Research article

Electrostatic features for nucleocapsid proteins of SARS-CoV and SARS-CoV-2

Wenhan Guo[1], Yixin Xie[1], Alan E Lopez-Hernandez[1], Shengjie Sun[1] and Lin Li[1,2,*]

[1] Computational Science Program, University of Texas at El Paso, El Paso, TX 79968, USA
[2] Department of Physics, University of Texas at El Paso, El Paso, TX 79968, USA

* **Correspondence:** Email: lli5@utep.edu.

Abstract: COVID-19 is increasingly affecting human health and global economy. Understanding the fundamental mechanisms of Severe Acute Respiratory Syndrome CoronaVirus 2 (SARS-CoV-2) is highly demanded to develop treatments for COVID-19. SARS-CoV and SARS-CoV-2 share 92.06% identity in their N protein RBDs' sequences, which results in very similar structures. However, the SARS-CoV-2 is more easily to spread. Utilizing multi-scale computational approaches, this work studied the fundamental mechanisms of the nucleocapsid (N) proteins of SARS-CoV and SARS-CoV-2, including their stabilities and binding strengths with RNAs at different pH values. Electrostatic potential on the surfaces of N proteins show that both the N proteins of SARS-CoV and SARS-CoV-2 have dominantly positive potential to attract RNAs. The binding forces between SARS-CoV N protein and RNAs at different distances are similar to that of SARS-CoV-2, both in directions and magnitudes. The electric filed lines between N proteins and RNAs are also similar for both SARS-CoV and SARS-CoV-2. The folding energy and binding energy dependence on pH revealed that the best environment for N proteins to perform their functions with RNAs is the weak acidic environment.

Keywords: SARS-CoV; SARS-CoV-2; COVID-19; protein-protein interactions; protein-RNA/DNA interactions; electrostatic force; DelPhi; DelPhiForce

1. Introduction

Severe Acute Respiratory Syndrome CoronaVirus 2 (SARS-CoV-2) is currently affecting human

health and global economy seriously. A similar situation happened in 2003 with SARS-CoV, which also belongs to Coronavirus family. SARS-CoV and SARS-CoV-2 share 92.06% identity in their N protein RBDs' sequences [1,2]. Both SARS-CoV and SARS-CoV-2's genomes encode nonstructural replicase polyproteins and structural proteins [1], including the nucleocapsid phosphoprotein (N protein). The main function of N protein is to link envelopes to the +RNA. The N protein of SARS had shown to play a crucial role in regulating viral RNA synthesis in replication and transcription [3]. Understanding the fundamental mechanisms of how N proteins Receptor Binding Domains (RBDs) of SARS-CoV and SARS-CoV-2 bind RNAs is highly demanded for developing new antiviral drugs and vaccines [4].

Some groups studied the N proteins of SARS-CoV and SARS-CoV-2 using experimental methods. Sisi Kang et al. [3] utilized chemical experiments and X-ray analysis to obtain the structure of N protein of SARS-CoV-2, which helped revealing potential drug targeting sites. Veverka V's laboratory [5] performed NMR-based titration experiments, combined with computational model, to build the complex model of the Nucleocapsid N-Terminal RNA binding Domains (N-NTD) with RNA. Unfortunately, only a few groups have conducted research on the structure and function of N proteins of SARS-CoV. Peter Kuhn's team [6] is one of them who characterized the structures of the N-NTD of SARS-CoV. Compared with experimental studies, some effort has been also made to investigate SARS-CoV and SARS-CoV-2 using computational approaches. Most of these computational studies focused on the spike (S) proteins of the SARS-CoV and SARS-CoV-2 [7,8], including discoveries of potential drug targets for SARS-CoV-2 [9–11], few works focused on N proteins. Some studies calculated the electrostatic potential on N protein surfaces in coronavirus [3,5,6]. Electrostatic features of N proteins help us understanding different mechanisms of RNA recognition and assembly. Other calculations between N proteins and RNAs explore more fundamental principles for their binding mechanisms.

