

Evolution and Pathogenicity of SARS-CoVs: A Microcanonical Analysis of Receptor-Binding Motifs

Rafael B. Frigori^{1,*}

¹*Universidade Tecnológica Federal do Paraná, Rua Cristo Rei 19, CEP 85902-490, Toledo (PR), Brazil*

The rapid evolution and global impact of coronaviruses, notably SARS-CoV-1 and SARS-CoV-2, underscore the importance of understanding their molecular mechanisms in detail. This study focuses on the receptor-binding motif (RBM) within the Spike protein of these viruses, a critical element for viral entry through interaction with the ACE2 receptor. We investigate the sequence variations in the RBM across SARS-CoV-1, SARS-CoV-2 and its early variants of concern (VOCs). Utilizing multicanonical simulations and microcanonical analysis, we examine how these variations influence the folding dynamics, thermostability, and solubility of the RBMs. Our methodology includes calculating the density of states (DoS) to identify structural phase transitions and assess thermodynamic properties. Furthermore, we solve the Poisson-Boltzmann equation to model the solubility of the RBMs in aqueous environments. This methodology is expected to elucidate structural and functional differences in viral evolution and pathogenicity, likely improving targeted treatments and vaccines.

I. INTRODUCTION

The recent emergence and rapid dissemination of coronaviruses [1, 2], particularly SARS-CoV-1 and SARS-CoV-2, have profoundly impacted global health, emphasizing the urgent need for a detailed understanding of their molecular mechanisms, especially those involved in viral entry and infection [3–5]. Both SARS-CoV-1, responsible for the Severe Acute Respiratory Syndrome (SARS) outbreak in 2002-2003, and SARS-CoV-2, which triggered the COVID-19 pandemic starting in late 2019, belong to the *Coronaviridae* family. These viruses share a large, enveloped, positive-sense single-stranded RNA genome, and their origins are closely linked to zoonotic spillovers, likely from bats to intermediate hosts before reaching humans [6]. Despite their similarities, the evolutionary paths of these viruses have led to significant differences in their biological behavior and impact on human health, making the study of their Spike (S) protein critical for understanding their pathogenicity and transmissibility [7].

The Spike protein is a trimeric transmembrane glycoprotein that mediates the attachment of the virus to the host cell surface and the subsequent fusion of the viral and cellular membranes, which is critical for viral entry. The receptor-binding domain (RBD) of the Spike protein is the region that directly interacts with the angiotensin-converting enzyme 2 (ACE2) receptor on host cells, determining the virus's ability to infect [8, 9]. Within the RBD lies receptor-binding motif (RBM), a key subdomain responsible for the direct interaction with ACE2. This motif is highly conserved in coronaviruses but exhibits subtle variations between different strains and variants, influencing the binding affinity and, consequently, the virus's transmissibility and pathogenicity [10].

SARS-CoV-1 and SARS-CoV-2 share a common an-

cestry but differ significantly in the structural features of their Spike proteins, particularly in the RBM. SARS-CoV-1's RBM binds to ACE2 with high affinity, which correlates with its ability to cause severe respiratory illness [11, 12]. In contrast, SARS-CoV-2's RBM, although it shares structural similarities with SARS-CoV-1, has evolved to enhance binding affinity to ACE2, contributing to its higher transmissibility [3, 7, 13, 14]. Notably, early variants of concern of SARS-CoV-2, such as the Beta (P.1) and Gamma (B.1.351) variants, exhibit mutations within the RBM that alter key residues involved in ACE2 binding. These mutations, including the N501Y and E484K, not only enhance the binding affinity to ACE2 [15, 16] but also confer the ability to evade neutralizing antibodies generated by natural infection or vaccination [17, 18].

Moreover, the impact of these mutations extends beyond immune evasion [19, 20]. They also influence the Spike protein's biophysical properties, such as charge distribution, solubility, and structural stability [21]. For instance, the E484K mutation, found in both the Beta (β) and Gamma (γ) variants, introduces a positive charge in the RBM, which can disrupt local electrostatic interactions and modify the overall stability of the protein [22, 23]. Similarly, mutations like N501Y enhance the Spike protein's structural stability, which may contribute to more efficient viral entry [21]. Therefore, the structural and functional differences in the RBM of those SARS-CoVs underscore the complex relationship between viral evolution, protein dynamics, and pathogenicity.

From a physical perspective the folding behavior of RBMs, crucial for their thermostability and Spike interaction potential, can be rigorously analyzed using multicanonical simulations [24–27], which enable precise calculations of the density of states (DoS). This approach, rooted in D.H.E. Gross's theoretical framework [28], facilitates microcanonical analyzes that uncovers critical phenomena such as phase transitions [29], often missed in simulations using the canonical ensemble [30]. RBMs,

* frigori@utfpr.edu.br

as protein motifs, qualify as “small systems” under this scenario, where the small number of particles such as the limited amino acid residues in RBMs allows surface effects and fluctuations to significantly impact their behavior [31]. Small systems, such as atomic nuclei [32, 33], nanoscale materials [34–36], and proteins [31, 37, 38] as the Spike RBM motifs, exhibit unique thermostatistical properties, including energy-dependent equilibrium states in the microcanonical ensemble that lack direct counterparts in the canonical ensemble due to the non-concavity of the entropy function [30, 31, 34]. This non-concavity can lead to phenomena like negative heat capacity and metastable or unstable states, which canonical methods often overlook [28]. Thus, microcanonical analysis is crucial for accurately capturing these states and understanding the detailed thermostatics and functional behavior of RBMs.

