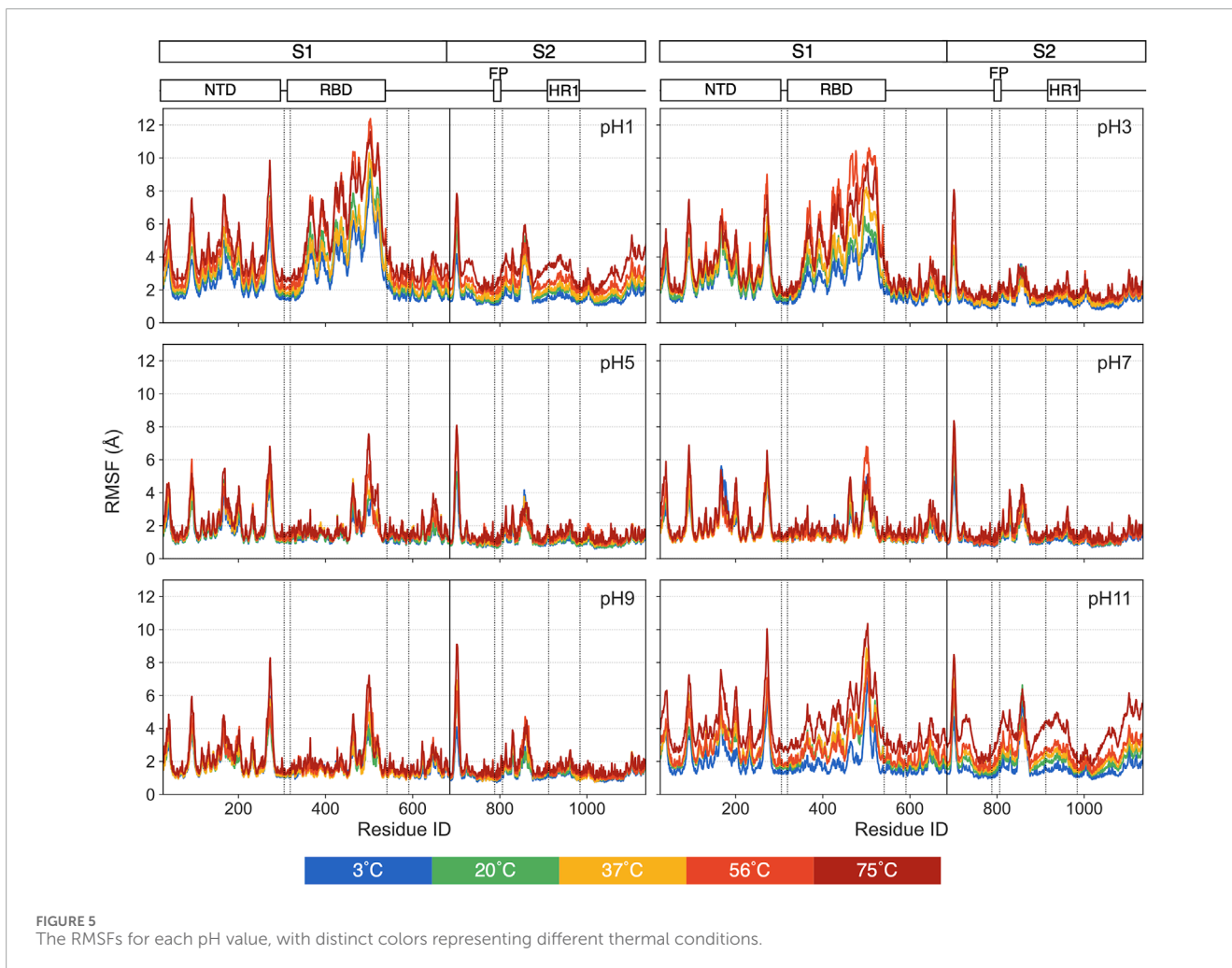
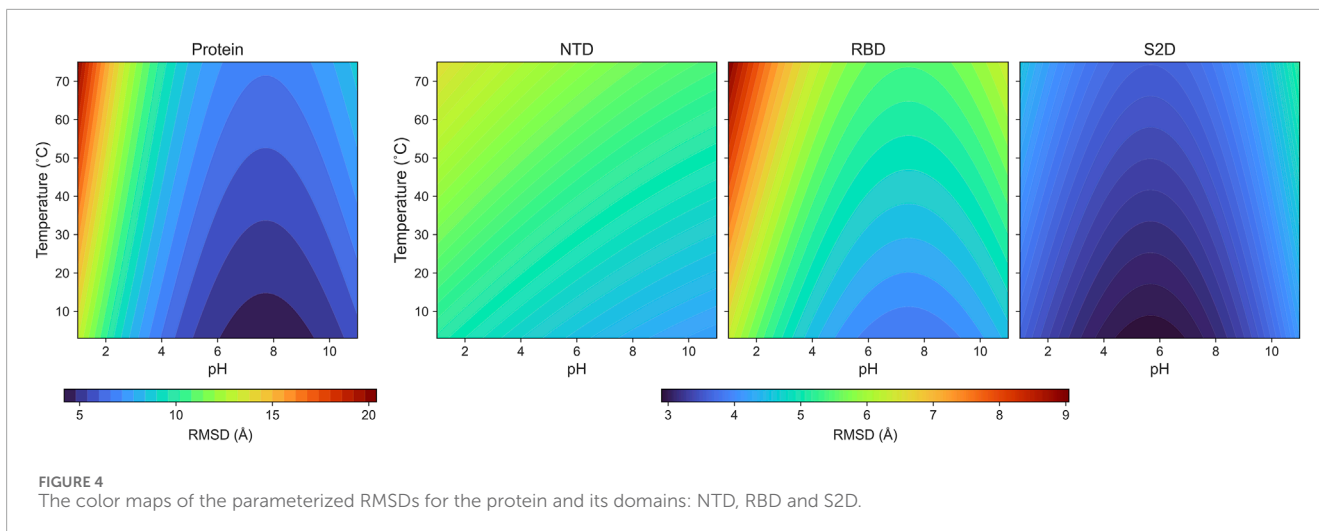
	R_0	α_T	β_a	β_b	R^2
R_{protein}	2.20	0.007	0.29	0.20	0.92
R_{NTD}	2.19	0.004	0.06	0.02	0.84
R_{RBD}	1.88	0.006	0.18	0.15	0.90
R_{S2D}	1.47	0.003	0.13	0.20	0.90

