Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Journal of Computational Physics 444 (2021) 110591

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

Computational modeling of protein conformational changes - Application to the opening SARS-CoV-2 spike

Anna Kucherova [a], Selma Strango [a], Shahar Sukenik [b], Maxime Theillard [a,*]

[a] Department of Applied Mathematics, University of California, Merced, CA 95343, USA
[b] Department of Chemistry and Chemical Biology, University of California, Merced, CA 95343, USA

ARTICLE INFO

Article history:
Available online 26 July 2021

Keywords:
SARS-CoV-2
Covid-19
Spike protein
Molecular trajectory
Poisson-Boltzmann
Multiscale modeling

ABSTRACT

We present a new approach to compute and analyze the dynamical electro-geometric properties of proteins undergoing conformational changes. The molecular trajectory is obtained from Markov state models, and the electrostatic potential is calculated using the continuum Poisson-Boltzmann equation. The numerical electric potential is constructed using a parallel sharp numerical solver implemented on adaptive Octree grids. We introduce novel *a posteriori* error estimates to quantify the solution's accuracy on the molecular surface. To illustrate the approach, we consider the opening of the SARS-CoV-2 spike protein using the recent molecular trajectory simulated through the Folding@home initiative. We analyze our results, focusing on the characteristics of the receptor-binding domain and its vicinity. This work lays the foundation for a new class of hybrid computational approaches, producing high-fidelity dynamical computational measurements serving as a basis for protein bio-mechanism investigations.

© 2021 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Proteins are polymers composed of amino acid chains that are folded into well-defined three-dimensional shapes. The structure and of this shape is crucial for proper protein function. However, in many proteins the three-dimensional structure is not enough, and function must be facilitated by specific intra-protein motions. [15,18,37,39]. The SARS-CoV-2 spike (S) protein is one such protein [15,18,37,39]. Naturally found as a homotrimer, the S-protein contains a buried receptor binding domain (RBD). The RBD has a high affinity to the human angiotensin-converting enzyme 2 (ACE2), and binding mediates viral entry to the cell. However, to facilitate ACE2 binding, the RBD must be exposed through a series of complex motions that occur in the unbound S protein trimer. [2,18,20,33]. Recently, Zimmerman et al. [43] produced the first Markov state models simulation of the SARS-CoV-2 spike opening. This computational *tour de force* involved millions of citizen scientists collaborating through the Folding@home initiative [1], and produced an overall 0.1s of molecular trajectories. For a full characterization of the protein interaction, the molecular trajectory may not be enough.

One aspect that may provide some insight into the interactions and function of a protein is the electrostatic potential it generates. For the SARS-CoV-2 spike, according to a previous study, the affinity constant (derived from the electrostatic potential) for the RBD of SARS-CoV-2 to the ACE2 is 10 to 15 times greater than that of SARS-CoV, potentially contributing to its transmission efficiency [37]. The reason for the higher binding affinity was attributed to several mutations, most notably from the residue Val404, found in SARS-CoV, to the positively charged Lys417 in SARS-CoV-2. This mutation resulted in an

* Corresponding author.
E-mail address: mtheillard@ucmerced.edu (M. Theillard).

https://doi.org/10.1016/j.jcp.2021.110591
0021-9991/© 2021 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

intensified electrostatic potential complementarity between the negatively charged ACE2 binding site and the now more positively charged RBD of SARS-CoV-2 [15,37]. To understand the contribution of protein charge repositioning and also to help inform drug design strategies that leverage this distribution, it is imperative to know how the protein's electrostatic potential changes as it deforms.

The Poisson-Boltzmann equation has long been recognized as the representation of choice to model the electric potential generated by proteins in solvents. It has drawn significant interest from the computational community since the pioneering calculations of Warwicker and Watson in the early 1980s [38], which has lead to the production of a broad variety of numerical solvers [4,3,6,7,13,21,17,22] and open source software [12,16,32], built over the traditional spectrum of numerical methods. Such tools are, for example, employed in the context of drug development and discovery to calculate solvation free energies [10,14,28]. Using massively parallel architectures [11], these calculations can be carried out on considerably large proteins, such as the entire HIV-1 capsid (4, 884, 312 atoms, Protein database entry: 3J3Q).

In this work, we combine Markov state model simulations with partial differential equation modeling to obtain dynamical electric potential maps of deforming proteins. To illustrate this novel approach, we leverage the S protein opening trajectory created by the Folding@home initiative to examine and characterize the potential of the SARS-CoV-2 S protein during trimer opening. At each frame of the simulated trajectory, we reconstruct the protein surface and calculate the generated electric potential with the approach developed by Mirzadeh et al. [26,27]. Using adaptive non-graded Octree grids and sharp discretizations, we efficiently produce high-fidelity solutions to the non-linear Poisson-Boltzmann equation. In the continuum model, all temporal variations are neglected, and as a consequence, all potential maps are independent, making their computation embarrassingly parallelizable. To verify the accuracy of our method, we construct practical *a posteriori* error estimates for the surface representation and electrostatic properties.

Our study of the electrostatic dynamics of the S protein opening reveals dramatic rearrangements of the electrostatic field during this process. These rearrangements act to localize a negatively charged field towards the interior of the S protein and expose a positive surface of its residue binding domain. This is in line with the negative charge of the target binding region on the ACE2 receptor and may aid the S protein in binding to its target with high affinity. The paper is arranged as follows: in section 2, we present the trajectory used for this study. Next, section 3 entails a thorough mathematical design for the calculation of the potential field that is done on each frame in the trajectory. Section 4 encapsulates the numerical method used along with a convergence study for computational validation. The dynamical electrostatic-geometric properties of the spike protein example are characterized in section 5, and conclusions are drawn in section 6.

2. Opening of the SARS-CoV-2 spike protein

The S protein exists primarily in the closed configuration, hiding the three identical RBDs in its core [43]. The Folding@home trajectory focuses on the opening of the unglycosylated, uncleaved spike configuration through the detachment of a single monomer from the spike's core, revealing its associated binding site. A similar (but not identical) monomeric-opening state has recently shown to be populated in $\approx 16\%$ of the unbound spike population in a recent cryoEM study [5].

2.1. Simulated trajectory

The trajectory, from Zimmerman et al. [43] contains a 71-frame sequence that captures the most populated pathway of the spike protein's transition from its closed to open state. This path was calculated utilizing a goal-oriented adaptive sampling algorithm (FAST, [42]) to favorably sample spike opening.

Each frame consists of a list of $N = 51,671$ atoms, represented by their positions $\mathbf{x}_{i=1..N}$, radius $r_{i=1..N}$ and fixed partial charge $z_{i=1..N}$. Throughout these frames, the protein undergoes continuous deformations, which shift the receptor-binding domain of one of the monomers from being hidden while the spike is in its closed conformation to being fully exposed once the spike opens. Fig. 1*a* and 1*b* represent the protein's molecular structure in the open and closed configurations. Each one of its chains is depicted in a different color to illustrate the protein's trimeric structure. The revealed receptor-binding domain is located at the opening extremity of the red chain (see Fig. 1*a*, 1*b*).

2.2. Relative opening measurement

To provide preliminary quantification of the spike opening, dozens of atom pairs on opposing sides of the spike were chosen, and the distance between them measured as the spike transitions from one conformation to another (i.e. as a function of frame) (Fig. 1*a* and *b*). One of the atoms in the pair was chosen from the binding interface, the monomer depicting in red in Fig. 1(*a* and *b*), while the second atom in the pair was chosen across the top of the spike opening. Although all pairs presented the same general trajectory, four of these atom pairs were arbitrarily selected as a subset for illustration.

The resulting relative distance between the atom pairs is measured in Angstroms (Å), as shown in Fig. 1*c*. The lines labeled 1, 2, 3, and 4 refer to the behavior of the four-pair subset, with their average behavior depicted in black. All four measurements generally follow the same trend. This extension happens continuously outside of frames 50 to 55, which show abrupt variations.