In addition to folding dynamics, the solubility of proteins in an aqueous environment plays a crucial role in their biological function [39]. The solubility of these peptides can be described using the Poisson-Boltzmann equation (PBE) [40], which accounts for the electrostatic interactions between the protein and the surrounding solvent. By solving the PBE [41, 42] for each RBM variant, we can gain insights into how changes in the amino acid sequence influence these protein’s domains electrostatic potential, and consequently, its solubility [43]. These mutations are also expected to influence the propensity of Spike to form aggregates or interact with other molecules [44, 45], with potential consequences for viral infectivity and immune recognition [10, 46].

The combination of multicanonical simulations [27, 47], microcanonical analysis [28, 31], and electrostatic modeling [40, 48] provides a powerful methodology for understanding the molecular mechanisms that underlie the differences between SARS-CoV-1 and various SARS-CoV-2 variants. By delving into the detailed thermostatics and solubility properties of the RBM, we aim to uncover the fundamental principles governing the folding, stability, and function of these critical viral motifs. This knowledge is essential not only for understanding the molecular basis of viral evolution and pathogenicity but also for informing the development of therapeutic strategies and vaccines aimed at combating current and future coronavirus outbreaks.

The article is organized as follows. In Sections 2 and 3, we describe the setup for multicanonical simulations used to investigate the folding dynamics of SARS-CoV RBMs and the microcanonical analysis framework employed to reveal structural phase transitions in folding, often missed by canonical ensembles. In Section 4, we present our results, focusing on the structural phase transitions identified through density of states calculations and the application of the Poisson-Boltzmann equation to model RBM solubility. We also discuss the broader implications of these findings, particularly how RBM mutations influence viral evolution, pathogenicity, and potential therapeutic strategies. Finally, in Section 5, we

conclude by summarizing our key insights and proposing directions for future research.

II. MODELING AND COMPUTATIONAL SETUP

We perform massive parallel Monte Carlo simulations [49] within a computational setup previously employed on modeling Amylin, Insulin, and Amyloid- β interactions [43, 48, 50]. Here, the focus is on simulating the receptor-binding motifs (RBMs) of different SARS-CoVs to explore their folding dynamics and thermodynamic properties. The simulations employ the ECEPP/3 force field and the implicit solvation model SCH2, as implemented in the SMMP 3.0 package [47], which provides an effective balance between accuracy and computational efficiency, consistent with Molecular Dynamics predictions [43, 51, 52]. While the absence of explicit solvation is known to introduce phase transitions at significantly elevated temperatures, this is a minor drawback that must be accepted as a tradeoff for the high-statistics data achieved. This approach involves up to 36 million molecular sweeps across 300 multicanonical (MUCA) recursions, enabling a robust microcanonical analysis.

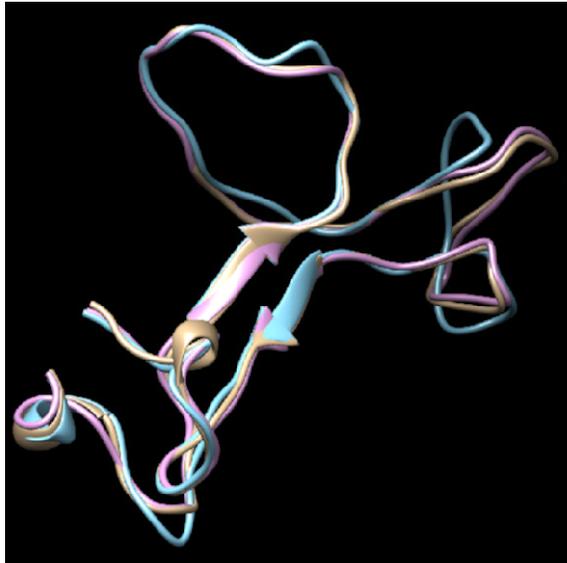
It shall be noted that RBMs of SARS-CoV-1 and SARS-CoV-2 has not been experimentally characterized in isolation, as available structural data pertain to the RBM embedded within the receptor-binding domain (RBD) of the Spike protein. The absence of experimental data on the isolated RBM poses a challenge, as its structural and dynamic properties may differ significantly from those observed in the native Spike context. This implies the need of computational modeling to generate initial structures suitable for simulation studies.

To begin the simulations, the RBMs for SARS-CoV-1, SARS-CoV-2 and its VOCs must first be encoded in the one-letter FASTA format, shown in (Table I). Noteworthy, the β/γ VOCs share the same RBM, whose mutations E484K and N501Y are prone to immunity evasiveness [19, 53]. Then, these sequences need to be converted into three-dimensional input structures (Figure 1), for instance using the I-TASSER (Iterative Threading ASSEmbling Refinement) software for homology modeling [54–56]. I-TASSER is a robust tool for protein structure and function prediction. It operates by initially threading the query sequence through a database of known protein structures to identify appropriate templates. The software then constructs full-length models via iterative fragment assembly simulations. This process not only predicts the 3D structure of the proteins but also offers potential insights into their biological functions, binding sites, and interactions.

The Multicanonical simulations are carried out in a rigid cubic box with 200Å sides, and the Boltzmann constant is set to $k_B = 1.987 \times 10^{-3}$ kcal/mol/K. Although initial molecular configurations are required to start the simulations, after thermalization, thermodynamic prop-

Virus	RBM Sequence (FASTA)
SARS-1	TRNIDATSTGNYNYKYR Y LRHGKLRPFERDISNVFSPDGKPC T PPALNCYWPLNDYGFYTTTIGIGYQPY
SARS-2 (WT)	SNNLDSKVGGN Y NYLYR L FRKSNLKP F ERDISTEIQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPY
SARS-2 (β/γ)	SNNLDSKVGGN Y NYLYR L FRKSNLKP F ERDISTEIQAGSTPCNGV K GFNCYFPLQSYGFQPT Y GVGYQPY

TABLE I. FASTA Sequences of RBMs of the studied SARS-CoVs

FIG. 1. RBMs modeled using I-TASSER serve as initial structures for SARS-1 (Blue), SARS-2 WT (Purple), SARS-2 β/γ (Brown) variants.

erties become independent of these initial conditions, as observed in ergodic Monte Carlo simulations. This allows for a comprehensive analysis of the folding dynamics and solubility of the RBMs across a wide temperature range, spanning from the homogeneous nucleation of ice crystals at 224.8 K to the critical point of water at 647 K.

