Computational Electrostatics Predict Variations in SARS-CoV-2 Spike and Human ACE2 Interactions

Scott P. Morton  
Center for Computational Science  
Middle Tennessee State University  
Murfreesboro, TN, USA  
0000-0002-3777-7089

Joshua L. Phillips  
Department of Computer Science  
Middle Tennessee State University  
Murfreesboro, TN, USA  
0000-0002-4619-6083

Abstract—SARS-CoV-2 is a novel virus that crossed over into humans in 2019 and declared a pandemic in early 2020. To understand how the virus infects a new host, we need to understand the mechanistic functions involved with the binding process. To address this need, we generate homology models of SARS-CoV-2 spikes as monomer and trimer to determine the feasibility of reduced computational requirements by using monomer structures. We further generate homology models of the conserved region of SARS-CoV-2 spike subunit s1 noted as the receptor binding domain (RBD) and the human angiotensin-converting enzyme 2 (ACE2). To determine functional breadth of spike monomer, trimer and RBD in relation with ACE2, we apply Coulomb's Law to determine an electric force between combinations with ACE2 across the range of pH from 3.0 to 9.0 in 0.1 increments. The results indicate that spike trimer should be used to determine mechanistic binding function and these data indicate that variations of spike sequence influence breadth of function. Our results also indicate the RBD has a broader range of function across pH compared to spike trimer, but is influenced by the range of function presented by the spike trimer.

Index Terms—COVID-19, SARS-CoV-2, electrostatics, binding, pH

I. INTRODUCTION

SARS-CoV-2 (CoV-2) is the latest novel coronavirus to have crossed the species boundary into humans over the past twenty years [1], [2]. Currently available information suggests an originating source to be the bat coronavirus SL-CoV-RaTG13 with 96% similarity to CoV-2 [3].

The World Health Organization (WHO) declared CoV-2 a public health emergency on January 30, 2020 and then pandemic status in the weeks following. WHO utilizes a tool to distinguish diseases based on the potential to bloom into epidemic/pandemic proportions and the availability of countermeasures to focus research and development (R&D) resources accordingly [4], [5].

We have previously performed extensive research using electrostatic analysis of HIV gp120 envelope glycoprotein interactions with human CD4 and broadly neutralizing antibodies [6]–[12]. Electrostatics are a novel way to evaluate protein interactions at the structural level with a focus on how environmental conditions, like pH, affect binding efficacy. Furthermore, our methods can predict how pH affects electric force, a measure of attractive or repulsive forces between two charged objects, known as Coulomb's Law.

The process can easily be adapted to any other protein-protein interaction which may be strengthened or weakened by modulating pH. Researchers can consume these data to evaluate what drives two proteins to interact and how solvent pH affects those interactions before performing laboratory experimentation.

To illustrate our methods, we generate homology models of spike as a monomer and trimer to compare the two structures analytically to determine if significant reductions in computational requirements can be achieved through similar results. We also evaluate the two structure types in relation with ACE2 to map forces that drive the two proteins to bind together. Additionally, we analyze data of the conserved s1 subunit known as the receptor binding domain (RBD) to produce a picture of the binding forces involved.

II. MATERIALS AND METHODS

CoV-2 spike sequences were sourced from NCBI GenBank accession numbers: 6VXX [13], 6XCM [14], 6XS6 [15], 6ZGF [16], 7BYR [17], QIS30165 [18], and QKM75579 (directly deposited to GenBank). The RBD is a highly conserved region of the CoV-2 spike subunit s1 extracted from accession number 6LZG. ACE2 is the highly conserved angiotensin-converting enzyme 2 (ACE2) extracted from NCBI GenBank accession number BAJ12180.1.

We curate spike sequences to isolate the trimer bulb from the attaching umbilical that allows undulation. We conclude that the distances involved present these forces at orders of magnitude less than those of the trimer bulb, however, our process addresses these forces as a single point because of computational requirements needed to evaluate accumulating forces over a range of distance.