domains, and Figure 4 shows a less pronounced parabolic curve for the NTD, indicating minimal influence of pH changes, particularly at basic conditions.

3.3 The residue-based analysis

Finally, our residue-based analysis delves into the behavior of individual amino acids within the S-protein, measuring their specific fluctuations. This level of detail helps us pinpoint particular residues that may be essential to the protein's overall stability.

For individual residues, we use RMSF to quantify time-averaged fluctuations from their initial positions within a time window of 250 ns–450 ns. The RMSF values for each pH values are shown in Figure 5, with distinct colors representing each



thermal condition. Residues in the heptad repeat sequence 2, transmembrane and cytoplasmic tail consistently display high RMSF across all pH and thermal conditions, reflecting their inherent flexibility in the S-protein stalk regions. The [Figure 6](#) displays the S-protein structure averaged over the 250 ns–450 ns

trajectory under selected conditions, highlighting the differences between low and high temperatures as well as acidic and basic environments. Residues are color-mapped to reflect their respective RMSF values. A comprehensive structure figure of all 30 conditions is provided in [Supplementary Figure S3](#).

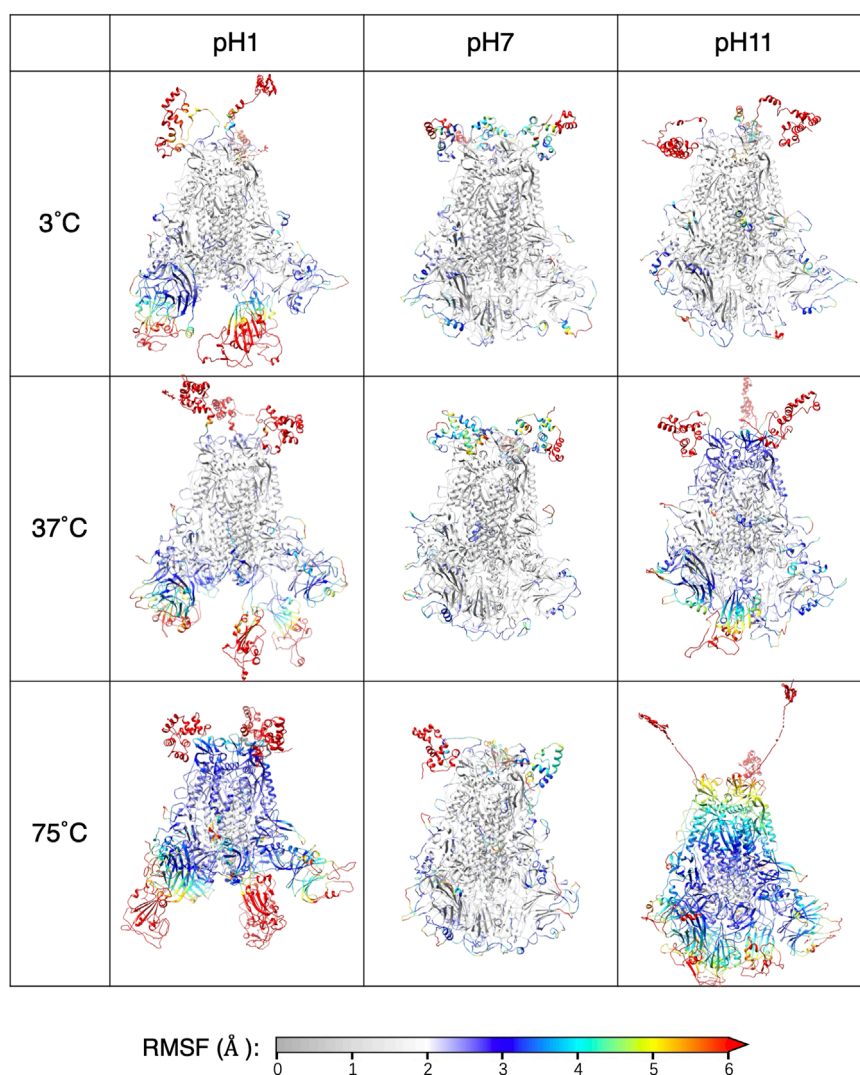


FIGURE 6
The averaged structure with each residue colored by the RMSF.

While most residues remain relatively stable, some display significant fluctuations under extreme pH, suggesting these residues may contribute to major structural changes in the S-protein. At moderate pH, even heat has minimal impact on RMSF. At pH = 1, 3 and 11, most of the highly fluctuating residues, with RMSF values exceeding 5 Å, are in the RBD and NTD, with increasing fluctuations at increasing temperatures. At pH = 1 and 11, a few residues in the S2 domain also exhibit unusually high RMSF values compared to pH = 7.

4 Discussions

While current experiments show that SARS-CoV-2 maintains high stability between pH = 3 and 10 at room temperature and remains stable for up to 2 weeks at 4°C, it rapidly loses activity within 5 min at 70°C (Chin et al., 2020). However, the