Due to the relatively high cost of experiments and the rapid development of computational algorithms [12,13], computational methods are now widely used to study biology phenomena, including biomolecular structures [14,15], biomolecular interactions [16–18], pH dependence of protein-protein/DNA/RNA interactions [19,20], etc. Using such state of art computing techniques, a lot of efforts have been contributed to study viruses [7,21,22]. In this work, several computational approaches are used to study nucleocapsid proteins of SARS-CoV and SARS-CoV-2, including DelPhi [23], DelPhiForce [24,25], DelPhiPKa [26,27]. The electrostatic features are critical in analyzing the interactions between the N protein and RNA. Thus, the electrostatic potential, electric field lines and electrostatic forces were analyzed based on the structures of N proteins of SARS-CoV/SARS-CoV-2 RBDs and RNAs. It was found that SARS-CoV and SARS-CoV-2 have similar electrostatic potential distributions on their binding surfaces, which demonstrated that the net charges play a significant role to attract the RNAs. In addition, DelPhiPKa was implemented to calculate the binding energy pH dependence. Such method has been proved successful and reliable [20,28–30]. The pH effects on the binding energies for N proteins' RBDs interacting with RNAs and folding energies of N proteins was analyzed, which demonstrated the optimal pH for N proteins' folding and binding with RNAs. Such details assist us to understand how the N proteins' RBDs recognize RNAs. These findings pave the way for research on future coronavirus-caused diseases. No experimental studies have been conducted to reveal the differences between the biology functions of SARS-CoV and SARS-CoV-2. Therefore, this work of comparing the N proteins of SARS-CoV and SARS-CoV-2 can also be useful for future experimental design.

2. Methods

2.1. Structure preparation

The complex structure of SARS-CoV-2 with the Double Strand RNA (dsRNA) was obtained from Protein Data Bank (pdb ID: 7ACS [5]). The SARS-CoV structure was obtained from Protein Data Bank (pdb ID: 2OFZ [6]), which does not include the dsRNA structure. Therefore, the complex structure of dsRNA combined with SARS-CoV N protein was modeled by aligning the SARS-CoV structure to SARS-CoV-2 based on the template of 7ACS using Chimera [31]. This study is mainly focused on the electrostatic features of Nucleocapsid N-Terminal RNA binding Domains (N-NTDs) of SARS-CoV and SARS-CoV-2. In the SARS-CoV N protein structure, the N and C terminals are not determined [6]. Figure S1 shows the complex structures of SARS-CoV-2 N protein RBD binding with RNA, which is determined by NMR experiments [5]. The NMR structures demonstrate none of the N or C terminals of SARS-CoV-2 binds to RNAs, therefore the N and C terminals are extremely flexible. Due to this experimental evidence, N and C terminals of SARS-CoV-2 were deleted in this work. After the deletion, we obtained the same length of N proteins for SARS-CoV-2 and SARS-CoV.

2.2. Electrostatic Calculations using DelPhi and DelPhiForce

DelPhi [23] and DelPhiForce [24,25] tools focus on accurate calculations and visualizations of the electrostatic potential and forces for biomolecules. They were used to calculate the electrostatic potential and total force for the N protein RBD and RNA binding domain. Finite difference (FD) method is implemented in the DelPhi and DelPhiForce tools to solve the Poisson-Boltzmann equation (PBE):

$$\nabla \cdot [\epsilon(r)\nabla\phi(r)] = -4\pi\rho(r) + \epsilon(r)\kappa^2(r)\sinh(\phi(r)/k_B T), \tag{1}$$

Where $\phi(r)$ is the electrostatic potential, $\epsilon(r)$ is the dielectric permittivity, $\rho(r)$ is the permanent charge density according to the atomic structure, κ is the Debye–Huckel parameter, k_B is the Boltzmann constant, and T is temperature.