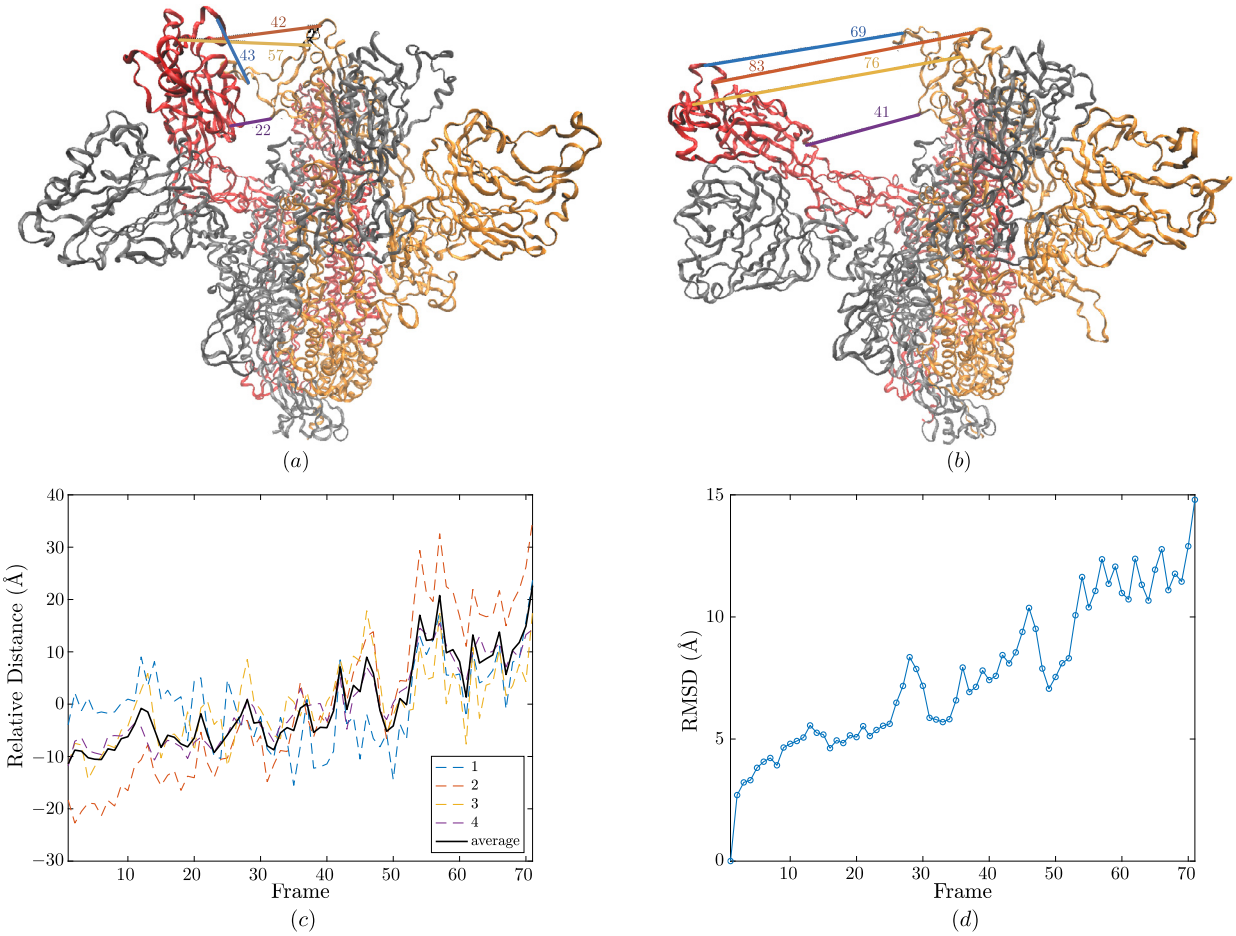

Fig. 1. SARS-CoV-2 Spike protein visualized in its closed (*a*) and open (*b*) states. The four pairs of atoms, along with the separating distances, used to quantify the gap opening are represented with straight solid colored lines. (*c*) The relative distance (Å) between arbitrarily chosen atom pairs is represented as a function of time. One atom is selected on the receptor-binding domain (RBD), and the other opposite the RBD across the opening of the spike protein. This distance is relative to the average distance between each respective atom pair (1, 2, 3, and 4 in the graph), hence the term' relative distance'. (*d*) The root mean squared deviation (RMSD) is calculated as a function of time with frame 1 (closed state) as the reference frame. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

2.3. Magnitude of the conformation change

To measure the total structural deformation, we compute the root-mean-square deviation (RMSD). It measures the average distance between atoms in the current position and some reference configuration, defined here as the initial closed configuration. Specifically

$$RMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N} |\mathbf{x}_i^n - \mathbf{x}_i^1|^2}, \tag{1}$$

where the superscripts n and 1 are for the current and initial states. Despite large variations at the initial and final stages, the RMSD evolution (depicted in Fig. 1*d*), shares similar features with the evolution of the relative opening. These similarities suggest that the significant transformations that occur during the opening are recapitulated by the four distances in Fig. 1*c*.

3. Continuum modeling

This section describes the reconstruction of the potential map at each iteration from the current protein structure using the non-linear Poisson-Boltzmann (PB) equation. Because of the way the trajectory has been obtained, all frames are independent. Thus, the potential maps are decoupled and can be computed separately. A more comprehensive model, such as the Poisson–Boltzmann–Nernst–Planck model [41,25], would consider the diffusion of the ions and introduce time derivatives in the partial differential equations. Because at the atomic scale ($L = 1$ Å), for typical diffusivities ($D \approx 10^{-9}$ m^2 s^{-1}),

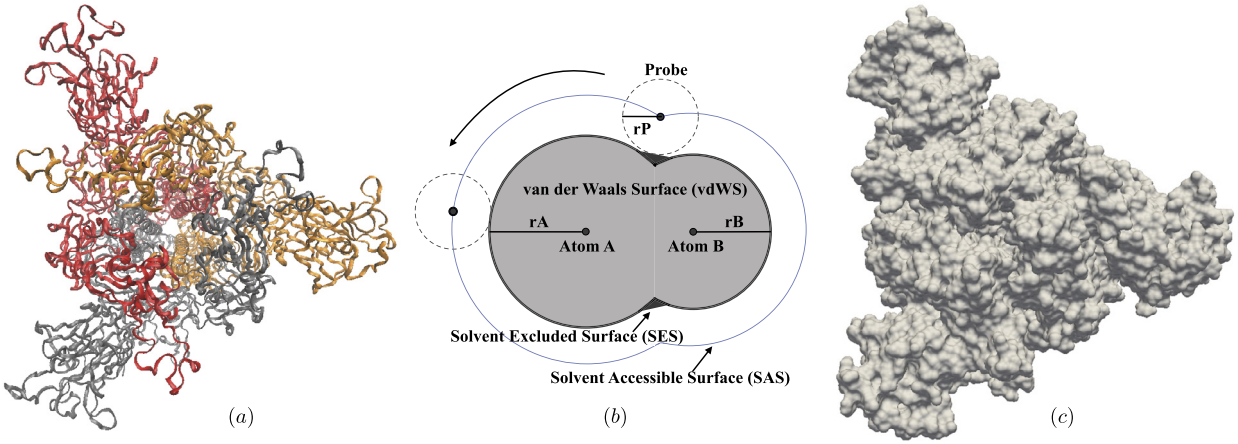

Fig. 2. (*a*) Visualization of the top view of the Spike protein molecule (colored by monomer). Color denotes the three monomers (*b*) A cutaway diagram of a general molecular surface depicting the Van der Waals Surface (vdWS), solvent-excluded surface (SES), and solvent-accessible surface (SAS). (*c*) SES representation of the Spike protein oriented the same way as Fig. (*a*) (top view).

the ionic diffusion time scale ($\tau_D \approx \frac{l^2}{D} = 10^{-11}$ s) is orders of magnitude smaller than the estimated opening duration ($\tau_O >> 10^{-9}$ s), these effects can be neglected and the static PB equation is a pertinent model. Readers not interested in the numerical specifics can skip the following sections (3 and 4).

3.1. Molecular surfaces

A common way of portraying a molecular surface is by use of the van der Waals Surface depiction (see Fig. 2b). Each atom in a molecule is depicted by a sphere with location $\mathbf{x}_{i=1\ldots N}$ and radius $r_{i=1\ldots N}$ that is defined by an isosurface on their electron density. The union of these spheres, depicted in gray in Fig. 2, forms the van der Waals Surfaces (vdWS) of that molecule. The nature of the vdWS means that some regions might be identified as being exposed to the solvent, while in fact, their geometry makes them inaccessible to the solvent particles. As a consequence, we define the Solvent Accessible Surface (SAS [19]) of the protein, depicted as the blue outline in Fig. 2 b; it is formed through the addition of the solvent's particle radius, r_P, to each $r_{i=1..N}$ resulting in a buffer around the vdWS. While the puffed-up SAS depiction may be useful for some areas of study, the Solvent Excluded Surface (SES) is preferable when discussing surface details of a molecule [31], as shown in Fig. 2b. As the name suggests, the inner molecule region defined by this surface includes all locations the solvent cannot occupy, including the vdWS and the tiny crevasses on its exterior, shown as concave black triangles at the meeting of atoms A and B in Fig. 2b.

3.2. Mathematical representation and numerical construction

To represent the biomolecule, we employ the level set method [30] and capture the SES location as the zero level set of an auxiliary field $\phi(\mathbf{x})$ defined over the domain of interest Ω. The solvent Ω^-, the Solvent Excluded Surface Γ, and the inside of the molecule Ω^+ are defined as

$$\Omega^- = \{\mathbf{x} \in \mathbb{R}^3 | \phi(\mathbf{x}) < 0\} \tag{2}$$

$$\Gamma \quad = \{\mathbf{x} \in \mathbb{R}^3 | \phi(\mathbf{x}) = 0\} \tag{3}$$

$$\Omega^+ = \{\mathbf{x} \in \mathbb{R}^3 | \phi(\mathbf{x}) > 0\} \tag{4}$$

The normal $\mathbf{n}$ to the interface Γ is defined as pointing toward Ω^+. It is calculated as the normalized level set gradient

$$\mathbf{n} = \frac{\nabla\phi}{|\nabla\phi|}. \tag{5}$$

The level set function is constructed following the approach proposed by Mirzadeh et al. in [27]. We start by constructing the level set function $\phi_{SAS}(\mathbf{x})$ representing the SAS as

$$\phi_{SAS}(\mathbf{x}) = \max_{i=1..N} \left(r_i + r_p - |\mathbf{x} - \mathbf{x}|\right). \tag{6}$$

The SAS level set function is then reinitialized to be a signed-distance function (*i.e.* $|\nabla\phi_{SAS}| = 1$) by solving the reinitialization equation in fictitious time τ until the steady state is reached

Table 1
Problem parameters and non-dimensional numbers.