III. MICROCANONICAL ANALYSIS AND SOLUBILITY MODELING

Canonical Monte Carlo (MC) simulations are limited in their ability to explore the full energy landscape of complex systems due to their reliance on the Boltzmann weight [27, 30], which is defined at a fixed temperature. This constraint often restricts the utility of re-weighting techniques for extrapolating thermodynamic quantities across different temperatures [57]. In contrast, Multicanonical (MUCA) simulations [24–27] offer a more versatile approach by employing generalized weights that allow the system to sample a wide range of energy states uniformly. The MUCA weight is given by

$$\omega_{muca} = \frac{1}{\Omega(E)} = e^{-S(E)} = e^{-\bar{\beta}(E)E + \bar{\alpha}(E)}, \quad (1)$$

where $\Omega(E)$ denotes the density of states (DoS), $S(E)$ is the microcanonical entropy, $\bar{\beta}(E)$ is the microcanonical inverse temperature, and $\bar{\alpha}(E)$ is a dimensionless free energy term.

The MUCA method comprises two key steps: (1) determining the appropriate multicanonical weights, usually through an iterative process that involves histogramming the internal energy, and (2) performing a Markov Chain Monte Carlo (MCMC) simulation using the weights from the first step, followed by reweighting to the Gibbs ensemble. In this study, we focus on microcanonical thermostats to analyze the folding and solubility of the RBM from different SARS-CoV variants. Thus, our objective is to concentrate on step 1, specifically on computing the multicanonical entropy

$$S_{muca}(E_k) = \beta_k E_k - \alpha_k, \quad (2)$$

where $\{\beta_k, \alpha_k\}$ are piecewise functions representing the MUCA parameters, which approximate the microcanonical entropy.

The calculation is implemented through an iterative process [25, 27] where initial MUCA weights are set to $\omega_{muca}^{(0)} = 1$, and the system undergoes a conventional Metropolis simulation. Data collected during this simulation is used to construct histograms of the energy distribution, $H(E)$, which are then used to update the MUCA weights. The iterative refinement of these weights follows the relation

$$\omega_{muca}^{(n+1)}(E) \equiv e^{-S_{muca}^{(n+1)}(E)} = c \cdot \frac{\omega_{muca}^{(n)}(E)}{H_{muca}^{(n)}(E)}, \quad (3)$$

where c is a normalization constant ensuring that $S_{muca}^{(n+1)}(E)$ correctly represents the microcanonical entropy. Traditionally, iterations aim for histogram flatness; here, we instead monitor for stabilization of $S_{muca}(E) \times E$ in the whole energetic range, as described in [48], which takes $\mathcal{O}(5M)$ sweeps.

After convergence, the microcanonical thermodynamics [28] of the RBMs folding is automatically analyzed using the PHAST package [58]. The temperature $T(E)$ as a function of energy is derived from the microcanonical inverse temperature

$$\beta(E)^* = k_B \cdot \beta(E) \equiv T^{-1}(E) = \frac{\partial S}{\partial E} \quad (4)$$

where k_B is the Boltzmann constant. The microcanonical specific heat c_V is then defined as

$$c_V = \frac{dE}{dT} = -\beta^2 / \left(\frac{\partial \beta}{\partial E} \right). \quad (5)$$

Finally, the Helmholtz free energy $F(E)$ can be obtained via a Legendre transform at a fixed temperature $T_c(E)$

$$F(E) = E - S(E) \cdot \left(\frac{\partial S}{\partial E} \right)^{-1} \Big|_{E=E^*(T_c)}. \quad (6)$$

We thoroughly investigate for potential first-order structural phase transitions involved in the folding of SARS-CoVs RBMs. In such cases, microcanonical caloric curves $\beta(E)^* \times E$ often exhibit backbends, or S-bends, which reveal metastable states that remain hidden in canonical ensemble simulations [28, 31, 34]. The pseudo-critical inverse temperatures $T_c(E)$ associated with these transitions are determined using the Maxwell construction across the energy range $\Delta \tilde{L} = E_{fold} - E_{unfold}$, where $\Delta \tilde{L}$ represents the latent heat of the folding transition, and E_{fold} and E_{unfold} correspond to the internal energies of the folded and unfolded states, respectively [31].

Furthermore, the solubility of RBMs in water, which is critical for their biological function, is analyzed using the Poisson-Boltzmann equation [40, 48, 52]. The solvation free energy ΔG_{solv} is decomposed into nonpolar (ΔG_{np}) and electrostatic (ΔG_{elec}) contributions

$$\Delta G_{solv} = \Delta G_{np} + \Delta G_{elec}. \quad (7)$$

The nonpolar contribution, associated with the formation energy of the molecular cavity, is often calculated using the solvent accessible surface area (SASA) model which is approximated as $\Delta G_{np} = \gamma \text{SASA}$, where γ represents the surface tension, typically treated as an atom-independent parameter [59]. The electrostatic contribution is given by

$$\Delta G_{elec} = \sum_k q_k \Delta \varphi_{vs}(r_k), \quad (8)$$

where q_k denotes the charge of atom k , and $\Delta \varphi_{vs}(r_k)$ represents the electrostatic potential difference between vacuum and solvent environments at the position of atom k . Accurately solving the Poisson-Boltzmann equation (PBEQ) to obtain these potentials is computationally demanding but necessary for precise estimation of the RBMs solvation free energy. In this study, we use the PBEQ-Solver [40–42] for these calculations, with the dielectric constants set to $\epsilon_s = 78.5$ for the solvent and $\epsilon_p = 2.0$ for the protein, and the salt concentration set to 150 mM, as in Ref. [48].