The electrostatic potential of the SARS-CoV N protein and SARS-CoV-2 N protein with RNA domain was calculated by DelPhi. Their surfaces were visualized by Chimera [31] using the color scale range from −1.0 to 1.0 kT/e (see Figure 1). In order to compare the directions and strengths of electrostatic forces, the N protein and RNA was separated from 5 Å to 40 Å with the step size of 5 Å using StructureMan [32]. Then at each position, the electrostatic force was calculated by DelPhiForce. The visual molecular dynamics (VMD) [33] was implemented to visualize the total forces and the electric field lines between N protein and RNA.

2.3. Folding energy calculation methods

DelPhiPKa [26,27] was implemented to calculate pKa values of nucleotides of RNAs' and proteins' ionizable residues. The net charge of the unfolded state was calculated with the following equation:

$$Q_u(pH) = \sum_{i=1}^{N} \frac{10^{-2.3y(i)(pH-pKa(i))}}{1+10^{-2.3y(i)(pH-pKa(i))}}, \tag{2}$$

where the summation is all titratable groups, y(i) is −1 for acidic groups and +1 for basic groups.

The pKa values of SARS-CoV N protein and SARS-CoV-2 N protein with RNA domain were calculated by DelPhiPKa. The pKa range was set from 0 to 14 with an interval of 0.5 in the calculations. The pH-dependence of the folding free energy using the equation:

$$\Delta N(pH_{folding}) = 2.3\text{RT} \int_{pH_i}^{pH_f} (Q_f(pH) - Q_u(pH)d(pH)), \tag{3}$$

Where $Q_f(pH)$ and $Q_u(pH)$ are the total net charge of folded and unfolded states. R is the universal gas constant taken as $1.9872 \times 10^{-3} \frac{kcal}{Mol*K}$. T is the temperature, which is 300 K.

2.4. Binding energy calculation methods

The pH-dependence of the binding energy of N proteins with RNAs was modeled by obtaining the pH-dependence of the net charge of the complexes and their components. The pH dependence of the stability of the complexes and their components using the equation:

$$\Delta N(pH_{binding}) = 2.3RT \int_{pH_i}^{pH_f} (Q_t(pH) - Q_n(pH) - Q_r(pH))d(pH) \tag{4}$$

where $\Delta N(pH_{binding})$ is the pH-dependence of the binding free energy, $Q_t(pH)$, $Q_n(pH)$, and $Q_r(pH)$ are the net charges of complex, N protein and RNA, R is the universal gas constant taken as $1.9872 \times 10^{-3} \frac{kcal}{Mol*K}$. T is the temperature, which is 300 K.

3. Results and discussion

There are no experimental studies which investigated the differences between the biology functions of SARS-CoV and SARS-CoV-2. Therefore, this work which compared the N proteins of SARS-CoV and SARS-CoV-2 can be used for future experimental design. The mutations between SARS-CoV and SARS-CoV-2 N proteins were analyzed to show the distribution of mutations on the SARS-CoV-2 N protein. Furthermore, the electrostatic features of N proteins of SARS-CoV and SARS-CoV-2 were investigated. Finally, the binding and folding energies of the complexes and their components were calculated and analyzed.

3.1. Mutations between SARS-CoV and SARS-CoV-2

The structures of N protein RBDs of SARS-CoV and SARS-CoV-2 are very similar (the RMSD is 0.967Å). We aligned the sequences of N protein RBDs of SARS-CoV and SARS-CoV-2 using clustal omega [34] to analyze their sequence differences. The positions of the mutation sites are marked in Figure 1 with orange color. Most of the mutation sites in the N protein RBD are distributed on or closed to the hairpin-like structure. It suggests that the flexibilities of the hairpin-like structure N protein RBDs may be different between these two viruses. The flexibility of the SARS-CoV-2 N protein's hairpin-like loop structure is shown in Figure S1.

Figure 1. Mutations on SARS-CoV-2 compared with SARS-CoV. A. The structure of SARS-CoV-2 N proteins. The mutation sites are highlighted in orange color. B. Sequence alignment of N proteins of SARS-CoV and SARS-CoV-2.