specific interactions between pH and thermal effects on the S-protein, and their implications for structural stability, are not fully understood. This study offers an in-depth examination of the structure and dynamics of the SARS-CoV-2 S-protein at pH = 1 through pH = 11 and temperatures 3°C through 75°C. Our findings reveal the individual, and coupling, effects of pH and thermal conditions.

The RMSD analysis reveals three distinct structural states of the SARS-CoV-2 S-protein. The protein remains structurally stable at pH = 5, 7, and 9, consistent with *in vitro* findings on other coronaviruses under similar pH conditions. At pH = 3 below 37°C and at pH = 11 across all thermal conditions, the S-protein exhibits transitional states where domain-specific structural changes begin, such as alterations in the RBD and NTD. At pH = 1 and at high temperatures for pH = 3, substantial structural deviations occur, suggesting that these conditions push the protein beyond its stable state. Despite these conformational changes, the secondary structure of the S-protein remains notably stable across a range of conditions,

as confirmed by DSSP and inter-mainchain H-bond analyses, consistent with studies showing secondary structure stability up to temperatures as high as 60°C (Khan et al., 2022).

The Rg, SASA and chain-chain interface area analyses provide additional detail, indicating that high temperatures lead to protein contraction while extreme pH promotes expansion. An intersectional effect emerges, with highly acidic conditions dominate the influence and thermal conditions become more impactful at moderate pH. This interaction suggests that S-protein stability depends on the combined effects of pH and thermal conditions. Acidic conditions weaken interchain interactions, potentially destabilizing the trimeric quaternary structure, as seen in the reduced chain-chain interface area, which may facilitate partial dissociation of the trimer into individual chains.