IV. RESULTS AND DISCUSSION

Parallel simulations were performed to determine the MUCA parameters $\{\beta_k, \alpha_k\}$, which were then utilized as input for the PHAST package [58]. These parameters enabled the derivation of the microcanonical thermostatics governing the folding structural phase transitions of the RBMs in SARS-CoV-1, SARS-CoV-2 WT, and its early VOCs β/γ . Thermostatistical results are summarized in extensive physical units in (Table II), they provide insights into the folding temperatures (T_{fold}), solvation energies (ΔG_{solv}), latent heats ($\Delta \tilde{L}$) and free energy barriers (ΔF) that dictate the structural stability and phase transition behavior of these motifs. The caloric curves and *intensive* energetic barriers ($\Delta f = \Delta F/\text{residues}$) of the Helmholtz free energy at the (pseudo) critical inverse temperature (β_c^*) are depicted in (Figure II) as functions of the internal energy per residue ($\varepsilon = E/\text{residues}$). These plots highlight the existence of regions with canonical metastabilities during the RBMs folding transitions, where using Maxwell’s constructions we identified the critical β_c^* and the associated *intensive* latent heats ($\Delta L = \Delta \tilde{L}/\text{residues}$) for the SARS-CoV-1 and SARS-CoV-2 WT RBMs. In contrast, the β/γ variants exhibit a second-order phase transition, marked by the absence of latent heat and Helmholtz free energy barriers, indicative of a smooth, continuous folding process similar to that observed in intrinsically disordered proteins such as Amylin and Amyloid- β [43, 48, 50]. The critical β_c^* (ε_c) of folding for β/γ variant is detected by locating the peak of its $c_v(\varepsilon) \times \varepsilon$ (Eq. 5) — not shown. Furthermore, (Fig. 3) presents the isoelectric surfaces of representative RBM configurations for the viruses, sampled at physiological temperatures (~ 310 K), which were used as input for computing the solvation energies via the PBEQ solver.

The correlation between thermostatical outcomes and thermo-structural features during RBM folding transitions is crucial, especially in the absence of experimentally resolved RBM structures. Our MUCA simulations, quantifying residues in α -helical and β -sheet conformations, illustrate how simulations might provide valuable guidance for future experimental design. Figure 4 shows the isocontours of internal energy per residue (ε) as a function of residues in these configurations. Residue classification follows the SMMP framework [47], using dihedral angles (ϕ, ψ) with ranges of $[-70^\circ \pm 30^\circ, -37^\circ \pm 30^\circ]$ for helices and $[-150^\circ \pm 30^\circ, -150^\circ \pm 30^\circ]$ for sheets. This figure highlights two notable findings: SARS-CoV-2 is more α -helix-prone than SARS-CoV-1, and the E484K and N501Y substitutions not only shift the folding phase transition to second order but also significantly enhance the α -helical content in the β/γ variants’ ground state compared to the WT. This trend, reminiscent of amyloidogenic Amylin isoforms studied using similar methods [51, 52, 58], can be related to the increased pathogenicity of these variants, as the Spike protein domains are

Virus	T_{fold} (K)	$\Delta G_{solv.}$ (kcal/mol)	$\Delta \tilde{L}$ (kcal/mol)	ΔF (kcal/mol)
SARS-1	402	-462	20.6	0.994
SARS-2 (WT)	541	-419	15.1	0.345
SARS-2 (β/γ)	513	-465	—	—

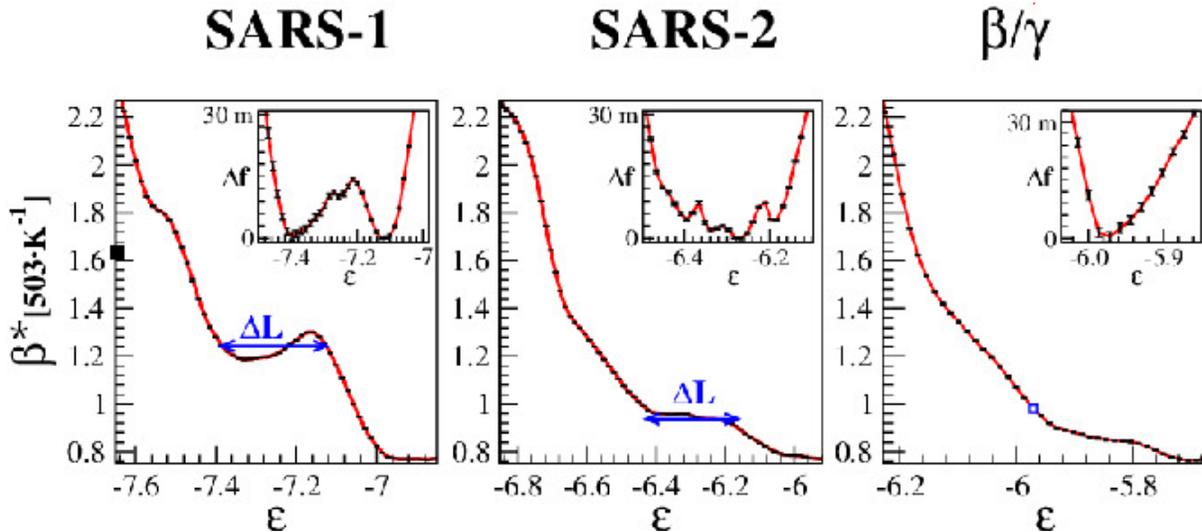
TABLE II. Thermodynamic results for RBMs of SARS-1, SARS-2 WT and its β/γ VOCs

FIG. 2. Microcanonical caloric curve for RBM motifs of SARS-CoVs. The black square denotes 310K, the bordered square denotes critical transition temperatures (when existing), ΔL is the *intensive* latent heat of folding obtained by Maxwell's equal-area constructions. The inserts show the respective *intensive* Helmholtz free energy Δf evaluated at the pseudo-critical temperature T_{fold} . Energy units in ϵ , Δf , and ΔL are normalized to kcal/mol/residues and the physical temperatures to $503 \cdot \beta^{-1}$ K.