3.2. Electrostatic potential on surfaces

The electrostatic features are important for protein structure and functions. We calculated the electrostatic potential of the N proteins of SARS-CoV and SARS-CoV-2. With the analysis, the binding interfaces of N proteins showed dominantly positive electrostatic potential (see Figure 2) while the RNAs are negatively charged. Thus, the N protein RBDs are attracted by RNAs because the two interfaces have opposite net charges. Such a phenomenon is common in the interactions between biomolecules [7,35]. The electrostatic features of N proteins and RNAs indicate that the electrostatic binding forces between N proteins and RNAs may enhance the stabilities of the complexes.

Figure 2. Structure and electrostatic surfaces of N proteins of SARS-CoV/SARS-CoV-2. (A) The side view of the structure of SARS-CoV (purple) and SARS-CoV-2 (cyan) N proteins. (B) The side view of the electrostatic potential on the surface of SARS-CoV; (C) The side view of the electrostatic potential on the surface of SARS-CoV-2; (D) The top view of the structure of SARS-CoV (purple) and SARS-CoV-2 (cyan) N proteins; (E) The top view of the electrostatic potential on the surface of SARS-CoV; (F) The top view of the electrostatic potential on the surface of SARS-CoV-2.

3.3. Electric field lines

The N protein structures of SARS-CoV and SARS-CoV-2 binding with RNAs are shown in Figure 3. From the complex structures, it is obvious that the RNAs bind to the hairpin-like loop of the N proteins. The structures of SARS-CoV and SARS-CoV-2 are very similar except the hairpin-like loops. Because the hairpin-loops are much more flexible than the rest of the N protein structures. Note that the N and C terminals have been removed from both of the N proteins to obtain more stable structures, because these terminals are too flexible and have no contribution to the binding interactions. The details are shown in the methods section.
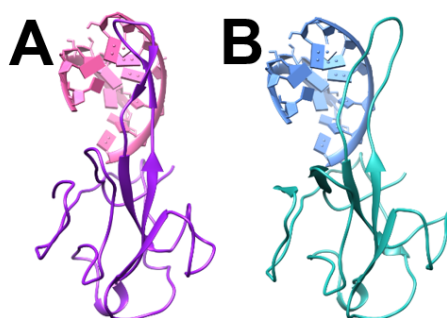
Figure 3. Complex structures of N protein RBDs and RNAs. (A) SARS-CoV N protein RBD (purple) bind with RNA (pink); (B) SARS-CoV-2 N protein RBD (cyan) bind with RNA (blue).
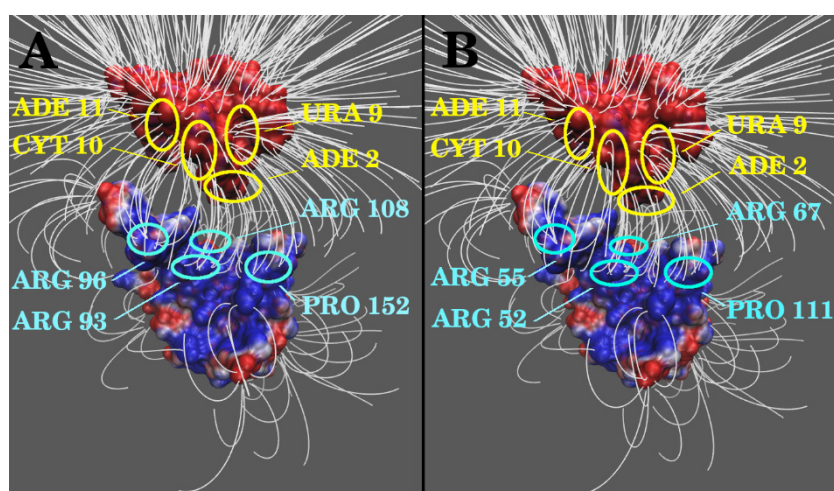
Figure 4. Electric field lines of complex binding domains. (A) Electric field lines of SARS-CoV N protein RBD and RNA; (B) Electric field lines of SARS-CoV-2 N protein RBD and RNA. The residues with high density of filed lines are circled and highlighted, including those on RNAs (yellow) and on N proteins (blue).