Parameter	Symbol	Definition	Value
Characteristic length	L	-	1 Å
Characteristic potential (thermal voltage)	ψ_0	$\frac{k_B T}{e}$	2.570×10^{-2} V
Vacuum permittivity	ϵ_0	-	8.854×10^{-12} $\mathrm{F m^{-2}}$
Spike relative permittivity	ϵ^+	-	2
Solvent relative permittivity	ϵ^-	-	7.830×10
Boltzmann constant	k_B	-	1.381×10^{-23} $\mathrm{J K^{-1}}$
Avogadro number	N_A	-	6.022×10^{23} $\mathrm{M^{-1}}$
Elementary charge	e	-	1.602×10^{-19} C
Temperature	T	-	298.15 K
Salt concentration	c^b	-	1×10.0^{-3} $\mathrm{Mm^{-3}}$
Probe radius	r_p	-	1.4 Å
Valence of the background electrolyte	z	-	1
Non-dimensional inverse of Debbye length	κ_D	$\sqrt{\frac{2c^b N_A e^2 z}{\epsilon_0 k_B T L^2}}$	9.207×10^{-3}
Non-dimensional Bjerrum length	λ_B	$\frac{e^2}{k_B T \epsilon_0 L}$	7.039×10^3

$$\frac{\partial \phi_{SAS}}{\partial \tau} + \mathrm{sign}\,(\phi_{SAS})(|\nabla \phi_{SAS}| - 1) = 0, \qquad \forall \mathbf{x} \in \Omega. \tag{7}$$

From the reinitialized function $\mathcal{R}(\phi_{SAS})$, the SES level set function is then obtained as

$$\phi(\mathbf{x}) = \mathcal{R}(\phi_{SAS})(\mathbf{x}) - r_p, \qquad \forall \mathbf{x} \in \Omega. \tag{8}$$

As it was pointed out in [27], this procedure can create non-physical inner cavities. They are identified by finding physical points disconnected from the contour of the computational domain $\partial\Omega$ where the level set function $\phi(\mathbf{x})$ is positive. In practice, such points are isolated by solving the following Laplace problem

$$-\triangle c = 0, \qquad \forall \mathbf{x} \in \Omega^-, \tag{9}$$

$$c = 0, \qquad \forall \mathbf{x} \in \Gamma, \tag{10}$$

$$c = 1, \qquad \forall \mathbf{x} \in \partial\Omega, \tag{11}$$

and detecting where $\phi(\mathbf{x}) < 0$ and $c = 0$. The cavities are removed by switching the sign of the level set function at these problematic positions. Finally, for computational purposes the SES level set is systematically reinitialized.

3.3. Poisson-Boltzmann equation

The electrostatic potential, Ψ, around a biomolecule immersed in a binary z:z electrolyte solution can be described by the following non-linear Poisson-Boltzmann (PB) equation,

$$-\nabla \cdot (\epsilon \epsilon_0 \nabla \Psi) + 2c^b(\mathbf{x}) N_A e z \sinh\left(\frac{e\Psi}{k_B T}\right) = \sum_{i=1}^{N} q_i \delta(\mathbf{x} - \mathbf{x}_i), \qquad \forall \mathbf{x} \in \Omega \setminus \Gamma, \tag{12}$$

where ϵ is the relative permittivity, being equal to ϵ^+ in the molecule (Ω^+) and ϵ^- in the electrolyte (Ω^-). ϵ_0 is the permittivity of a vacuum, N_A is the Avogadro number, c^b is the bulk salt concentration, e is the elementary charge, k_B is the Boltzmann constant, T is the temperature, z is the valence of the background electrolyte, q_i and x_i are the partial charge and position of the i^{th} atom respectively, N is the total number of atoms in the molecule. $c^b(x) = 0$ inside the molecule. In non-dimensional form, the Poisson-Boltzmann equation takes the following form

$$-\nabla \cdot (\epsilon \nabla \psi) + \kappa_D^2(\mathbf{x}) \sinh(\psi) = \sum_{i=1}^{N} \lambda_B z_i \delta(\mathbf{x} - \mathbf{x}_i), \qquad \forall \mathbf{x} \in \Omega \setminus \Gamma \tag{13}$$

where the characteristic length L is chosen to be 1 Å, the potential is scaled by the thermal voltage ($\frac{k_B T}{e}$), z_i is the non-dimensional partial charge on the i^{th} atom, and $\kappa_D = \sqrt{\frac{2c^b N_A e^2 z}{\epsilon_0 k_B T L^2}}$ is the non-dimensional inverse of the Debye length. Inside the molecule, κ_D is null. The constant $\lambda_B = \frac{e^2}{k_B T \epsilon_0 L}$ is the non-dimensional Bjerrum length. The above non-dimensional PDE is completed with the following jump conditions on non-dimensional potential

$$[\psi]_\Gamma = 0, \qquad [\epsilon \nabla \psi \cdot \mathbf{n}]_\Gamma = 0, \qquad \forall \mathbf{x} \in \Gamma, \tag{14}$$

where the jump operator is defined for any quantity ζ defined in both domain as $[\zeta]_\Gamma = \zeta^+ - \zeta^-$. All parameters and non-dimensional numbers, along with their values for this study, are summarized in Table 1.

3.4. Solution decomposition

Following [27], we treat the singularities arising in the solution due to the singular charges inside the molecules by using the decomposition proposed by Chern et al. [8]. Doing so we split the non-dimensional potential ψ, into regular and singular parts: $\hat{\psi}$ and $\bar{\psi}$, respectively

$$\psi = \hat{\psi} + \bar{\psi}. \tag{15}$$

The singular part is itself split into two parts ψ^* and ψ_0

$$\bar{\psi}(\mathbf{x}) = \begin{cases} \psi^*(\mathbf{x}) + \psi_0(\mathbf{x}) & \text{if } \mathbf{x} \in \Omega^+, \\ 0 & \text{if } \mathbf{x} \in \Omega^- \end{cases} \tag{16}$$

where ψ^* is the Coulombic potential due to singular charges,

$$\psi^*(\mathbf{x}) = \frac{\lambda_B}{4\pi\epsilon^+} \sum_{i=1}^{N} \frac{z_i}{|\mathbf{x} - \mathbf{x}_i|}, \tag{17}$$

and ψ_0 satisfies the following Poisson's problem

$$\nabla\psi_0 = \quad 0, \qquad \forall \mathbf{x} \in \Omega^+, \tag{18}$$

$$\psi_0 = -\psi^*, \qquad \forall \mathbf{x} \in \Gamma. \tag{19}$$

Utilizing the decomposition shown above, the regular (non-singular) part of the solution is given by solving

$$-\nabla \cdot (\epsilon \nabla \hat{\psi}) + \kappa^2(\mathbf{x}) \sinh(\hat{\psi}) = 0, \qquad \forall \mathbf{x} \in \Omega \setminus \Gamma, \tag{20}$$

subject to the following jump conditions:

$$[\hat{\psi}]_\Gamma = 0, \qquad \forall \mathbf{x} \in \Gamma \tag{21}$$

$$[\epsilon \nabla \hat{\psi} \cdot \mathbf{n}]_\Gamma = -\epsilon^+ \nabla(\psi^* + \psi_0) \cdot \mathbf{n}|_\Gamma, \qquad \forall \mathbf{x} \in \Gamma \tag{22}$$

Since ψ^* is known analytically, the gradient, $\nabla\psi^*$, appearing in the right hand side of (22) can be computed exactly

$$\nabla\psi^* = -\frac{\lambda_B}{4\pi\epsilon^+} \sum_{i=1}^{N} z_i \frac{\mathbf{x} - \mathbf{x}_i}{|\mathbf{x} - \mathbf{x}_i|^{3/2}}, \tag{23}$$

while $\nabla\psi_0$ must be numerically approximated.

4. Numerical method

The numerical approach for the resolution of the above Poisson-Boltzmann problem is presented in this section, along with novel practical *a posteriori* error estimates. We then verify the method for the entire trajectory by conducting a systematic convergence study for these estimators.

4.1. Implementation

The numerical method, implemented on non-graded adaptive Octree grids, follows the general description given in [27], with each snapshot of the protein treated independently. The surface and grid generation is done using the level set framework developed by Min and Gibou [23]. In particular, as Fig. 3 illustrates, the mesh is systematically adapted to the SES location. All quantities are stored at the nodes of the mesh for improved accuracy and facilitated manipulation. The numerical solutions of the Poisson systems (9)-(10)-(11) and (18)-(19) (for the cavities detection and the construction of the regular part of the solution ψ_0 respectively) are obtained using the second-order approach presented in [35], itself based on [24]. The solution, $\hat{\psi}$, to the problem defined by Eqs. (20), (21) and (22) is constructed using a nodal version of the jump solver presented in [34]. The non-linearity of Eq. (20) is addressed using Newton's method [8,26,27], with a relative error tolerance of 10^{-6}, chosen to be orders of magnitude smaller than the desired overall numerical error. The whole method is parallelized in a shared memory fashion using OpenMP [9,29].

The entire computational domain is defined as the cube of side length 400 Å (about twice the size of the whole protein) and center $\mathbf{x}_c = \frac{1}{N}\sum_{i=1}^{N} \mathbf{x}_i$. Calculations were performed on Octrees of maximum level ranging from 7 to 11. On the finest grids, the minimal spatial resolution is 0.18 Å.