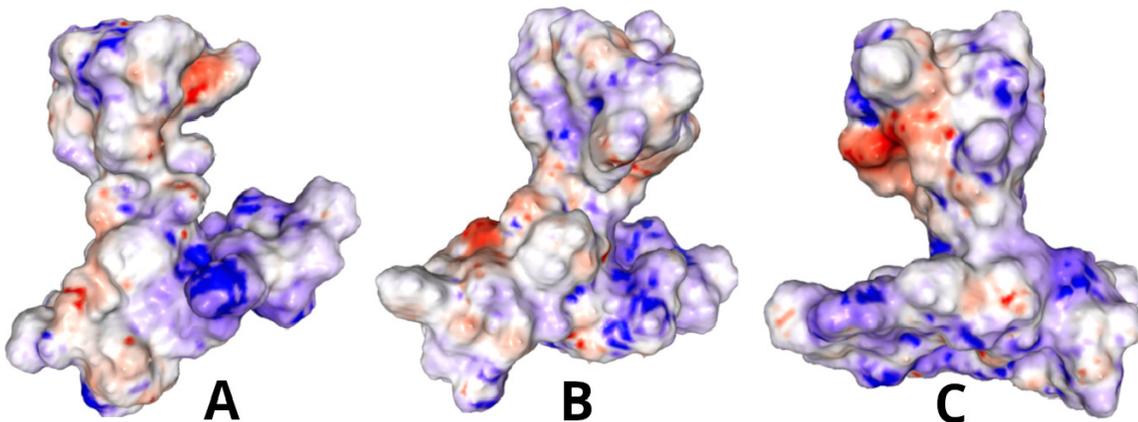


FIG. 3. Isoelectric surface solutions of PBEQ associated to representative configurations of RBMs molecular motifs of SARS-1 (A), SARS-2 WT (B) and SARS-2 β/γ (C).

known to induce protein aggregation under suitable biological conditions [44, 60, 61].

Next, we delve into the specific thermostistical and solvation properties of each RBM, providing a detailed analysis and discussion.

SARS-1: High Solubility and Structural Rigidity as Evolutionary Constraints

The RBM of SARS-CoV-1 exhibits a pronounced first-order folding transition, as indicated by a distinct back-bend in its caloric curve, presenting the largest latent heat ($\Delta L = 20.6$ kcal/mol), and also highest Helmholtz

free energy barrier among the RBMs studied ($\Delta F = 0.994$ kcal/mol). This strong first-order transition points to a stable folded state, further reinforced by the high folding temperature ($T_{fold} = 402$ K). The significant solvation energy of -462 kcal/mol reflects robust interactions with the solvent, resulting in appreciable solubility.

These thermodynamic characteristics suggest that SARS-CoV-1 RBM is structurally rigid and resistant to mutations. The rigidity and high structural stability of the RBM limit its ability to undergo significant conformational changes without compromising its stability, which likely contributed to the virus’s slower evolutionary diversification. The stringent structural constraints imposed by the RBMs thermodynamic profile may have hindered the virus’s ability to rapidly adapt, leading to a lower mutation rate compared to SARS-CoV-2. While this rigidity ensures stable binding to the ACE2 receptor, it also likely reduces the virus’s capacity to evade the host immune system, as the RBM structure is less flexible in altering epitopes for immune escape.

SARS-2 (WT): Structural Flexibility Facilitates Evolutionary Adaptation

The RBM of SARS-CoV-2 (RBM-WT) displays a weaker first-order phase transition compared to SARS-CoV-1, as evidenced by the absence of backbends in its caloric curve, which displays a latent heat ($\Delta L = 15.1$ kcal/mol), and a smaller Helmholtz free energy barrier ($\Delta F = 0.345$ kcal/mol). The folding temperature ($T_{fold} = 541$ K), derived from the flat plateau of the caloric curve, reflects a less distinct transition. This weaker phase transition implies greater structural flexibility of this RBM, allowing it to adapt to a broader range of structural perturbations.

Despite this increased flexibility, the solvation energy of the RBM-WT (-419 kcal/mol) suggests lower solubility compared to SARS-CoV-1, potentially driving evolutionary pressure for variants with enhanced RBM solubility. The structural flexibility of the RBM-WT is critical for the viruses evolutionary adaptability, enabling it to tolerate mutations while maintaining effective ACE2 binding. This adaptability promotes rapid mutation and evolution, which are vital for immune evasion and sustained transmission. Additionally, the RBM-WT’s lower solubility indicates a potential for further mutations that could enhance solubility while allowing structural modifications that alter antibody recognition sites, thereby improving immune evasion without significantly compromising ACE2 binding.