The domain and residue-based analysis, analyses reveal further insights into these structural changes. The RBD shows significant structural deviations at pH = 1 and 3, including translation and rotation that could affect binding efficiency. While no obvious change can be detected under moderate pH = 5, 7, and 9, consistent with prior observations (Xie et al., 2022) that the S-protein RBD is most stable between pH = 6 and 9. Conversely, the S2D shows increased sensitivity at basic pH = 11. High RMSF values in the RBD and NTD at pH = 1, 3, and 11 suggest local structural instability. The differential responses of RBD and S2D to acidic *versus* basic environments, as seen in the parametrized RMSD analysis, underscore each domain's unique adaptability to environmental changes.

We recognize the crucial role that glycan groups play in providing structural stability and aiding immune evasion of the S-protein by shielding it and influencing its interactions with host receptors (Casalino et al., 2020; Watanabe et al., 2020). In this study, we exclude glycan groups to concentrate solely on the core protein, a choice made to mitigate computational burden while focusing on our objective: to evaluate the bioactivity and conformational transitions of the S-protein and to examine the correlation between these transitions at varying pH and thermal conditions.

This study expands upon prior research (Niu et al., 2022; Niu et al., 2024) by comprehensively exploring the full range of pH and thermal impacts on the S-protein. Our findings are consistent with recent studies on the stability of coronavirus proteins at various conditions, underscoring the critical roles of pH and temperature in modulating protein structure and stability. Gaining insight into these conformational dynamics is essential for understanding the S-protein's behavior across different physiological and pathological environments. Additionally, the high computational costs associated with all-atom MD simulations highlight the promise of AI-driven models (Han et al., 2021; Liang et al., 2021; Zhang et al., 2024) in enhancing our ability to simulate complex biological systems. These models offer the potential to alleviate computational demands while effectively capturing the dynamics of the S-protein across a wide range of conditions and states.

5 Conclusion

In conclusion, this study elucidates the coupling effects of pH and thermal conditions on the structures and dynamics of the SARS-CoV-2 S-protein. Our analyses reveal that acidic conditions (pH =

1 and 3) lead to significant structural changes, particularly in the RBD and NTD, while heat intensifies these effects. Basic conditions, especially at pH = 11, also alter structures, particularly in the S2D. The data underscore the S-protein's differential response to acidic and basic environments, with extreme pH conditions dominate the structure impact while, at moderate pH, thermal effects lead.

This work highlights the potential impact of environmental factors on the structures of the S-protein and these insights could be critical in forming antiviral strategies, in devising storage conditions for vaccines, and even in designing therapeutics. Further studies, leveraging on potential advances in modeling technologies, could explore the longer-term effects at extreme pH and thermal conditions on the S-protein and, even, the entire virus.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://github.com/Niuzyx/TpH-Sprotein-simulation.git>, The data that support the findings of this study are available in github.

Author contributions

ZN: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Software, Visualization, Writing—original draft, Writing—review and editing. GK: Formal Analysis, Investigation, Methodology, Writing—review and editing. MR: Writing—review and editing, Project administration. MS: Writing—review and editing. EP: Writing—review and editing, Validation. BA: Writing—review and editing, Validation. YD: Conceptualization, Resources, Supervision, Writing—review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. We gratefully acknowledge support for this project from the SUNY-IBM Consortium Award, and the Stony Brook University OVPR and IEDM COVID-19 Seed Grant.

Acknowledgments

All simulations are conducted on the AiMOS supercomputer at Rensselaer Polytechnic Institute and the WSC Cluster at the IBM T. J. Watson Research Center through an IBM Faculty Award FP0002468 (Y. Deng).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the