To further explore the electrostatic interactions, we calculated the electric field lines between N protein RBDs and RNAs (see Figure 4). Densities of field lines represent the strengths of

electrostatic interactions. From the electric field line distributions, it is clearly shown that both N proteins of SARS-CoV and SARS-CoV-2 have strong attractive binding forces to RNAs. The residues with dense field lines on the RNA interface areas are the same, which are ADE2, URA9, CYT10 and ADE11. On the other side, the key residues generating dense field lines on N proteins of SARS-CoV and SARS-CoV-2 are also the same (note that the sequence numbers of SARS-CoV and SARS-CoV-2 have 41 residues difference). For SARS-CoV, the key residues are: ARG93, ARG96, ARG108, PRO152; for SARS-CoV-2, the corresponding identical key residues are: ARG52, ARG55, ARG67, PRO111. In each case, three out of four N proteins' dense field lines generating residues are arginine. Also, those key residues that produce the dense electric field lines do not have any mutation from SARS-CoV to SARS-CoV-2, which means that these residues are conserved.
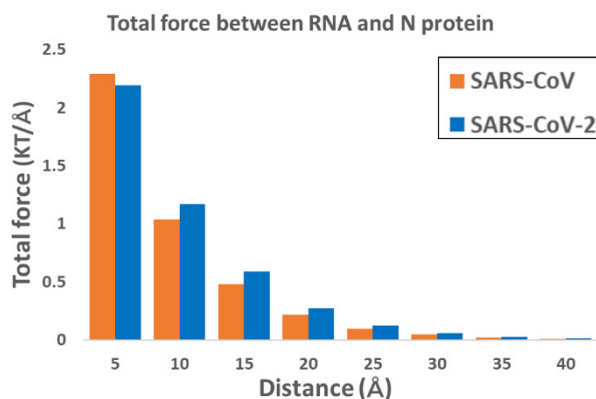
3.4. Electrostatic forces

Electrostatic forces of SARS-CoV and SARS-CoV-2 N proteins' RBDs with RNAs at distances from 5 Å to 40 Å with a step size of 5 Å were separated by StructureMan [32] and calculated by DelPhiForce at each position (see Figure 5). The directions of the blue arrows are illustrated to show the directions of net forces between N proteins and RNAs. The arrows are normalized to the same size for better visualizations. From the figures, the electrostatic forces of N proteins attract the corresponding RNAs. It clearly showed that the directions of arrows are different by comparing SARS-CoV and SARS-CoV-2 N protein RBDs with RNAs at distance of 5 Å. It may because of the hairpin-like loop structure at the top of the N protein, which is more flexible as shown in Figure S1.
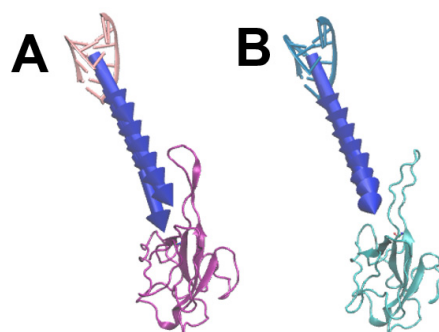
Figure 5. Electrostatic forces of SARS-CoV and SARS-CoV-2 N protein RBDs with RNAs at distances from 5 Å to 40 Å. (A) Electrostatic force of SARS-CoV N protein RBD (purple) and RNA(pink) at distances from 5 Å to 40 Å with a step size of 5 Å; (B) Electrostatic force of SARS-CoV-2 N protein RBD (cyan) and RNA (light blue) at distances from 5 Å to 40 Å with a step size of 5 Å.

While Figure 5 only focuses on the directions of the electrostatic forces, the magnitudes of these electrostatic forces of SARS-CoV and SARS-CoV-2 are shown in Figure 6. The electrostatic forces between N proteins' RBDs and RNAs decrease as the distances increase. It is obviously shown that SARS-CoV and SARS-CoV-2 have similar electrostatic forces at different distances. SARS-CoV-2 has relatively stronger forces than SARS-CoV, except at the distance of 5 Å. Figure 6 only compares

the magnitudes of the forces between SARS-CoV and SARS-CoV-2. However, the directions of the forces are also important for electrostatic forces, which is shown in Figure 5.