SARS-2 VOCs: Enhanced Pathogenicity and Immune Evasion

The RBM of β/γ variants, characterized by the E484K and N501Y mutations, present a particular thermody-

amic profile, undergoing a second-order phase transition. Unlike the first-order transitions observed in the RBMs of SARS-CoV-1 and SARS-CoV-2 WT, these variants do not exhibit latent heat of folding or Helmholtz free energy barriers, features typically also observed in fast folders as Intrinsically Disordered Proteins (IDPs) such as Amylin and Amyloid- β [43, 48]. The absence of these thermodynamic signatures indicates smoother, continuous transitions between conformational states, enhancing the flexibility and adaptability of the RBM. The increased solvation energy of -465 kcal/mol suggests that these variants have a higher solubility profile, similar to SARS-CoV-1, which might have emerged by evolutionary pressure as a mechanism to improve immune evasion of WT antibodies. The E484K mutation introduces a positively charged lysine at a key binding interface, altering electrostatic interactions and facilitating immune evasion [17, 18, 62]. The N501Y mutation replaces asparagine with tyrosine, which introduces a bulkier, hydrophobic side chain capable of forming stronger π -stacking interactions with ACE2 residues, particularly Y41 and K353 [63]. This enhances the affinity for ACE2 binding, stabilizing the RBM-ACE2 interface and further contributing to immune escape by promoting tighter receptor engagement. Together, these mutations synergistically exploit structural flexibility and solvent interactions. Therefore, our data confirm experimental findings in which such mutations boost immune evasion and ACE2 binding in these variants [17–19, 53], highlighting the underlying role of thermodynamic principles guiding viral evolution.

V. CONCLUSION

This study provides a detailed microcanonical analysis of the RBMs in SARS-CoV-1, SARS-CoV-2 WT, and the β/γ VOCs, offering critical insights into how thermodynamic properties influence viral evolution and pathogenicity. By utilizing the ECEPP/3 force field with implicit solvation, the analysis effectively balances computational efficiency with molecular accuracy, enabling robust data collection despite trade-offs such as elevated structural transition temperatures. The results highlight distinct thermodynamic and structural differences across these RBMs: SARS-CoV-1’s RBM, with a strong first-order phase transition and high solvation energy, exhibits structural rigidity that limits mutational flexibility and impedes immune evasion, reducing transmissibility. In contrast, the SARS-CoV-2 WT RBM shows greater thermodynamic flexibility, marked by a weaker first-order phase transition and reduced solvation energy, which promotes a broader mutational spectrum, facilitating the emergence of variants with enhanced immune evasion and transmissibility under selective pressures. The β/γ VOCs, including the E484K and N501Y mutations, illustrate these evolutionary dynamics, increasing RBM solvation energy and shifting folding transitions to

second-order, similar to intrinsically disordered proteins. This structural adaptation supports improved immune evasion and enhanced ACE2 receptor binding. Furthermore, the analysis reveals that SARS-CoV-2 RBMs exhibit a stronger propensity for α -helix formation compared to SARS-CoV-1, with the β/γ variants showing significant increases in α -helical content in their ground states. This trend, reminiscent of amyloidogenic proteins like Amylin, may be linked to the increased pathogenicity of these variants, as the Spike protein domains are known to promote aggregation under certain biological conditions. These findings underscore the importance of ongoing surveillance and targeted interventions to mitigate the risks posed by emerging variants, as these thermodynamic properties suggest continued potential for viral

evolution under selective pressures.

ACKNOWLEDGMENT

The Brazilian National Laboratory for Scientific Computing (LNCC) is acknowledged by supercomputing time granted at the Santos Dumont facility under project PHAST2. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. *This article is a tribute to the memory of my esteemed colleague and friend, Dr. Marcelo Fernandes, who passed away prematurely due to COVID-19.*