References

- Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., et al. (2015). GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* 1–2, 19–25. doi:10.1016/j.softx.2015.06.001
- Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G. R., et al. (2021). Accurate prediction of protein structures and interactions using a three-track neural network. *Science* 373 (6557), 871–876. doi:10.1126/science.abc8754
- Becker, O. M. (1997). Geometric versus topological clustering: an insight into conformation mapping. *Proteins* 27 (2), 213–226. doi:10.1002/(sici)1097-0134(199702)27:2<213::aid-prot8>3.0.co;2-g
- Best, R. B., Zhu, X., Shim, J., Lopes, P. E., Mittal, J., Feig, M., et al. (2012). Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ , ψ and side-chain $\chi(1)$ and $\chi(2)$ dihedral angles. *J. Chem. Theory Comput.* 8 (9), 3257–3273. doi:10.1021/ct300400x
- Brooks, B. R., Brooks, C. L., Mackerell, A. D., Jr., Nilsson, L., Petrella, R. J., Roux, B., et al. (2009). CHARMM: the biomolecular simulation program. *J. Comput. Chem.* 30 (10), 1545–1614. doi:10.1002/jcc.21287
- Buslav, P., Aho, N., Jansen, A., Bauer, P., Hess, B., and Groenhof, G. (2022). Best practices in constant pH MD simulations: accuracy and sampling. *J. Chem. Theory Comput.* 18 (10), 6134–6147. doi:10.1021/acs.jctc.2c00517
- Casalino, L., Gaieb, Z., Goldsmith, J. A., Hjorth, C. K., Dommer, A. C., Harbison, A. M., et al. (2020). Beyond shielding: the roles of glycans in the SARS-CoV-2 spike protein. *ACS central Sci.* 6 (10), 1722–1734. doi:10.1021/acscentsci.0c01056
- Chan, K. H., Peiris, J. S., Lam, S. Y., Poon, L. L., Yuen, K. Y., and Seto, W. H. (2011). The effects of temperature and relative humidity on the viability of the SARS coronavirus. *Adv. Virol.* 2011, 734690. doi:10.1155/2011/734690
- Chen, W., and Shen, J. K. (2014). Effects of system net charge and electrostatic truncation on all-atom constant pH molecular dynamics. *J. Comput. Chem.* 35 (27), 1986–1996. doi:10.1002/jcc.23713
- Chen, X., and Geiger, J. D. (2020). Janus sword actions of chloroquine and hydroxychloroquine against COVID-19. *Cell. Signal* 73, 109706. doi:10.1016/j.cellsig.2020.109706
- Chin, A. W. H., Chu, J. T. S., Perera, M. R. A., Hui, K. P. Y., Yen, H. L., Chan, M. C. W., et al. (2020). Stability of SARS-CoV-2 in different environmental conditions. *Lancet Microbe* 1 (1), e10. doi:10.1016/S2666-5247(20)30003-3
- Chu, V. C., McElroy, L. J., Chu, V., Bauman, B. E., and Whittaker, G. R. (2006). The avian coronavirus infectious bronchitis virus undergoes direct low-pH-dependent fusion activation during entry into host cells. *J. Virol.* 80 (7), 3180–3188. doi:10.1128/JVI.80.7.3180-3188.2006
- Cia, G., Pucci, F., and Rooman, M. (2022). Analysis of the neutralizing activity of antibodies targeting open or closed SARS-CoV-2 spike protein conformations. *Int. J. Mol. Sci.* 23 (4), 2078. doi:10.3390/ijms23042078
- Costa, L. B., Perez, L. G., Palmeira, V. A., Macedo, E. C. T., Ribeiro, V. T., Lanza, K., et al. (2020). Insights on SARS-CoV-2 molecular interactions with the renin-angiotensin system. *Front. Cell. Dev. Biol.* 8, 559841. doi:10.3389/fcell.2020.559841
- Coutinho, J. V. P., Macedo-da-Silva, J., Mule, S. N., Kronenberger, T., Rosa-Fernandes, L., Wrenger, C., et al. (2022). Glycoprotein molecular dynamics analysis: SARS-CoV-2 spike glycoprotein case study. *Adv. Protein Chem. Struct. Biol.* 131, 277–309. doi:10.1016/bs.apcsb.2022.05.004
- Darnell, M. E., Subbarao, K., Feinstone, S. M., and Taylor, D. R. (2004). Inactivation of the coronavirus that induces severe acute respiratory syndrome, SARS-CoV. *J. Virol. Methods* 121 (1), 85–91. doi:10.1016/j.jviromet.2004.06.006
- de Oliveira, V. M., Liu, R., and Shen, J. (2022). Constant pH molecular dynamics simulations: current status and recent applications. *Curr. Opin. Struct. Biol.* 77, 102498. doi:10.1016/j.sbi.2022.102498