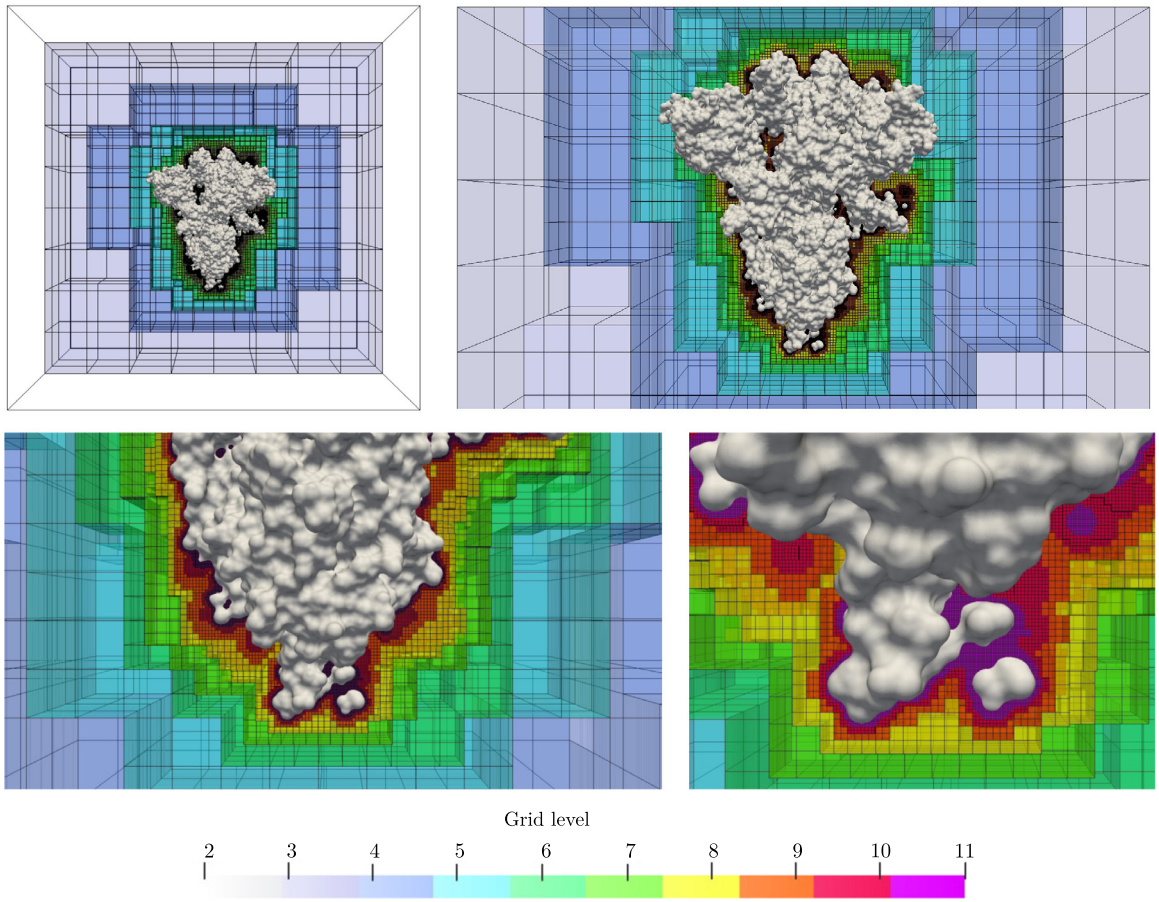

Fig. 3. Visualizations of the entire computational domain, SARS-CoV-2 spike protein (frame 1), and finest mesh used for this study (level 11 with 34, 362, 796 nodes). Computational cells are colored by their corresponding tree level (*i.e.* the number of successive mesh subdivisions required for their construction). For visual purposes only half of the mesh is depicted.

4.2. Error estimates

To monitor the convergence of the overall method, we construct *a posteriori* numerical error estimators. They only rely on the decomposition presented in section 3.4, and can therefore be employed independently of the numerical approach. For the current implementation, detailed in section 4, we refer the interested reader to [26,27] for formal convergence studies, using analytic solutions and order estimations.

The error on the interface representation, gradient of the solution (*i.e.* electric field) and solution itself can be estimated using the following metrics e_Γ, e_E, e_ψ

$$e_\Gamma = \frac{1}{|\Gamma|}\left| \int_\Gamma \nabla\psi^* \cdot \mathbf{n} - \sum_i^N \frac{\lambda z_i}{4\pi\epsilon^+} \right|, \tag{24}$$

$$e_E = \frac{1}{|\Gamma|}\left| \int_\Gamma \nabla\psi_0 \cdot \mathbf{n} \right|, \tag{25}$$

$$e_\psi = \frac{1}{|\Gamma|} \int_\Gamma \left| \psi^* + \psi_0 \right|, \tag{26}$$

where $|\Gamma|$ denotes the surface area of Γ. In virtue of Gauss's Theorem the first two integrals are null. Because of the boundary condition (19), the third quantity should also be null. When computing e_Γ, because the gradient of the Coulombic and the total charge are known exactly, numerical errors can only arise from calculating the local normal $\mathbf{n}$ or approximating the surface integral. Therefore, this metric focuses on the geometry and its manipulation only.

The numerical errors in the gradient of the component ψ_0 are the primary source of errors in the second metric e_E. Thus, it can be interpreted as a lower bound estimate for the average total normal electric field on the interface. The metric

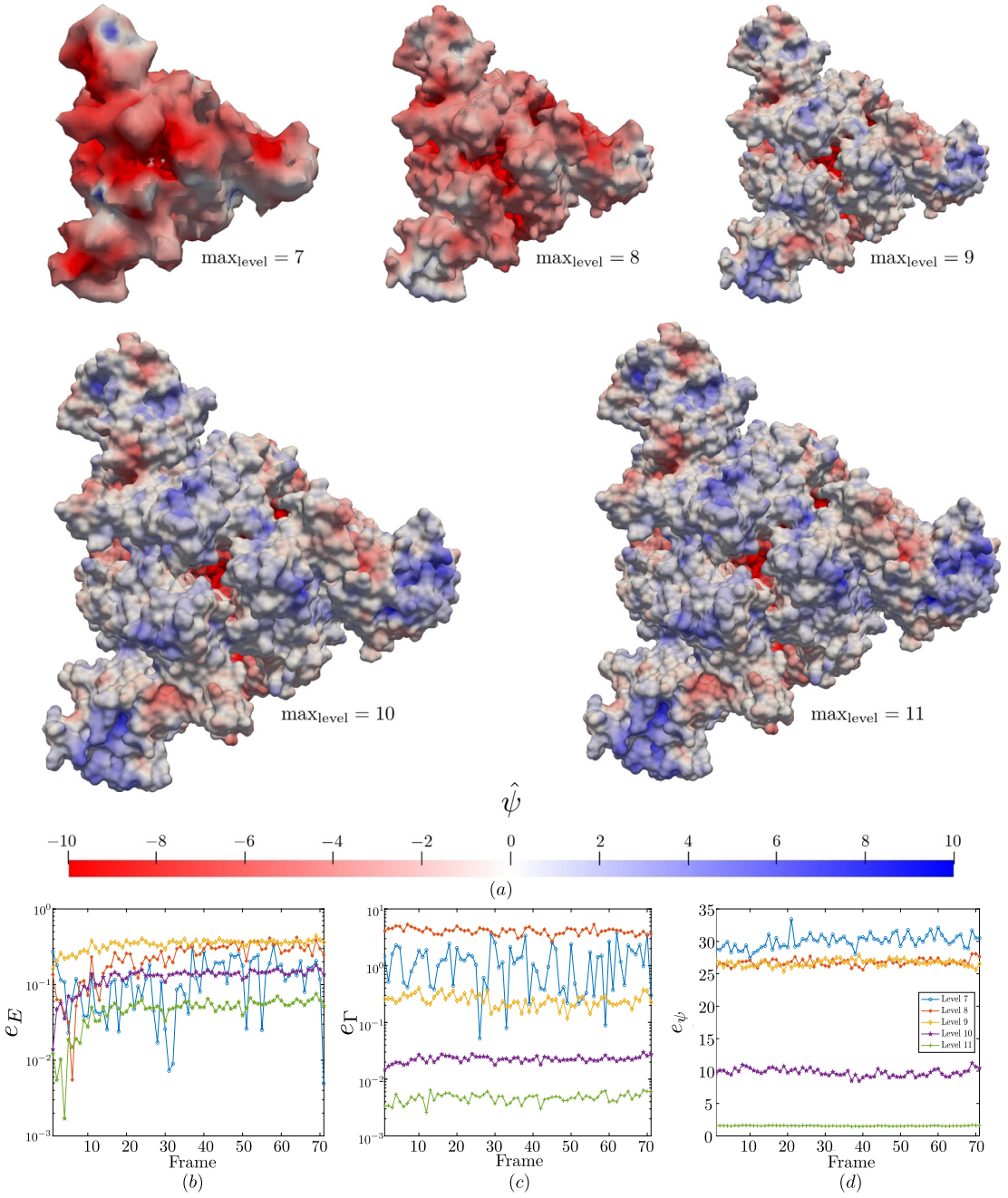

Fig. 4. (*a*) Evolution of the spike's Solvent Excluded Surface and electrostatic potential as the maximal resolution increases. (top down view in the closed configuration (frame 1)). ($b-d$) Convergence of the interface representation (e_Γ), surface normal electric potential (e_E), and surface potential (e_ψ).

e_ψ, the average error in ψ_0 on the interface, can similarly be used to estimate the numerical error in the total potential on the interface.

4.3. Convergence study

Fig. 4*a* depicts the spike protein's electrostatic potential at the initial frame (closed configuration) as the mesh is refined. Positive electrostatic potential is colored in blue, neutral (no charge) in white, and the negative electrostatic potential is shown in red. The coarsest simulations ($\max_{\text{level}} = 7, 8$) are only able to reproduce the general structure of the protein but fail at creating an accurate electrostatic map. As the maximal resolution reaches the characteristic atomic radius ($r_{\min} = 1$ Å), the finest geometrical features are correctly reproduced, leading to appreciably more accurate results ($\max_{\text{level}} = 9$, minimal

resolution 0.79 Å). Further increasing the spatial resolution refines these molecular structures and the small scale potential variations even more ($\max_{\text{level}} = 10, 11$).

The time evolution of our three error estimates for all examined resolutions is presented in Fig. 4. As expected, all three metrics converge with increasing resolution. The impact of using subatomic resolution (*i.e.* $\max_{\text{level}} \geq 9$) is well illustrated with the convergence of the electric field and potential error estimates: it is unclear for super-atomic resolutions ($\max_{\text{level}} = 7, 8$), and evident for subatomic ones ($\max_{\text{level}} = 9, 10, 11$).

The error estimation for the total potential (e_ψ) is significantly larger than the variations between consecutive maximum levels observed in Fig. 4*a*, which is a proxy for the error on the regular part of the solution $\hat{\psi}$. Since e_ψ involves the singular part of the solution, which exhibits large spatial variations over small length scales, it is prone to higher numerical errors and expected to be larger than the actual error on the regular part of the solution. Closer inspection of our measurements reveals that these two metrics may differ by at least one order of magnitude. The error estimation for ψ, using $\max_{\text{level}} = 10$ is approximately equal to the maximum absolute surface potential value (≈ 10) observed on Fig. 4*a*. This misleadingly suggests the relative error on the total solution is as large as 100%. The comparison between the potential maps for $\max_{\text{level}} = 10$ and $\max_{\text{level}} = 11$ indicates that in practice the relative error on the surface potential probably lies between 1% and 10%.

From all these remarks, we are confident that the method is correctly implemented and that the most refined simulations accurately capture the S protein's electrostatic potential. For the finest resolution, the estimated average error on the total potential is close to 2, which in light of the above discussion, suggests that the average practical relative error on the surface potential is under 2%.