Figure 6. The trends of electrostatic forces of SARS-CoV and SARS-CoV-2 N proteins' RBDs with RNAs at distances from 5 Å to 40 Å. Orange and Blue colors represent SARS-CoV and SARS-CoV-2, respectively.

3.5. Binding energies

In protein-DNA/RNA complexes, it is common that the binding energies depend on the pH environment [29,30]. The pH-optimum is the pH at which the complex has maximal electrostatic binding energy [28]. To demonstrate the pH dependence in the binding process of N proteins and RNAs, DelPhiPKa was implemented to calculate the binding energies. It should be mentioned that the binding energies calculated using DelPhiPKa method are relative binding energies rather than absolute energies. By default, the binding energy at pH 0 is set as reference, which is 0 kcal/mol. The relative energy profile can be used to study the binding energy dependence on pHs.

The results are shown in Figure 7. From the binding energy curves, it is obvious that for both SARS-CoV and SARS-CoV-2, the binding energy is stable within a wide range of pH (from 5.5 to 10). Such pH independent binding energy phenomena were also found in some other related studies [19].
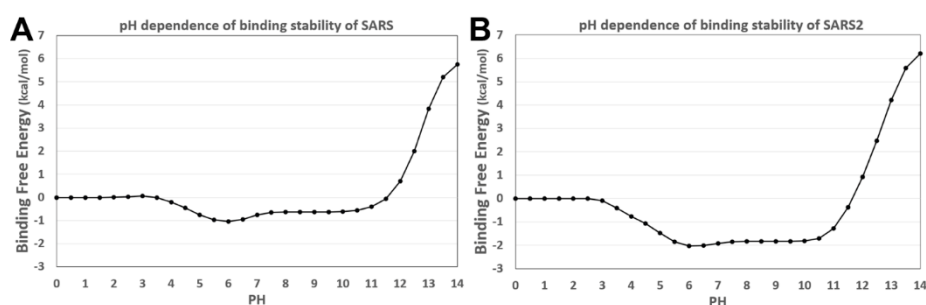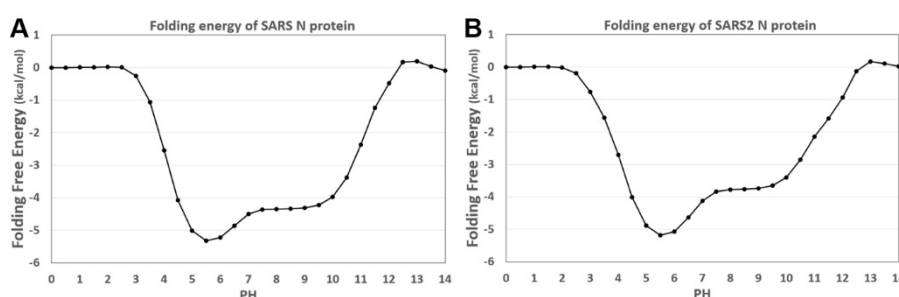

Figure 7. The calculated pH dependence of the binding energy between N proteins' RBDs and RNAs. (A) The binding energy of SARS-CoV N protein RBD and RNA; (B) The binding energy of SARS-CoV-2 N protein RBD and RNA. The pH range was set from 0.0 to 14.0 with an interval size of 0.5.

3.6. Folding energies

The net charges of SARS-CoV and SARS-CoV-2 are calculated with DelPhiPKa [26,27]. The pH range was set from 0 to 14 with an interval of 0.5. Figure 8 shows the calculated folding energies of SARS-CoV and SARS-CoV-2 at different pH values. The pH-dependence of the folding free energy demonstra that SARS-CoV and SARS-CoV-2 have the same pH-optimum value where the folding energy is the most favorable at this pH (here the pH-optimum value is 5.5). In addition, N proteins of SARS-CoV and SARS-CoV-2 have similar inverted funnel-shaped folding energy curves. These curves indicate that the pH-dependences of folding energies of SARS-CoV/SARS-CoV-2 N proteins binding with RNAs are very similar. And the combination of the folding energy and binding energy profiles demonstrates that the N proteins perform their functions best at pH 5.5.