- Edwards, R. J., Mansouri, K., Stalls, V., Manne, K., Watts, B., Parks, R., et al. (2021). Cold sensitivity of the SARS-CoV-2 spike ectodomain. *Nat. Struct. Mol. Biol.* 28 (2), 128–131. doi:10.1038/s41594-020-00547-5
- Frances-Monerris, A., Hognon, C., Miclot, T., Garcia-Iriepa, C., Iriepa, I., Terenzi, A., et al. (2020). Molecular basis of SARS-CoV-2 infection and rational design of potential antiviral agents: modeling and simulation approaches. *J. Proteome Res.* 19 (11), 4291–4315. doi:10.1021/acs.jproteome.0c00779
- Gur, M., Taka, E., Yilmaz, S. Z., Kilinc, C., Aktas, U., and Golcuk, M. (2020). Conformational transition of SARS-CoV-2 spike glycoprotein between its closed and open states. *J. Chem. Phys.* 153 (7), 075101. doi:10.1063/5.0011141
- Han, C., Zhang, P., Bluestein, D., Cong, G., and Deng, Y. (2021). Artificial intelligence for accelerating time integrations in multiscale modeling. *J. Comput. Phys.* 427, 110053. doi:10.1016/j.jcp.2020.110053
- Hanson, W. A. (2019). The CORAL supercomputer systems. *IBM J. Res. Dev.* 64 (3/4), 1–1. doi:10.1147/jrd.2019.2960220
- He, J., Tao, H., Yan, Y., Huang, S. Y., and Xiao, Y. (2020). Molecular mechanism of evolution and human infection with SARS-CoV-2. *Viruses* 12 (4), 428. doi:10.3390/v12040428
- Islam, M. A., Islam, S., Haque, E., Rahman, M., Uddin, A., Khasruzzaman, A., et al. (2020). Thermal and pH sensitivity of avian corona and influenza viruses: a model study for inactivation of SARS-CoV-2 (COVID-19) and other flu viruses. *Int. Res. J. Med. Med. Sci.* 8, 42–56. doi:10.30918/irjmmms.82.20.028
- Juraszek, J., Ruten, L., Blokland, S., Bouchier, P., Voorzaat, R., Ritschel, T., et al. (2021). Stabilizing the closed SARS-CoV-2 spike trimer. *Nat. Commun.* 12 (1), 244. doi:10.1038/s41467-020-20321-x
- Kabsch, W., and Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22 (12), 2577–2637. doi:10.1002/bip.360221211
- Kandzia, F., Ostermeier, K., and Zacharias, M. (2019). Global dynamics of yeast Hsp90 middle and C-terminal dimer studied by advanced sampling simulations. *Front. Mol. Biosci.* 6, 93. doi:10.3389/fmolb.2019.00093
- Ke, Z., Oton, J., Qu, K., Cortese, M., Zila, V., McKeane, L., et al. (2020). Structures and distributions of SARS-CoV-2 spike proteins on intact virions. *Nature* 588 (7838), 498–502. doi:10.1038/s41586-020-2665-2
- Khan, F. I., Lobb, K. A., and Lai, D. (2022). The molecular basis of the effect of temperature on the structure and function of SARS-CoV-2 spike protein. *Front. Mol. Biosci.* 9, 794960. doi:10.3389/fmolb.2022.794960
- Lamarre, A., and Talbot, P. J. (1989). Effect of pH and temperature on the infectivity of human coronavirus 229E. *Can. J. Microbiol.* 35 (10), 972–974. doi:10.1139/m89-160
- Lasham, J., Djurabekova, A., Zickermann, V., Vonck, J., and Sharma, V. (2024). Role of protonation states in the stability of molecular dynamics simulations of high-resolution membrane protein structures. *J. Phys. Chem. B* 128 (10), 2304–2316. doi:10.1021/acs.jpcc.3c07421
- Leclercq, I., Batejat, C., Burguière, A. M., and Manuguerra, J. C. (2014). Heat inactivation of the Middle East respiratory syndrome coronavirus. *Influenza other Respir. viruses* 8 (5), 585–586. doi:10.1111/irv.12261
- Leonard, C., Phillips, C., and McCarty, J. (2021). Insight into seeded tau fibril growth from molecular dynamics simulation of the alzheimer's disease protofibril core. *Front. Mol. Biosci.* 8, 624302. doi:10.3389/fmolb.2021.624302
- Liang, D., Song, M., Niu, Z., Zhang, P., Rafailovich, M., and Deng, Y. (2021). Supervised machine learning approach to molecular dynamics forecast of SARS-CoV-2 spike glycoproteins at varying temperatures. *MRS Adv.* 6 (13), 362–367. doi:10.1557/s43580-021-00021-4

reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2025.1545041/full#supplementary-material>

- Maisuradze, G. G., and Leitner, D. M. (2007). Free energy landscape of a biomolecule in dihedral principal component space: sampling convergence and correspondence between structures and minima. *Proteins* 67 (3), 569–578. doi:10.1002/prot.21344
- Malaspina, D. C., and Farauo, J. (2020). Computer simulations of the interaction between SARS-CoV-2 spike glycoprotein and different surfaces. *Biointerphases* 15 (5), 051008. doi:10.1116/6.0000502
- Mark, P., and Nilsson, L. (2001). Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *J. Phys. Chem. A* 105 (43), 9954–9960. doi:10.1021/jp003020w
- Marti, D., Torras, J., Bertran, O., Turon, P., and Aleman, C. (2021). Temperature effect on the SARS-CoV-2: a molecular dynamics study of the spike homotrimeric glycoprotein. *Comput. Struct. Biotechnol. J.* 19, 1848–1862. doi:10.1016/j.csbj.2021.03.037
- Niu, Z., Hasegawa, K., Deng, Y., Zhang, Z., Rafailovich, M., Simon, M., et al. (2022). Modeling of the thermal properties of SARS-CoV-2 S-protein. *Front. Mol. Biosci.* 9, 953064. doi:10.3389/fmolb.2022.953064
- Niu, Z., Kementzidis, G., Zhang, P., Zhang, Z., Essuman, B., Hasegawa, K., et al. (2024). Modelling of SARS-CoV-2 spike protein structures at varying pH values. *Mol. Simul.* 50 (17–18), 1540–1552. doi:10.1080/08927022.2024.2415524
- Olsson, M. H., Sondergaard, C. R., Rostkowski, M., and Jensen, J. H. (2011). PROPKA3: consistent treatment of internal and surface residues in empirical pKa predictions. *J. Chem. Theory Comput.* 7 (2), 525–537. doi:10.1021/ct100578z
- Ostrowska, N., Feig, M., and Trylska, J. (2019). Modeling crowded environment in molecular simulations. *Front. Mol. Biosci.* 6, 86. doi:10.3389/fmolb.2019.00086
- Pace, C. N., Fu, H., Lee Fryar, K., Landua, J., Trevino, S. R., Schell, D., et al. (2014). Contribution of hydrogen bonds to protein stability. *Protein Sci.* 23 (5), 652–661. doi:10.1002/pro.2449
- Papaleo, E., Mereghetti, P., Fantucci, P., Grandori, R., and De Gioia, L. (2009). Free-energy landscape, principal component analysis, and structural clustering to identify representative conformations from molecular dynamics simulations: the myoglobin case. *J. Mol. Graph. Model.* 27 (8), 889–899. doi:10.1016/j.jmgs.2009.01.006
- Pipito, L., Rujan, R. M., Reynolds, C. A., and Deganutti, G. (2022). Molecular dynamics studies reveal structural and functional features of the SARS-CoV-2 spike protein. *Bioessays* 44 (9), e2200060. doi:10.1002/bies.202200060
- Rath, S. L., and Kumar, K. (2020). Investigation of the effect of temperature on the structure of SARS-CoV-2 spike protein by molecular dynamics simulations. *Front. Mol. Biosci.* 7, 583523. doi:10.3389/fmolb.2020.583523
- Rowan, N. J., Meade, E., and Garvey, M. (2021). Efficacy of frontline chemical biocides and disinfection approaches for inactivating SARS-CoV-2 variants of concern that cause coronavirus disease with the emergence of opportunities for green eco-solutions. *Curr. Opin. Environ. Sci. Health* 23, 100290. doi:10.1016/j.coesh.2021.100290
- Sahihi, M., and Farauo, J. (2022). Molecular dynamics simulations of adsorption of SARS-CoV-2 spike protein on polystyrene surface. *J. Chem. Inf. Model.* 62 (16), 3814–3824. doi:10.1021/acs.jcim.2c00562
- Shang, J., Wan, Y., Luo, C., Ye, G., Geng, Q., Auerbach, A., et al. (2020a). Cell entry mechanisms of SARS-CoV-2. *Proc. Natl. Acad. Sci. U. S. A.* 117 (21), 11727–11734. doi:10.1073/pnas.2003138117
- Shang, J., Ye, G., Shi, K., Wan, Y., Luo, C., Aihara, H., et al. (2020b). Structural basis of receptor recognition by SARS-CoV-2. *Nature* 581 (7807), 221–224. doi:10.1038/s41586-020-2179-y
- Sturman, L. S., Ricard, C. S., and Holmes, K. V. (1990). Conformational change of the coronavirus peplomer glycoprotein at pH 8.0 and 37 degrees C correlates with virus aggregation and virus-induced cell fusion. *J. Virol.* 64 (6), 3042–3050. doi:10.1128/JVI.64.6.3042-3050.1990