5. Results

The S protein remains predominantly in the closed conformation to mask its receptor-binding domains (RBDs), thereby impeding their binding. To bind with ACE2, the S protein transforms into its open conformation, revealing its binding interface.

In describing our results, we refer to the part of the spike containing the three RBDs as the top part and the predominantly negatively charged portion binding to the virus membrane as the bottom part. [43].

Our simulations (see Fig. 5*a*) illustrate that the top part of the spike protein is predominantly positively charged, in accordance with the negative charge of the ACE2 binding site (see Fig. 6). Surprisingly, the top of the spike also reveals an underlying core with a surface area that has a dense negative charge. As the spike opens, this area could be exposed to the solvent, generating a negatively charged electrostatic cloud in the upper part of the protein, which may repulse the negatively charged ACE2 receptor. Therefore, we characterize the dynamics of the geometrical and electric properties of this negatively charged core χ (Fig. 5*b*), the resulting repulsive cloud $\mathcal{C}$ (Fig. 5*c*), the binding site $\mathcal{B}$ (Fig. 6*c*), and the far electric field of the spike protein as they shift from the dynamics that occur during spike opening.

5.1. Negatively charged core

We define the negatively charged core χ, illustrated in Fig. 5*b*, as the area of the SES in the top of the protein where the electric potential is more negative than a threshold value $c_\chi = -5$. Fig. 5 *c* and *d* depict the evolution of its surface and average charge. As the spike opens, the area shrinks by about 50%, while the variation in the average potential remains small ($\approx 10\%$). In both cases, the most significant variations happen during the first ten iterations. In comparison, the molecular structure analysis (see Fig. 1) displays continuous variation over the entirety of the trajectory, indicating that the structural and electro-geometric transformations are non-trivially coupled. The diminution of both quantities could be explained by the fact that the spike opening removes a barrier, one of the monomers, between the core and solvent, causing the dilution of the surface charges in the solvent, and therefore a contraction of the core and its charge.

5.2. Repulsive electric cloud

We define the negatively charged electric cloud $\mathcal{C}$, generated by the core χ, as the region in the upper part of the solvent where the potential is below the threshold value $c_\mathcal{C} = -2$ (see Fig. 5*c*). The measurements presented in Fig. 5*d* and *e* indicate $\mathcal{C}$ expands by $\approx 42\%$ as the spike opens, while its average potential decreases in magnitude by 30%. Again we interpret this phenomenon as an effective dissolution of the negative charges induced by removing the protecting monomer.

For both geometries χ and $\mathcal{C}$, we observe large variations in the first ten frames. The abrupt change (around frame 50) in the protein structure, discussed in section 2, is only perceptible in the size variations of the electric cloud $\mathcal{C}$.

5.3. Binding site characterization

The interface between the SARS-CoV-2 receptor-binding domain (RBD) and ACE2 are of particular interest as the binding of the two facilitates virus entry into cells [2,18,20,33]. Lan et al. [18] undertook the task of determining the residues of the RBD and ACE2 interface which form a connection, finding the location of the binding site on the spike, $\mathcal{B}$, to consist of 17 residues between Lys417-Tyr505 and the ACE2 site to be formed of 20 residues between Gln24-Tyr83 and Asn330-Arg393.

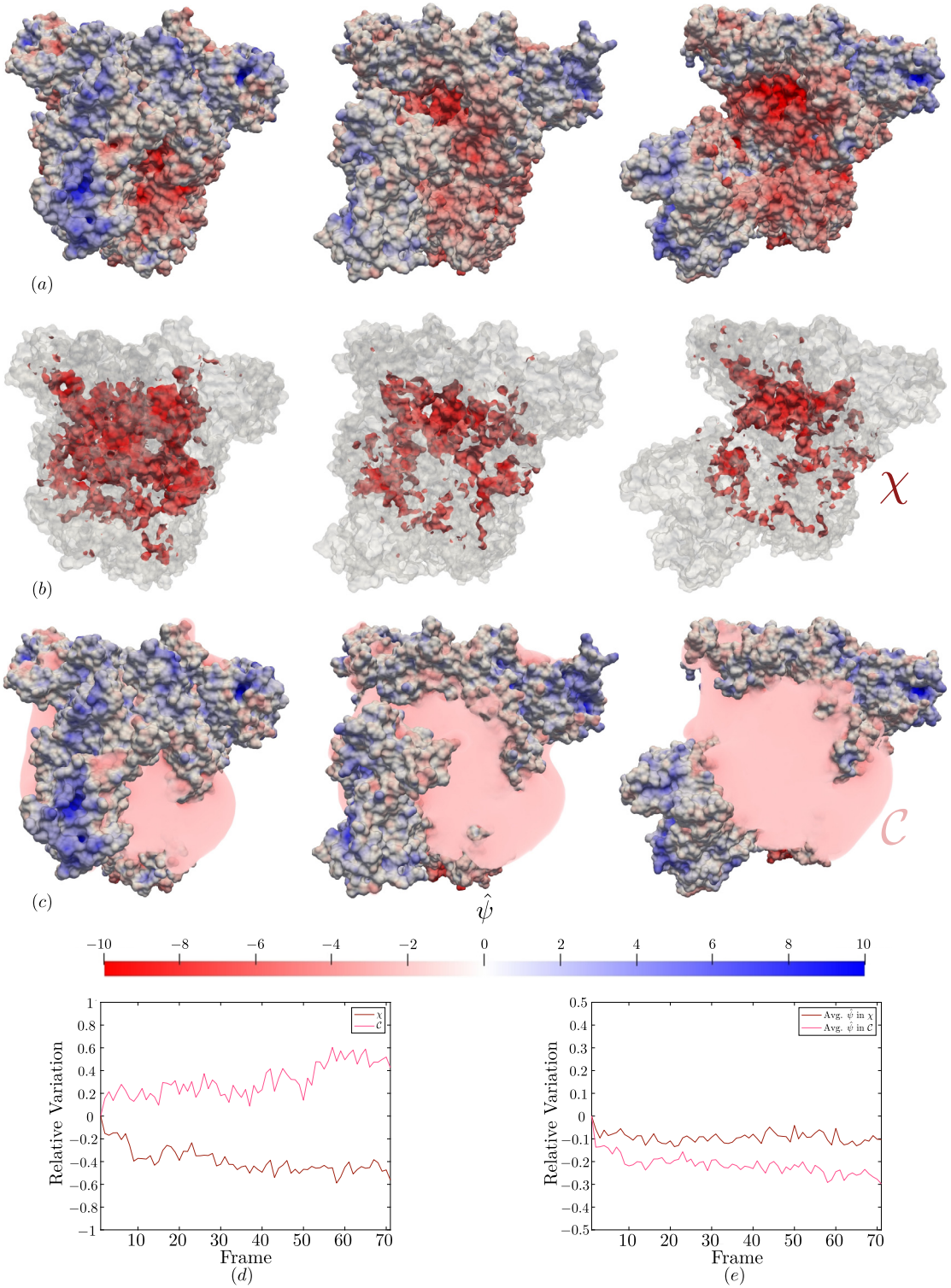

Fig. 5. (*a*) Potential map of the SARS-Cov-2 spike open (frame 1), intermediate (frame 35), and close configuration (frame 71). (*b*) Evolution of the negatively charged core as the spike opens. (*c*) Repulsive electric cloud generated by the negatively charged core. ($d-e$) Characterization of the negatively charged core χ and repulsive electric cloud $\mathcal{C}$: relative size (*d*) and potential (*e*) variations over the entire molecular trajectory.