The electrostatic surface charge (ESSC) pipeline calculates the average static-charge across the solvent accessible surface (SAS) of the molecule being analyzed by determining all grid points containing SAS atoms using APBS [19], [20] data generated for the entire molecule. Complete details of the methods related to the electrostatic pipeline can be found in [6]–[11].
Template PDB used for spike trimer is NCBI GenBank accession number 7A98 and for spike monomer 7A98_A. To target Frodan for specific conformations of spike trimer we used accession numbers: 6VXX and 7A98 for closed and opened conformations respectively. For spike monomer we use 6ZB5_B and 6ZGG_B for closed and opened conformations respectively.

Template PDB’s used to model RBD of subunit s1 are: 6LZG, 6M0J, 6VW1, and 6VYB [13], [21]–[23] and for ACE2 the templates are: 1R4L, 1R42, 6M0J, 6VW1 [22]–[24]. To target Frodan for specific conformations of s1 we used: 6LGZ as open and 6VYB as closed conformations. For ACE2 we used 6LZG for open and closed target conformations. Furthermore, our results indicate that no conformational changes take place for s1 or ACE2 during binding and are therefore indistinguishable in simulation.

For the specific calculation of Electric Force, APBS returns a factor for $V$ from Equation 1:

$$V = \frac{k_B T}{e_c}$$

Where $V$ is voltage of the system per unit, $k_B$ is the Boltzmann’s constant: 1.3806504 E-23 JK⁻¹, $T$ is the temperature in Kelvin and $e_c$ is the charge of an electron: 1.60217646 E-19 C.

For the calculation of Electric Force ($F_e$) we have Coulomb's Law presented as Equation 2:

$$F_e = k_c \frac{|q_1 q_2|}{r^2}$$

Where $F_e$ is the Electric Force in Newtons, $k_c$ is Coulomb's constant: 8.987551E+09, $q_1$ is the charge in Coulombs of the first mass, $q_2$ is the charge in Coulombs of the second mass, and $r$ is the distance in meters between the two masses.

To derive Coulombs of charge from APBS [19], [20] requires a simple transposition of the terms $V$ and $e_c$ and multiplication of the results with values returned by APBS. For this model, the variables required to complete the calculations are: $T = 310 K$ and $r = 10 \text{Å}$ for all electric force calculations noting that a change in distance only affects force amplitude.

III. RESULTS

A. Comparing Spike Monomer to Spike Trimer

We evaluate our processes for spike monomer in contrast with spike trimer to determine the value of reduced computational requirements by allowing the use of spike monomer exclusively.

We begin by examining box plots of electrostatic data in Fig. 1 for a spike trimer (top), spike monomer (middle), and from ACE2 (bottom); plots are comprised of 10 models each. We can observe that charge values for spike trimer (top) transition to negative between pH 6.0 and 6.5, for spike monomer (middle) between pH 5.5 and 6.0 and ACE2 (bottom) transitions to negative between pH 4.5 and 5.0. This presents an opportunity to evaluate potential differences and from there, electric force.

By taking the absolute values of the difference between combinations of ACE2 with spike trimer and spike monomer, surface charges expose a predicted electric potential difference for the two complexes. A typical representation is displayed in Fig. 2 showing the trimer (top) and the monomer (bottom) vary greatly. Color bars indicate the pH range of the following: normal nasal secretions of pH 5.5 to 6.5 is depicted in red; yellow indicates inflamed nasal pH and blue indicates lung pH.
list 3.5 to 4.5 (0.1 increments), however, the graphs scale in 0.1 increments and displays the range as pH 7.3 to 7.5 as inclusive of the specific range.

The presence of electric potential difference introduces a force that can be determined by applying Coulomb's Law to obtain the Electric Force ($F_e$) in Newtons ($N$) that act upon the two structures. We can graph the changing forces, either repulsive (+), or attractive (-), between two structures across the pH range of 3.0 to 9.0 in 0.1 increments, and at a distance of 10 Å. Our model is non-specific in that, specific interactions involve close-proximity contacts between proteins such as: hydrogen bonds, ionic bonds, and hydrophobic residue stacking. However, non-specific interactions are not close-proximity, and can be attractive or repulsive beyond the range of 3 angstroms or more. Electric force is a good example of the latter case.