- Takemura, K., and Kitao, A. (2007). Effects of water model and simulation box size on protein diffusional motions. *J. Phys. Chem. B* 111 (41), 11870–11872. doi:10.1021/jp0756247
- Talley, K., and Alexov, E. (2010). On the pH-optimum of activity and stability of proteins. *Proteins* 78 (12), 2699–2706. doi:10.1002/prot.22786
- Tilton, R. F., Dewan, J. C., and Petsko, G. A. (1992). Effects of temperature on protein structure and dynamics: X-ray crystallographic studies of the protein ribonuclease-A at nine different temperatures from 98 to 320K. *Biochemistry* 31 (9), 2469–2481. doi:10.1021/bi00124a006
- Touw, W. G., Baakman, C., Black, J., te Beek, T. A., Krieger, E., Joosten, R. P., et al. (2015). A series of PDB-related databanks for everyday needs. *Nucleic Acids Res.* 43 (Database issue), D364–D368. doi:10.1093/nar/gku1028
- Turonova, B., Sikora, M., Schurmann, C., Hagen, W. J. H., Welsch, S., Blanc, F. E. C., et al. (2020). *In situ* structural analysis of SARS-CoV-2 spike reveals flexibility mediated by three hinges. *Science* 370 (6513), 203–208. doi:10.1126/science.abd5223
- Van Doremalen, N., Bushmaker, T., Morris, D. H., Holbrook, M. G., Gamble, A., Williamson, B. N., et al. (2020). Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1. *N. Engl. J. Med.* 382 (16), 1564–1567. doi:10.1056/NEJMc2004973
- Walls, A. C., Park, Y. J., Tortorici, M. A., Wall, A., McGuire, A. T., and Veesler, D. (2020). Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* 181 (2), 281–292.e6. doi:10.1016/j.cell.2020.02.058
- Walls, A. C., Tortorici, M. A., Bosch, B. J., Frenz, B., Rottier, P. J. M., DiMaio, F., et al. (2016). Cryo-electron microscopy structure of a coronavirus spike glycoprotein trimer. *Nature* 531 (7592), 114–117. doi:10.1038/nature16988
- Warwicker, J. (2021). A model for pH coupling of the SARS-CoV-2 spike protein open/closed equilibrium. *Brief. Bioinform* 22 (2), 1499–1507. doi:10.1093/bib/bbab056
- Watanabe, Y., Allen, J. D., Wrapp, D., McLellan, J. S., and Crispin, M. (2020). Site-specific glycan analysis of the SARS-CoV-2 spike. *Science* 369 (6501), 330–333. doi:10.1126/science.abb9983
- Wrapp, D., Wang, N., Corbett, K. S., Goldsmith, J. A., Hsieh, C.-L., Abiona, O., et al. (2020). Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* 367 (6483), 1260–1263. doi:10.1126/science.abb2507
- Xia, S., Zhu, Y., Liu, M., Lan, Q., Xu, W., Wu, Y., et al. (2020). Fusion mechanism of 2019-nCoV and fusion inhibitors targeting HR1 domain in spike protein. *Cell. Mol. Immunol.* 17 (7), 765–767. doi:10.1038/s41423-020-0374-2
- Xie, Y., Guo, W., Lopez-Hernandez, A., Teng, S., and Li, L. (2022). The pH effects on SARS-CoV and SARS-CoV-2 spike proteins in the process of binding to hACE2. *Pathogens* 11 (2), 238. doi:10.3390/pathogens11020238
- Yan, R., Zhang, Y., Li, Y., Ye, F., Guo, Y., Xia, L., et al. (2021). Structural basis for the different states of the spike protein of SARS-CoV-2 in complex with ACE2. *Cell Res.* 31 (6), 717–719. doi:10.1038/s41422-021-00490-0
- Zhang, C., and Huang, J. (2021). Interactions between nucleosomes: from atomistic simulation to polymer model. *Front. Mol. Biosci.* 8, 624679. doi:10.3389/fmolb.2021.624679
- Zhang, Z., Kementzidis, G., Zhang, P., Zhang, L., Kozloski, J., Hansen, A., et al. (2024). Learning coarse-grained force fields for fibrogenesis modeling. *Comput. Phys. Commun.* 295, 108964. doi:10.1016/j.cpc.2023.108964
- Zhou, T., Tsybovsky, Y., Gorman, J., Rapp, M., Cerutti, G., Chuang, G. Y., et al. (2020). Cryo-EM structures of SARS-CoV-2 spike without and with ACE2 reveal a pH-dependent switch to mediate endosomal positioning of receptor-binding domains. *Cell. Host Microbe* 28 (6), 867–879.e5. doi:10.1016/j.chom.2020.11.004