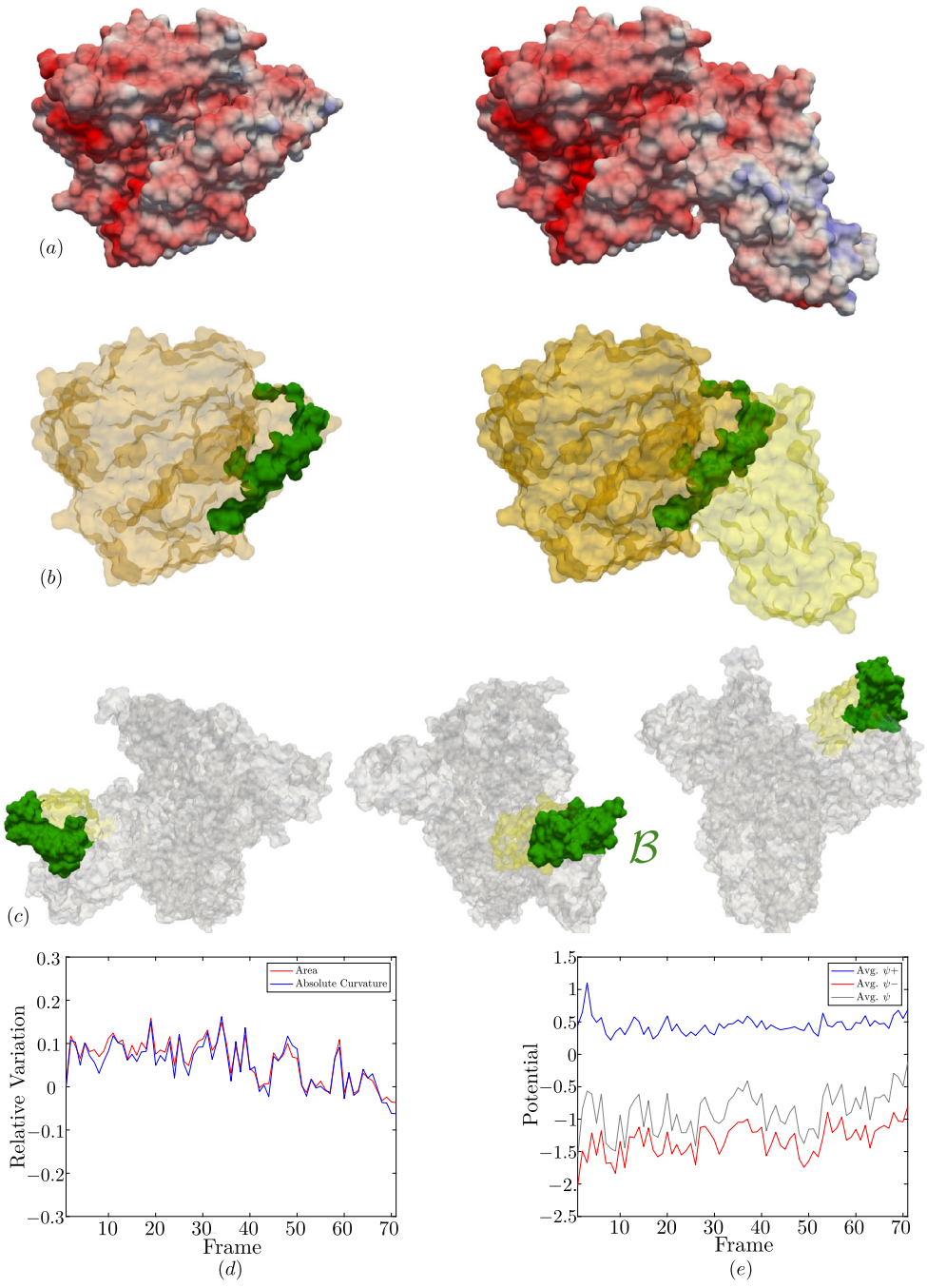

Fig. 6. (*a*) Portion of the ACE2 receptor displaying surface charge (left). ACE2 and RBD of the spike protein connected (right). (*b*) Binding site of ACE2 highlighted in green (left). The binding site of ACE2 is highlighted within the ACE2 (orange) and RBD (yellow) connection (right). (*c*) SARS-Cov-2 spike in open conformation with RBD (yellow) and binding site, $\mathcal{B}$ (green) highlighted. (*d* – *e*) Characteristics of the binding site, $\mathcal{B}$, of the spike protein. (*d*) Relative variance of the area and absolute curvature along with average potential. (e) average positive and negative potential.

For this analysis, we define the binding site, $\mathcal{B}$, for the spike and the ACE2 receptor as the portion of the proteins SES generated by the residues sequences $[417-505]$ and $[24-42]\cup[79-83]\cup[330]\cup[354-357]\cup[393]$ respectively. The spike RBD is generated by the sequence $[333-527]$. When the spike is in the closed position $\mathcal{B}$ is located near its center, but as the spike shifts to its open conformation, $\mathcal{B}$ is moved outward and exposed to the solvent, as seen in Fig. 7. As the spike opens, we monitor the evolution of the binding site area, average absolute curvature, and potential. The average is computed over the binding site surface. We interpret the average absolute curvatures as measures of the global convexity of

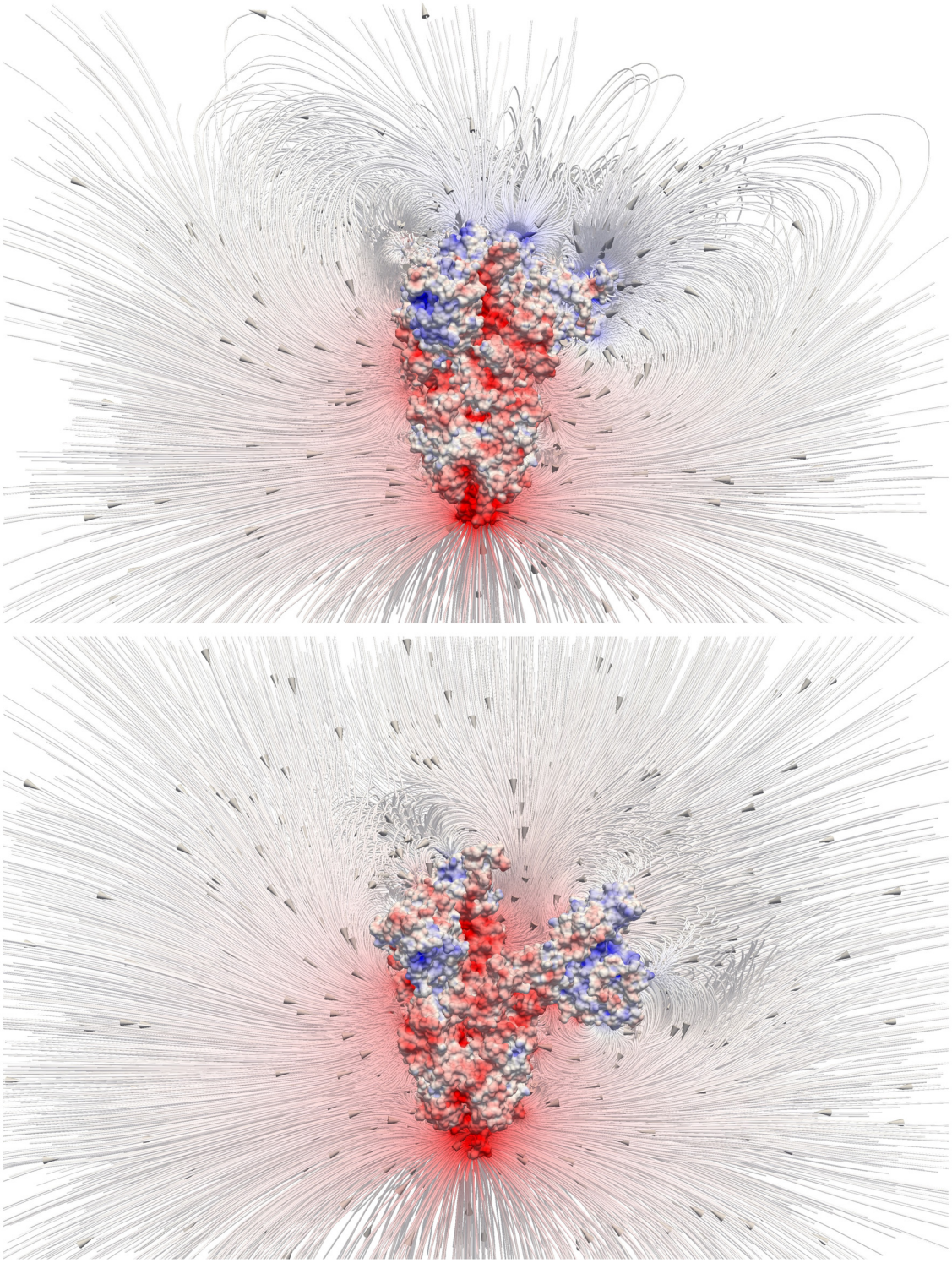

Fig. 7. Streamlines and direction of the electric field in the close (top) and open (bottom) configurations. As the spike opens, the streamlines emanating from the upper part of the molecule are observed to open as well. This effect is particularly noticeable for the lines emerging from the RBD domain (opening extremity on the right).

the binding site. For the potential evolution, we distinguish between the average potential, average positive potential, and average negative potential.

Fig. 6 provides a depiction of the resulting information. A relative decrease is observed for the area and average absolute curvature: as the spike opens, the binding site shrinks and flattens, in each case by $10-15\%$, suggesting minor conformational changes. The average negative potential also experiences a decrease, becoming less negative as the spike opens. Meanwhile, the average positive potential remains relatively constant, resulting in the total average potential of $\mathcal{B}$ following very closely with the changes in average ψ^- and ultimately nearing 0 in the final frame. This increase in average potential is consistent with the knowledge that the ACE2 receptor is designed to bind to negative charges, and may be necessary in directing the ACE2 binding to the correct region.

5.4. Electric field structure

The spike electric far-field streamlines are depicted in Fig. 7. In the closed configuration, most streamlines emanating from the bottom part of the spike are pointing away from the spike. Only a few of them, caused by the rare presence of positive charges, return rapidly to the protein. In the upper part, the streamlines are predominantly closed, suggesting that in this configuration, the spike protein may not be able to attract charges of any polarity at long range, and therefore has limited attracting potential. This may cause a lower affinity. As the spike opens, the streamlines in the bottom remain unchanged. However, above the top of the protein, the streamlines are now predominantly open.

Fig. 8(a)-(b) depicts the entire electric field structure along with the region in the solvent where the electric field point away from the protein, or equivalently where negative charges would be drawn to the spike. In the closed configuration, this region is predominantly located above the spike, while in the open configuration, it is split into three sub-regions, each centered around one of the RBDs. In the latter, the sub-region centered around the revealed RBD represents 45% of the entire region and carries 34% of its total electric potential energy.

To characterize the restructuring of the electric field of the S protein, we define the first four moments of the non-dimensional charge distribution $q(\mathbf{x}) = \sinh(-\psi(\mathbf{x}))$ in the solvent

$$\text{Total Charge} \quad C = \int_{\Omega^-} q(\mathbf{x}) d^{(3)}\mathbf{x}, \tag{27}$$

$$\text{Dipole} \quad D_i = \int_{\Omega^-} q(\mathbf{x}) x_i d^{(3)}\mathbf{x}, \tag{28}$$

$$\text{Quadrupole} \quad Q_{ij} = \int_{\Omega^-} q(\mathbf{x}) \left(3x_i x_j - \mathbf{x}\cdot\mathbf{x}\delta_{ij}\right) d^{(3)}\mathbf{x}, \tag{29}$$

$$\text{Octupole} \quad O_{ijk} = \int_{\Omega^-} q(\mathbf{x}) \left(15x_i x_j x_k - \mathbf{x}\cdot\mathbf{x}\left(\delta_{ij}x_k + \delta_{ik}x_j + \delta_{jk}x_i\right)\right) d^{(3)}\mathbf{x}. \tag{30}$$

Note that the last two moments are defined in their symmetric traceless form, as they would be for a standard electrostatic multipole expansion. Because here our continuum model is more complex, it should be reminded that these four moments are not guaranteed to be the ones appearing in the multipole expansion. Nonetheless, they remain a pertinent tool to characterize the structure of the electrostatic solution. Their evolution throughout the opening, illustrated in Fig. 8*c*, show that the strength of the dipole (D) is decreasing, while all other moments norms are increasing. The two most informative moments for the structure of the electric field, the dipole and the quadrupole, are diminished by 9% and increased by 67% respectively. The principal direction of Q associated to the positive eigenvector appears to remain quasi-parallel to the dipole direction. In fact, all four directions (the dipole direction and the three principal directions of Q) appear to undergo the same rotation. The opening also reinforces the anisotropy in the tensor Q, by amplifying the difference between the magnitude of its eigenvalues, compressing one of its directions and effectively turning it into a two-dimensional quadrupole.