-
- [1] J. Cui, F. Li, and Z.-L. Shi, *Nature reviews microbiology* **17**, 181 (2019).
- [2] Z. Song, Y. Xu, L. Bao, L. Zhang, P. Yu, Y. Qu, H. Zhu, W. Zhao, Y. Han, and C. Qin, *viruses* **11**, 59 (2019).
- [3] J. Lan, J. Ge, J. Yu, S. Shan, H. Zhou, S. Fan, Q. Zhang, X. Shi, Q. Wang, L. Zhang, *et al.*, *nature* **581**, 215 (2020).
- [4] J. Shang, Y. Wan, C. Luo, G. Ye, Q. Geng, A. Auerbach, and F. Li, *Proceedings of the National Academy of Sciences* **117**, 11727 (2020).
- [5] C. B. Jackson, M. Farzan, B. Chen, and H. Choe, *Nature reviews Molecular cell biology* **23**, 3 (2022).
- [6] J. Xu, S. Zhao, T. Teng, A. E. Abdalla, W. Zhu, L. Xie, Y. Wang, and X. Guo, *Viruses* **12**, 244 (2020).
- [7] M. M. Hatmal, W. Alshaer, M. A. Al-Hatamleh, M. Hatmal, O. Smadi, M. O. Taha, A. J. Oweida, J. C. Boer, R. Mohamud, and M. Plebanski, *Cells* **9**, 2638 (2020).
- [8] F. Li, W. Li, M. Farzan, and S. C. Harrison, *Science* **309**, 1864 (2005).
- [9] W. Tai, L. He, X. Zhang, J. Pu, D. Voronin, S. Jiang, Y. Zhou, and L. Du, *Cellular & molecular immunology* **17**, 613 (2020).
- [10] C. Yi, X. Sun, J. Ye, L. Ding, M. Liu, Z. Yang, X. Lu, Y. Zhang, L. Ma, W. Gu, *et al.*, *Cellular & molecular immunology* **17**, 621 (2020).
- [11] W. Li, M. J. Moore, N. Vasilieva, J. Sui, S. K. Wong, M. A. Berne, M. Somasundaran, J. L. Sullivan, K. Luzuriaga, T. C. Greenough, *et al.*, *Nature* **426**, 450 (2003).
- [12] M. Hoffmann, H. Kleine-Weber, S. Schroeder, N. Krüger, T. Herrler, S. Erichsen, T. S. Schiergens, G. Herrler, N.-H. Wu, A. Nitsche, *et al.*, *cell* **181**, 271 (2020).
- [13] A. C. Walls, Y.-J. Park, M. A. Tortorici, A. Wall, A. T. McGuire, and D. Velesler, *Cell* **181**, 281 (2020).
- [14] B. Korber, W. M. Fischer, S. Gnanakaran, H. Yoon, J. Theiler, W. Abfalterer, N. Hengartner, E. E. Giorgi, T. Bhattacharya, B. Foley, *et al.*, *Cell* **182**, 812 (2020).
- [15] H. Tegally, E. Wilkinson, M. Giovanetti, A. Iranzadeh, V. Fonseca, J. Giandhari, D. Doolabh, S. Pillay, E. J. San, N. Msomi, *et al.*, *Nature* **592**, 438 (2021).
- [16] N. R. Faria, T. A. Mellan, C. Whittaker, I. M. Claro, D. d. S. Candido, S. Mishra, M. A. Crispim, F. C. Sales, I. Hawryluk, J. T. McCrone, *et al.*, *Science* **372**, 815 (2021).
- [17] W. F. Garcia-Beltran, E. C. Lam, K. S. Denis, A. D. Nitido, Z. H. Garcia, B. M. Hauser, J. Feldman, M. N. Pavlovic, D. J. Gregory, M. C. Poznansky, *et al.*, *Cell* **184**, 2372 (2021).
- [18] D. Zhou, W. Dejnirattisai, P. Supasa, C. Liu, A. J. Mentzer, H. M. Ginn, Y. Zhao, H. M. Duyvesteyn, A. Tuekprakhon, R. Nutalai, *et al.*, *Cell* **184**, 2348 (2021).
- [19] W. T. Harvey, A. M. Carabelli, B. Jackson, R. K. Gupta, E. C. Thomson, E. M. Harrison, C. Ludden, R. Reeve, A. Rambaut, S. J. Peacock, *et al.*, *Nature Reviews Microbiology* **19**, 409 (2021).
- [20] A. M. Carabelli, T. P. Peacock, L. G. Thorne, W. T. Harvey, J. Hughes, S. J. Peacock, W. S. Barclay, T. I. De Silva, G. J. Towers, and D. L. Robertson, *Nature Reviews Microbiology* **21**, 162 (2023).
- [21] R. Pondé, *Virology* **572**, 44 (2022).
- [22] Z. Zhang, J. Zhang, and J. Wang, *Frontiers in Public Health* **10**, 952916 (2022).
- [23] W. B. Wang, Y. Liang, Y. Q. Jin, J. Zhang, J. G. Su, and Q. M. Li, *Journal of Molecular Graphics and Modelling* **109**, 108035 (2021).
- [24] B. A. Berg and T. Neuhaus, *Physics Letters B* **267**, 249 (1991).
- [25] B. A. Berg, arXiv preprint hep-lat/9503019 (1995).
- [26] B. A. Berg, arXiv preprint cond-mat/9909236 (1999).
- [27] B. A. Berg, *Computer physics communications* **153**, 397 (2003).
- [28] D. H. Gross, *Microcanonical thermodynamics: phase transitions in "small" systems*, Vol. 66 (World Scientific, 2001).
- [29] H. E. Stanley, *Phase transitions and critical phenomena*, Vol. 7 (Clarendon Press, Oxford, 1971).
- [30] J. Hernández-Rojas and J. Gomez Llorente, *Physical review letters* **100**, 258104 (2008).
- [31] M. Bachmann, *Thermodynamics and statistical mechanics of macromolecular systems* (Cambridge University Press, 2014).
- [32] P. Chomaz, in *AIP Conference Proceedings*, Vol. 610 (American Institute of Physics, 2002) pp. 167–181.
- [33] N. Büyükcizmeci, N. Eren, R. Oğul, and R. Oğul, *Cankaya University Journal of Arts and Sciences* **1**, 55 (2005).