Fig. 3 displays a typical electric force across the evaluated pH range where any positive value represents a repulsive force and any negative value indicates an attractive force. The interesting portion of the graph is below zero where an attractive force is the desirable result that invokes binding.

In Fig. 4 we compare the interesting portion of electric force, trimer versus monomer, that exposes differences in force amplitude and pH range.

Additionally, we evaluate spike trimer using Electrostatic Variance Masking (EVM) to map residues that display significant changes to charge as solvent pH varies. Fig. 5 displays variance maps for trimer (top) and monomer (bottom) in open conformations. The amplitude of variance is significantly reduced for spike monomer results and this would cause EVM to select a wider range of residues. Closed conformations show similar differences as open (data not shown).

To better understand these differences, Fig. 6 expresses a typical trimer in tricolor format to contrast the three spikes and show entanglement in their natural state. The woven structure reduces overall solvent accessible surface area of each spike and forces the trimer to function as a unit.

With s1 in the closed state, the RBD overlaps the neighboring spike at the non terminating domain (NTD) as show in Fig. 7. The forces acting upon s1 by the interaction with NTD in the closed conformation would need to be calculated by extracting the NTD from spike and performing the same process performed for all other structures.

We assert that electrostatic surface charge data, potential difference data, electric force data and EVM data suggest that spike monomer cannot be used to evaluate the forces involved in the mechanistic function of protein binding.
Fig. 5. Variance maps of a spike trimer (top) and spike monomer (bottom) for the open conformation. Reduced amplitude returned for spike monomer (bottom) generates a larger list of pH sensitive residues showing that the results are significantly different compared to that of spike trimer (top).

B. Spike Trimer

1) Electric Force: As previously shown, electric force can be calculated to determine an attractive or repulsive force between two charged objects. We continue this analysis using graphs that remove repulsive forces, to the extent possible without manipulation of data. Fig. 8 presents electric force determined for a trimer in relation with ACE2. Ignoring amplitude of the force, we can suggest that the breadth of function for this combination of spike trimer and ACE2 begins where results cross 0.0 to a negative value and ends where values cross 0.0 back to positive.

Combining results for all seven combinations of spike trimer and ACE2 presents an interesting view displaying breadth of function and differences of each combination. Fig. 9 represents spike trimer in the closed conformation while Fig. 10 displays spike trimer in the open conformation.

Looking at the extremes of Fig. 9, we select sequences 6VXX and QKM75579 and note that QKM75579 has a greater breadth of function over 6VXX. This could be construed as QKM75579 has a gain of function over 6VXX exceeding two full points of pH. In Fig. 10, we can see the same unique characteristics as in Fig. 9, and again QKM75579 has a gain of function in excess of two full points of pH over 6VXX.

Further observations of these figures suggest that sequence QIS30165 has minimal change in breadth of function transforming from closed to open conformation, but does have a variation in amplitude during conformational change. We further note that breadth of function for QKM75579 increases beyond the range of pH associated with human lung tissue and includes a broader range of higher amplitude as do all variants.

2) Electrostatic Variance Masking: EVM provides the ability to view pH sensitive residues as an overlay on spike trimer using VMD [27] and rendering with the internal Tachyon driver [28] to produce the imagery.

Fig. 11 displays a side view of spike trimer with EVM...
selected residues highlighted with a transparent red surface. The image expresses that most of the residues that respond to varying pH are contained in the core of the trimer bulb and residues associated with fusion peptides. The NTD also indicates some changes to specific residues in the open conformation. Fig. 12 displays the same spike trimer in the closed conformation with reduced activity at the NTD suggesting that s1 would have to move towards the open state to activate NTD residues observed in Fig. 11.

3) Receptor Binding Domain: This analysis uses the highly conserved region of spike subunit s1 described as the receptor binding domain. This region is subject to change or may have already changed at some point in time, however, this does not diminish these observations in anyway.