6. Conclusions

We examined the SARS-CoV-2 spike protein's transition from its open to closed conformation, observing the protein's electrostatic potential dynamics with a particular focus on its receptor-binding domains. In the Folding@home trajectory we analyzed, one of the three monomers detaches from the core of the complex and becomes visible to the surrounding environment, consistent with recent cryo-EM derived structures [5]. Our results show that despite the dramatic molecular displacement engendered by this strategic repositioning, the geometric and electric properties of the RBD itself remain largely unaltered. Instead, as we describe below, changes emanating from the exposure of the core of the S protein cause a change in the electric fields surrounding the RBD.

Our continuum analysis, both of the surface potential and the volumetric potential in the vicinity of the binding region, points to the existence of an inner negatively charged core on the surface of the spike, which is revealed to the solvent as

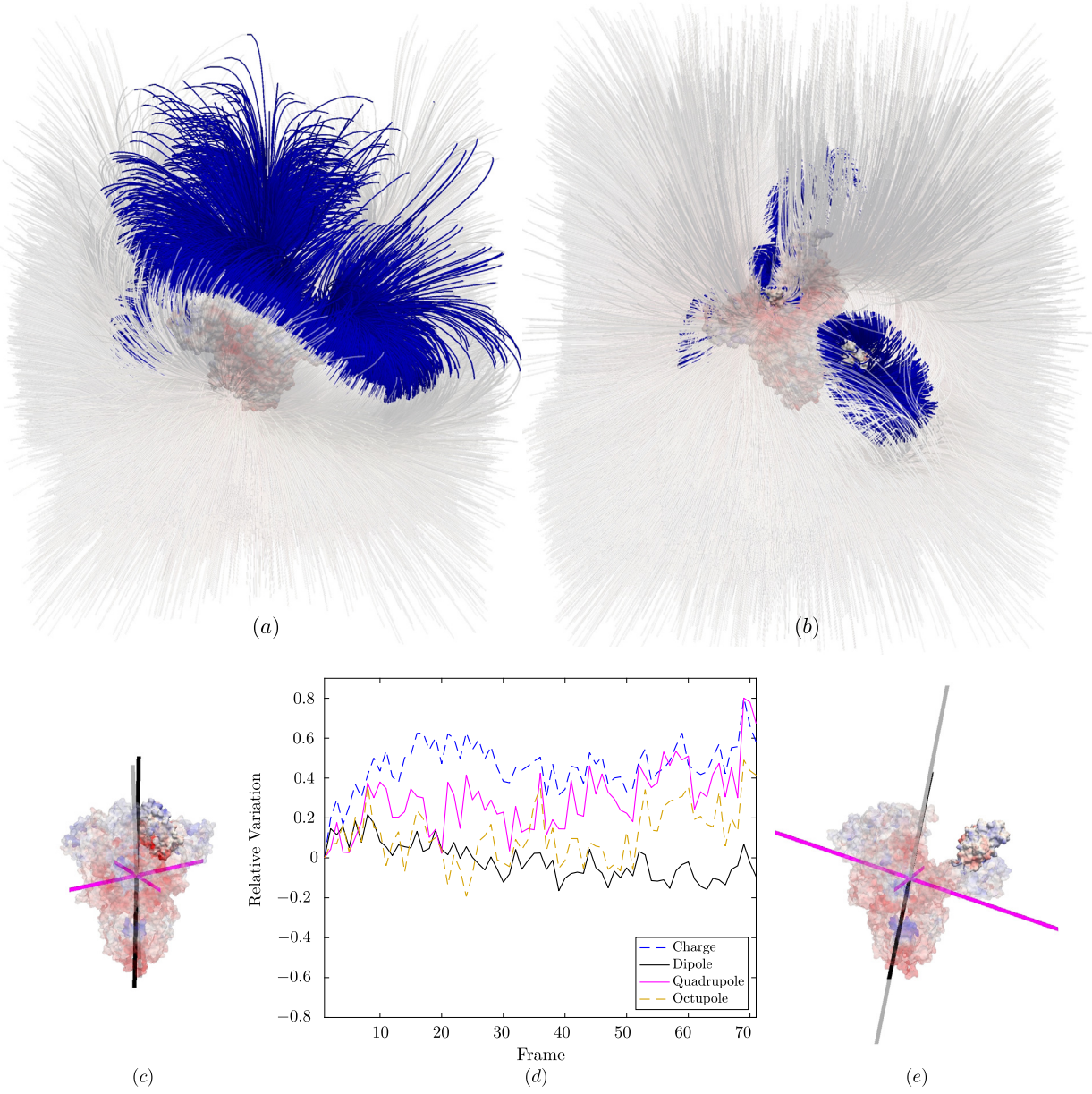

Fig. 8. Streamlines of the electric field ($\mathbf{E} = -\nabla\hat{\psi}$) in the close ($a$) and open ($b$) configurations over the entire computational domain. The portion of the streamlines where the electric field is pointing away from the protein (*i.e.* $\nabla\psi \cdot \mathbf{x} = -\mathbf{E} \cdot \mathbf{x} < 0$), and therefore negative charges would be drawn to it, is depicted in blue. ($c - e$) Electric potential multipole structure: charge density decomposition over its first four moments C, D, Q and O. (c) and (e) represent the principal directions of the dipole D (in black) and the quadrupole Q, in the initial and final configurations, respectively. For the quadrupole, the purple directions (eigenvectors) are associated with negative eigenvalues. The gray one has a positive eigenvalue. All segment lengths are proportional to the strength of the corresponding pole (*i.e.* the norm of the dipole or of the corresponding eigenvalue). The protein binding site is depicted as opaqued. (d) Depicts the evolution of the norm of all four first moments. We use the standard L^2 entry-wise norm.

the spike opens. This negatively charged surface shrinks as the spike opens, inducing a negatively charged region between the exposed RBD and the two hidden ones. The emergence of this electric cloud is *a priori* puzzling: the binding site of the ACE2 receptor being almost entirely negatively charged, we would expect this cloud to repulse the receptor. However, this cloud does not envelop the spike's binding site, $\mathcal{B}$, which becomes more positively charged. This is in line with what we expect is a strong binder of the ACE2 receptor and indicates that targeting the open state of the S protein may be a more viable drug design strategy than the closed configuration. Indeed, recent cryoEM studies showed that $\approx 16\%$ of a recombinantly expressed S protein population is in an open state, with a single monomer "erected" from the core [5].

The electric field, which we observe uncoiling above the protein, exhibits a dramatic rearrangement. This is correlated to the emergence of the negatively charged cloud and manifests in the multipole moments decomposition. In particular,

we observe the quadrupole moment growing in magnitude, rotating, and compressing one of its principal directions. This transformation can be interpreted as a strategic transition from an undirected configuration where the entire top part of the spike attracts negatively charged structures, such as the ACE2 binding site, to one where the attraction is directed toward a specific RBD.

The progression of time will uncover a multitude of breakthroughs regarding the behavior of the SARS-CoV-2 S protein. In the short time since the conception of our investigation, numerous discoveries by other researchers have been made public already. For example, the trajectory presented in [43] contains glycans, chemical compounds known to coat the exterior of many viruses, which may have an impact on the results presented here. Recall that the trajectory we explored here is simply one possible path the spike may follow. So there is a possibility that a different trajectory would be a more accurate representation of the true function of the spike protein. Recent studies displayed cryo_EM structures for the spike, presenting an open configuration that differs from the open configuration represented here [40,36]. A similar analysis to the one presented here may be required on different structures, as more probable trajectories are predicted and novel molecular structures uncovered, to capture a more relevant image of the electric potential on and around the spike. Utilizing the same scientific pipeline, we are confident high-fidelity quantitative insights — into the electric potential of molecules unrestricted to the one investigated here — could be obtained efficiently, supporting the global scientific effort.

At the computational level, we have developed a novel framework, combining Markov states models simulations with a continuum physics numerical solver to produce high-resolution multiscale dynamical potential maps of deforming proteins and their surrounding environment. Our framework includes practical mathematical tools to quantify the numerical error and characterize protein electro-geometric properties. The molecular trajectories alone are insufficient for understanding protein electrostatic interactions and, thus, protein core bio-mechanisms. On the other hand, it is unrealistic to envision an approach where trajectories of such a large protein could be obtained solely through continuum modeling. We believe hybrid approaches, such as the one presented here, can provide the scientific community with invaluable information, which we hope can be used to elucidate how changes to physical properties such as electrostatics translate into function.

CRediT authorship contribution statement

M. Theillard and S. Sukenik conceived the presented project. M. Theillard developed and implemented the numerical method. S. Strango, A. Kucherova, and M. Theillard carried the computations and analyzed the Results. All authors contributed to the redaction of the manuscript and approved its final form.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank M. Zimmerman, G. Bowman, and the Folding@Home project for creating and providing us with the S protein opening trajectory. We also thank D. Strubbe and J. Grasis for valuable discussions. This research was supported by a COVID-19 seed grant from the Center for Information Technology Research in the Interest of Society (CITRIS) at UC Merced awarded to M. Theillard and S. Sukenik. The authors acknowledge computing time on the Multi-Environment Computer for Exploration and Discovery (MERCED) cluster at the University of California, Merced, which was funded by National Science Foundation Grant No. ACI-1429783. M. Theillard and S. Sukenik are members of the NSF-CREST Center for Cellular and Biomolecular Machines at the University of California, Merced (NSF-HRD-1547848).