Figure 8. The calculated pH-dependence of the folding energy of N proteins' RBDs and RNAs. (A) The folding energy of SARS-CoV N protein RBD and RNA; (B) The folding energy of SARS-CoV-2 N protein RBD and RNA. pH ranges from 0.0 to 14.0 with an interval of 0.5.

4. Conclusions

Due to the sequence similarity, SARS-CoV and SARS-CoV-2 have very similar functions and structures. Each of their genes encodes four types of structural proteins, including N protein which is studied in this work. The N proteins of SARS-CoV and SARS-CoV-2 are similar in sequence and almost identical in structure. This study revealed some fundamental mechanisms of these N proteins, including their stabilities and binding strengths with RNAs at different pHs.

Multiple computational approaches were utilized in this work to investigate the N proteins. Electrostatic potential of the surfaces of N proteins show that both of the N proteins from SARS-CoV and SARS-CoV-2 have similar electrostatic potential distributions. The binding interfaces are dominantly positively charged, which results in attractive electrostatic interactions to RNAs. The electrostatic force analyses validated such attractive interactions. The binding forces between SARS-CoV N protein and RNA at different distances are similar to that of SARS-CoV-2, in both directions and magnitudes. Electric filed lines between N proteins and RNAs are also similar between SARS-CoV and SARS-CoV-2. The binding energy dependence to pHs shows that the binding of both N proteins with RNAs are stable in a wide range of pH (from pH 5.5 to 10). For folding energy dependence to pH, the optimal pH is found as 5.5 for both N proteins. This indicates that the N proteins perform their functions best in a weak acidic environment, which is perfect for

theses N proteins to maintain their structures and perform functions surrounding RNAs.

Acknowledgments

Conflict of interest

Authors declare no conflicts of interest in this paper.

References

1. P. Zhou, X. L. Yang, X. G. Wang, B. Hu, L. Zhang, W. Zhang, et al., A pneumonia outbreak associated with a new coronavirus of probable bat origin, *Nature*, **579** (2020), 270–273.
2. M. A. Marra, S. J. Jones, C. R. Astell, R. A. Holt, A. Brooks-Wilson, Y. S. Butterfield, et al., The genome sequence of the SARS-associated coronavirus, *Science*, **300** (5624), 1399–404
3. S. Kang, M. Yang, Z. Hong, L. Zhang, S. Chen, Crystal structure of SARS-CoV-2 nucleocapsid protein RNA binding domain reveals potential unique drug targeting sites, *Acta Pharm. Sin. B*, **10** (2020), 1228–1238.
4. N. K. Dutta, K. Mazumdar, J. T. Gordy, The nucleocapsid protein of SARS–CoV-2: a target for vaccine development, *J. Virol.*, **94** (2020), e00647-20.
5. D. C. Dinesh, D. Chalupska, J. Silhan, E. Koutna, R. Nencka, V. Veverka, et al., Structural basis of RNA recognition by the SARS-CoV-2 nucleocapsid phosphoprotein, *PLoS Pathog.*, **16** (2020), e1009100.
6. K. S. Saikatendu, J. S. Joseph, V. Subramanian, B. W. Neuman, M. J. Buchmeier, R. C. Stevens, et al., Ribonucleocapsid formation of severe acute respiratory syndrome coronavirus through molecular action of the N-Terminal domain of N protein, *J. Virol.*, **81** (2007), 3913–3921.
7. Y. Xie, D. Du, C. B. Karki, W. Guo, A. E. Lopez-Hernandez, S. Sun, et al., Revealing the mechanism of SARS-CoV-2 spike protein binding with ACE2, *Comput. Sci. Eng.*, **22** (2020), 21–29.
8. Y. Xie, C. B. Karki, D. Du, H. Li, J. Wang, A. Sobitan, et al., Spike proteins of SARS-CoV and SARS-CoV-2 utilize different mechanisms to bind with human ACE2, *Front. Mol. Biosci.*, **7** (2020), e591873.
9. W. Dai, B. Zhang, X. Jiang, H. Su, J. Li, Y. Zhao, et al., Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease, *Science*, **368** (2020), 1331–1335.
10. D. D. Nguyen, K. Gao, J. Chen, R. Wang, G. Wei, Unveiling the molecular mechanism of SARS-CoV-2 main protease inhibition from 137 crystal structures using algebraic topology and deep learning, *Chem. Sci.*, **11** (2020), 12036–12046.
11. K. Gao, D. D. Nguyen, J. Chen, R. Wang, G. Wei, Repositioning of 8565 existing drugs for COVID-19, *J. Phys. Chem. Lett.*, **11** (2020), 5373–5382.