- [34] J. Barré, D. Mukamel, and S. Ruffo, *Physical Review Letters* **87**, 030601 (2001).
- [35] R. B. Frigori, L. G. Rizzi, and N. A. Alves, *The European Physical Journal B* **75**, 311 (2010).
- [36] N. A. Alves and R. B. Frigori, *Physica A: Statistical Mechanics and its Applications* **446**, 195 (2016).
- [37] R. B. Frigori, L. G. Rizzi, and N. A. Alves, *The Journal of Chemical Physics* **138** (2013).
- [38] R. B. Frigori, *Physical Review E* **90**, 052716 (2014).
- [39] Y. Levy and J. N. Onuchic, *Annu. Rev. Biophys. Biomol. Struct.* **35**, 389 (2006).
- [40] W. Im, D. Beglov, and B. Roux, *Computer physics communications* **111**, 59 (1998).
- [41] S. Jo, T. Kim, V. G. Iyer, and W. Im, *Journal of computational chemistry* **29**, 1859 (2008).
- [42] S. Jo, M. Vargyas, J. Vasko-Szedlar, B. Roux, and W. Im, *Nucleic acids research* **36**, W270 (2008).
- [43] R. B. Frigori, *Physical Chemistry Chemical Physics* **19**, 25617 (2017).
- [44] D. Idrees and V. Kumar, *Biochemical and Biophysical Research Communications* **554**, 94 (2021).
- [45] T. Bhardwaj, K. Gadhawe, S. K. Kapuganti, P. Kumar, Z. F. Brotzakis, K. U. Saumya, N. Nayak, A. Kumar, R. Joshi, B. Mukherjee, *et al.*, *Nature Communications* **14**, 945 (2023).
- [46] A. J. Greaney, A. N. Loes, K. H. Crawford, T. N. Starr, K. D. Malone, H. Y. Chu, and J. D. Bloom, *Cell host & microbe* **29**, 463 (2021).
- [47] J. H. Meinke, S. Mohanty, F. Eisenmenger, and U. H. Hansmann, *Computer Physics Communications* **178**, 459 (2008).
- [48] R. B. Frigori and F. Rodrigues, *Journal of Molecular Modeling* **27**, 1 (2021).
- [49] J. Zierenberg, M. Marenz, and W. Janke, *Physics Procedia* **53**, 55 (2014).
- [50] R. B. Frigori, *The European Physical Journal B* **97**, 156 (2024).
- [51] N. A. Alves and R. B. Frigori, *The Journal of Physical Chemistry B* **122**, 10714 (2018).
- [52] N. A. Alves, L. G. Dias, and R. B. Frigori, *Journal of Molecular Modeling* **25**, 1 (2019).
- [53] P. Masson, C. Hulo, E. De Castro, H. Bitter, L. Grunbaum, L. Essioux, L. Bougueleret, I. Xenarios, and P. Le Mercier, *Nucleic acids research* **41**, D579 (2012).
- [54] J. Yang and Y. Zhang, *Nucleic acids research* **43**, W174 (2015).
- [55] W. Zheng, C. Zhang, Y. Li, R. Pearce, E. W. Bell, and Y. Zhang, *Cell reports methods* **1** (2021).
- [56] X. Zhou, W. Zheng, Y. Li, R. Pearce, C. Zhang, E. W. Bell, G. Zhang, and Y. Zhang, *Nature Protocols* **17**, 2326 (2022).
- [57] D. Landau and K. Binder, *A guide to Monte Carlo simulations in statistical physics* (Cambridge university press, 2021).
- [58] R. B. Frigori, *Computer Physics Communications* **215**, 165 (2017).
- [59] B. Aguilar and A. V. Onufriev, *Journal of chemical theory and computation* **8**, 2404 (2012).
- [60] S. Nystrom and P. Hammarstrom, *Journal of the American Chemical Society* **144**, 8945 (2022).
- [61] M. H. Chang, J. H. Park, H. K. Lee, J. Y. Choi, and Y. H. Koh, *Biomedicines* **12**, 1223 (2024).
- [62] A. G. Wrobel, D. J. Benton, C. Roustan, A. Borg, S. Husain, S. R. Martin, P. B. Rosenthal, J. J. Skehel, and S. J. Gamblin, *Nature communications* **13**, 1178 (2022).
- [63] P. Han, C. Su, Y. Zhang, C. Bai, A. Zheng, C. Qiao, Q. Wang, S. Niu, Q. Chen, Y. Zhang, *et al.*, *Nature communications* **12**, 6103 (2021).

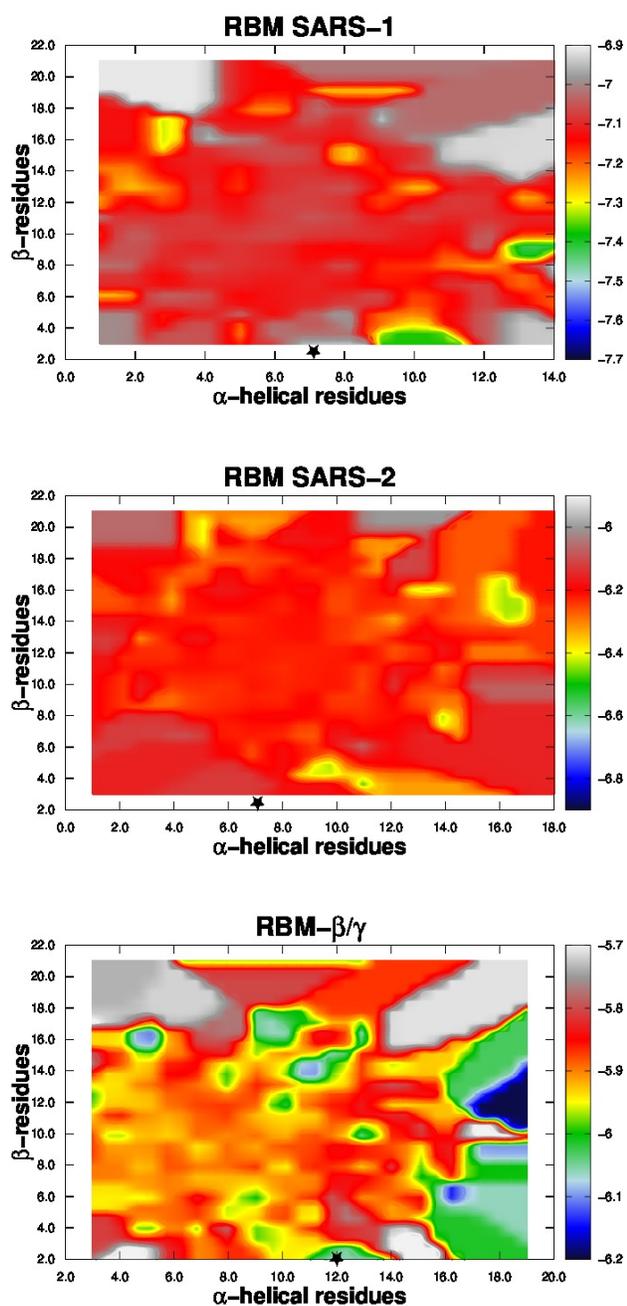


FIG. 4. Isocontours of the intensive internal energy ε (kcal/mol/residue) expressed in terms of the number of residues in α -helical and β -sheet molecular configurations sampled for RBMs of SARS-1, SARS-2 WT and SARS-2 β/γ . The star (\star) shows structural prevalence for I-TASSER initial RBM models.