References

[1] Folding@home, http://foldingathome.org.
[2] M. Amin, M.K. Sorour, A. Kasry, Comparing the binding interactions in the receptor binding domains of sars-cov-2 and sars-cov, J. Phys. Chem. Lett. 11 (12) (2020) 4897–4900, PMID: 32478523.
[3] N.A. Baker, Improving implicit solvent simulations: a Poisson-centric view, Curr. Opin. Struct. Biol. 15 (2) (April 2005) 137–143.
[4] N.A. Baker, D. Bashford, D.A. Case, Implicit solvent electrostatics in biomolecular simulation, in: New Algorithms for Macromolecular Simulation, 2006, pp. 263–295.
[5] D. Benton, A. Wrobel, P. Xu, C. Roustan, S. Martin, P. Rosenthal, J. Skehel, S. Gamblin, Receptor binding and priming of the spike protein of SARS-CoV-2 for membrane fusion, Nature (September 2020).
[6] A.H. Boschitsch, M.O. Fenley, A fast and robust Poisson–Boltzmann solver based on adaptive Cartesian grids, J. Chem. Theory Comput. 7 (5) (2011) 1524–1540.
[7] J. Chen, C.L. Brooks, J. Khandogin, Recent advances in implicit solvent-based methods for biomolecular simulations, Curr. Opin. Struct. Biol. 18 (2) (April 2008) 140–148.
[8] I.-L. Chern, J.-G. Liu, W.C. Wang, Accurate evaluation of electrostatics for macromolecules in solution, Methods Appl. Anal. 10 (2003) 309–328.
[9] L. Dagum, R. Menon Openmp, An industry-standard api for shared-memory programming, IEEE Comput. Sci. Eng. 5 (1) (January 1998) 46–55.
[10] S. Decherchi, M. Masetti, I. Vyalov, W. Rocchia, Implicit solvent methods for free energy estimation, Eur. J. Med. Chem. 91 (Feb 2015) 27–42, PMID: 25193298.
[11] R. Egan, F. Gibou, Fast and scalable algorithms for constructing solvent-excluded surfaces of large biomolecules, J. Comput. Phys. 374 (2018) 91–120.

[12] L.E. Felberg, D.H. Brookes, E.-H. Yap, E. Jurrus, N.A. Baker, T. Head-Gordon Pb-am, An open-source, fully analytical linear Poisson-Boltzmann solver, J. Comput. Chem. 38 (15) (June 2017) 1275–1282.
[13] W. Geng, S. Yu, G.W. Wei, Treatment of charge singularities in implicit solvent models, J. Chem. Phys. 127 (11) (2007) 114106.
[14] R.C. Harris, T. Mackoy, M.O. Fenley, Problems of robustness in Poisson-Boltzmann binding free energies, J. Chem. Theory Comput. 11 (2) (Feb 2015) 705–712, PMID: 26528091.
[15] K. Hassanzadeh, H. Perez Pena, J. Dragotto, L. Buccarello, F. Iorio, S. Pieraccini, G. Sancini, M. Feligioni, Considerations around the sars-cov-2 spike protein with particular attention to covid-19 brain infection and neurological symptoms, ACS Chem. Neurosci. 11 (15) (Aug 2020) 2361–2369, PMID: 32627524.
[16] E. Jurrus, D. Engel, K. Star, K. Monson, J. Brandi, L.E. Felberg, D.H. Brookes, L. Wilson, J. Chen, K. Liles, M. Chun, P. Li, D.W. Gohara, T. Dolinsky, R. Konecny, D.R. Koes, J.E. Nielsen, T. Head-Gordon, W. Geng, R. Krasny, G. Wei, M.J. Holst, J.A. McCammon, N.A. Baker, Improvements to the apbs biomolecular solvation software suite, Protein Sci. 27 (1) (2018) 112–128.
[17] P. Koehl, Electrostatics calculations: latest methodological advances, Curr. Opin. Struct. Biol. (2006) 142–151.
[18] J. Lan, J. Ge, J. Yu, S. Shan, H. Zhou, S. Fan, Q. Zhang, X. Shi, Q. Wang, L. Zhang, X. Wang, Structure of the sars-cov-2 spike receptor-binding domain bound to the ace2 receptor, Nature 581 (7807) (May 2020) 215–220.
[19] B. Lee, F.M. Richards, The interpretation of protein structures: estimation of static accessibility, J. Mol. Biol. 55 (3) (1971) 379–400.
[20] M. Letko, A. Marzi, V. Munster, Functional assessment of cell entry and receptor usage for sars-cov-2 and other lineage b betacoronaviruses, Nat. Microbiol. 5 (4) (2020) 562–569.
[21] C. Li, L. Li, M. Petukh, E. Alexov, Progress in developing Poisson-Boltzmann equation solvers, in: Molecular Based Mathematical Biology, Comput. Math. Biophys. 1 (Mar 2013), https://doi.org/10.2478/mlbmb-2013-0002, PMID: 24199185.
[22] B.Z. Lu, Y.C. Zhou, M.J. Holst, J.A. McCammon, Recent progress in numerical methods for the Poisson Boltzmann equation in biophysical applications, Commun. Comput. Phys. 3 (5) (2008) 973–1009.
[23] C. Min, F. Gibou, A second order accurate level set method on non-graded adaptive Cartesian grids, J. Comput. Phys. 225 (2007) 300–321.
[24] C. Min, F. Gibou, H. Ceniceros, A supra-convergent finite difference scheme for the variable coefficient Poisson equation on non-graded grids, J. Comput. Phys. 218 (2006) 123–140.
[25] M. Mirzadeh, F. Gibou, T.M. Squires, Enhanced charging kinetics of porous electrodes: surface conduction as a short-circuit mechanism, Phys. Rev. Lett. 113 (Aug 2014) 097701.
[26] M. Mirzadeh, M. Theillard, F. Gibou, A second-order discretization of the nonlinear Poisson-Boltzmann equation over irregular geometries using non-graded adaptive Cartesian grids, J. Comput. Phys. 230 (2010) 2125–2140.
[27] M. Mirzadeh, M. Theillard, A. Helgadöttir, D. Boy, F. Gibou, An adaptive, finite difference solver for the nonlinear Poisson-Boltzmann equation with applications to biomolecular computations, Commun. Comput. Phys. 13 (2013) 150–173.
[28] D.D. Nguyen, B. Wang, G.-W. Wei, Accurate, robust, and reliable calculations of Poisson-Boltzmann binding energies, J. Comput. Chem. 38 (13) (May 2017) 941–948, PMID: 28211071.
[29] OpenMP Architecture Review Board, OpenMP Application Program Interface, version 5.0 May 2018.
[30] S. Osher, J. Sethian, Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations, J. Comput. Phys. 79 (1988) 12–49.
[31] Q. Pan, X.-C. Tai, Model the solvent-excluded surface of 3D protein molecular structures using geometric pde-based level-set method, Commun. Comput. Phys. 6 (2009) 777–792.
[32] W. Rocchia, E. Alexov, B. Honig, Extending the applicability of the nonlinear Poisson-Boltzmann equation: multiple dielectric constants and multivalent ions, J. Phys. Chem. B 105 (28) (Jul 2001) 6507–6514.
[33] W. Tai, L. He, X. Zhang, J. Pu, D. Voronin, S. Jiang, Y. Zhou, L. Du, Characterization of the receptor-binding domain (rbd) of 2019 novel coronavirus: implication for development of rbd protein as a viral attachment inhibitor and vaccine, Cell. Mol. Immunol. 17 (6) (Jun 2020) 613–620.
[34] M. Theillard, F. Gibou, D. Saintillan, Sharp numerical simulation of incompressible two-phase flows, J. Comput. Phys. 391 (2019) 91–118.
[35] M. Theillard, C. Rycroft, F. Gibou, A multigrid method on non-graded adaptive Octree and Quadtree Cartesian grids, J. Sci. Comput. 55 (2013) 1–15.
[36] A.C. Walls, Y.-J. Park, M.A. Tortorici, A. Wall, A.T. McGuire, D. Veesler, Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein, Cell 181 (2) (April 2020) 281–292.
[37] Y. Wang, M. Liu, J. Gao, Enhanced receptor binding of sars-cov-2 through networks of hydrogen-bonding and hydrophobic interactions, Proc. Natl. Acad. Sci. 117 (25) (2020) 13967–13974.
[38] J. Warwicker, H.C. Watson, Calculation of the electric potential in the active site cleft due to alpha-helix dipoles, J. Mol. Biol. 157 (4) (1982) 671–679.
[39] D. Wrapp, N. Wang, K.S. Corbett, J.A. Goldsmith, C.-L. Hsieh, O. Abiona, B.S. Graham, J.S. McLellan, Cryo-em structure of the 2019-ncov spike in the prefusion conformation, Science 367 (6483) (2020) 1260–1263.
[40] A.G. Wrobel, D.J. Benton, P. Xu, C. Roustan, S.R. Martin, P.B. Rosenthal, J.J. Skehel, S.J. Gamblin, Sars-cov-2 and bat ratg13 spike glycoprotein structures inform on virus evolution and furin-cleavage effects, Nat. Struct. Mol. Biol. 27 (8) (Aug 2020) 763–767.
[41] Q. Zheng, G.-W. Wei. Poisson-boltzmann-nernst-planck model, J. Chem. Phys. 134 (19) (May 2011) 194101, PMID: 21599038.
[42] M.I. Zimmerman, G.R. Bowman, Fast conformational searches by balancing exploration/exploitation trade-offs, J. Chem. Theory Comput. 11 (12) (2015) 5747–5757.
[43] M.I. Zimmerman, J.R. Porter, M.D. Ward, S. Singh, N. Vithani, A. Meller, U.L. Mallimadugula, C.E. Kuhn, J.H. Borowsky, R.P. Wiewiora, M.F.D. Hurley, A.M. Harbison, C.A. Fogarty, J.E. Coffland, E. Fadda, V.A. Voelz, J.D. Chodera, G.R. Bowman, Citizen scientists create an exascale computer to combat covid-19, bioRxiv, 2020.