We begin with Fig. 13 to show the motif of electric potential across the pH range of 3.0 to 9.0 in 0.1 increments for the RBD (top) and ACE2 (bottom). We observe that charge values for the RBD never cross zero to negative. We also observe the charge values for ACE2 across the same pH range and note that charge values cross to negative between pH 4.5 and 5.0.

We further note the charge motifs present a potential difference between the two structures and display these data in

IV. CONCLUSIONS

Our results indicate that spike monomer cannot be used to predict the mechanics involved in the binding process between CoV-2 and ACE2. We assert that entanglement of spike as a trimer shields many residues from charge modifications due to pH and this requires the consideration of spike trimer functioning as a unit.

Fig. 8. Electric force of a typical CoV-2 spike trimer in the open conformation focusing on the range of attractive force. Red color bar indicates normal nasal pH, yellow represents inflamed nasal pH and blue indicates lung pH.

Fig. 9. Electric force of spike trimer variants with s1 in the closed conformation displaying unique characteristics of each. Red color bar indicates normal nasal pH, yellow represents inflamed nasal pH and blue indicates lung pH.

Fig. 10. Electric force of spike trimer variants with s1 in the open conformation displaying unique characteristics of each. Red color bar indicates normal nasal pH, yellow represents inflamed nasal pH and blue indicates lung pH.

Fig. 11. Side view of typical spike trimer in the open conformation detailing pH sensitive residues shown in transparent red as a surface rendering.

Fig. 14. This motif establishes the foundation for an electric force to be calculated and presented in Fig. 15. Again, this is the full range of evaluated pH and the interesting portion is the attractive forces beginning at the transition from positive to negative values.

In Fig. 16, we highlight this attraction to show greater detail. The motif clearly shows that the RBD and ACE2 would be attracted to each other in any close proximity scenario starting between pH 4.5 and 5.0. This should be an expected circumstance as any repulsive forces would disrupt binding efficacy.
Fig. 12. Side view of typical spike trimer in the closed conformation detailing pH sensitive residues shown in transparent red as a surface rendering.

Fig. 13. Charge data as a box plot for the receptor binding domain of CoV-2 spike (top) showing the charge values remain positive across the pH range. Charge data as a box plot for the angiotensin-converting enzyme 2 (bottom) showing the charges cross into negative values between pH 4.5 and 5.0; each plot presents data from ten models. Red color bar indicates normal nasal pH, yellow represents inflamed nasal pH and blue indicates lung pH.

We evaluate these structures to determine electric forces involved across the relevant range of pH 3.0 to 9.0 in 0.1 increments. We note that electric force motifs of CoV-2 spike variants have differing breadth of function and amplitude of force and this drives binding efficacy. We further assert that variants can be considered to have a gain of function from one variant to the next that can be observed by calculating electric force as described in this research. These data can assist in the understanding of binding mechanisms and how these proteins are tuned for specific environments.

Considering the many common knowledge functions of nasal passages, notably as a filter for particulate prior to entering the lungs, spike trimer is highly tuned for pH 5.5 to 6.5 (normal nasal pH range) making nasal passages the primary target for initial infection.

We also evaluated interactions of the RBD and ACE2 across the relevant pH range to express electric forces involved. These data indicate a wide range of functionality for the RBD to interact with ACE2 under varying environmental pH levels. Data for the RBD in concert with spike trimer bulb indicates the potential for higher rates of lung infection for variants with a broader pH range of function as shown in Fig. 9 and Fig. 10.

Additionally, we assert the need to evaluate all variants of CoV-2 spike and any variations in the RBD to understand the gain of function the virus establishes over time, regardless of naturally occurring mutations or clinical engineering.

We also suggest that further research is required to determine other functional aspects of spike, namely the afore-
mentioned s1, NTD interactions and how pH may influence opening and closing of s1 to protect the RBD while seeking a target ACE2 protein to interact with.

We provide access to the full set of figures generated for this study as a tar package located at: https://github.com/spmorton/SARS-CoV-2-

**REFERENCES**