12. L. Li, C. Li, Z. Zhang, E. Alexov, On the dielectric "constant" of proteins: smooth dielectric function for macromolecular modeling and its implementation in DelPhi, *J. Chem. Theory Comput.*, **9** (2013), 2126–2136.

13. L. Li, J. Alper, E. Alexov, Multiscale method for modeling binding phenomena involving large objects: application to kinesin motor domains motion along microtubules, *Sci Rep.*, **6** (2016), 1–12.

14. A. W. Senior, R. Evans, J. Jumper, J. Kirkpatrick, L. Sifre, T. Green, et al., Improved protein structure prediction using potentials from deep learning, *Nature*, **577** (2020), 706–710.

15. J. Wang, C. Karki, Y. Xiao, L. Li, Electrostatics of prokaryotic ribosome and its biological implication, *Biophys J.*, **118** (2020), 1205–1212.

16. M. F. Lensink, S. Velankar, S. J. Wodak, Modeling protein–protein and protein–peptide complexes: CAPRI 6th edition, *Proteins*, **85** (2017), 359–377.

17. L. Li, L. Wang, E. Alexov, On the energy components governing molecular recognition in the framework of continuum approaches, *Front Mol Biosci.*, **2** (2015), e00005.

18. L. Li, D. Guo, Y. Huang, S. Liu, Y. Xiao, ASPDock: protein-protein docking algorithm using atomic solvation parameters model, *BMC Bioinform.*, **12** (2011), 1–9.

19. Y. Peng, E. Alexov, Computational investigation of proton transfer, pKa shifts and pH-optimum of protein–DNA and protein–RNA complexes, *Proteins*, **85** (2017), 282–295.

20. L. Wang, S. Witham, Z. Zhang, L. Li, M. Hodsdon, E. Alexov, In silico investigation of pH-dependence of prolactin and human growth hormone binding to human prolactin receptor, *Commun. Comput. Phys.*, **13** (2013), 207–222.

21. J. Wang, Fast identification of possible drug treatment of Coronavirus Disease-19 (COVID-19) through computational drug repurosing study, *J. Chem. Inf. Model.*, **60** (2020), 3277–3286.

22. G. Eraslan, Ž. Avsec, J. Gagneur, F. J. Theis, Deep learning: new computational modelling techniques for genomics, *Nat Rev Genet.*, **20** (2019), 389–403.

23. L. Li, C. Li, S. Sarkar, J. Zhang, S. Witham, Z. Zhang, et al., DelPhi: a comprehensive suite for DelPhi software and associated resources, *BMC Biophys.*, **5** (2012), 1–11.

24. L. Li, Z. Jia, Y. Peng, A. Chakravorty, L. Sun, E. Alexov, DelPhiForce web server: electrostatic forces and energy calculations and visualization, *Bioinformatics*, **33** (2017), 3661–3663.

25. L. Li, A. Chakravorty, E. Alexov, DelPhiForce, a tool for electrostatic force calculations: Applications to macromolecular binding, *J. Comput. Chem.*, **38** (2017), 584–593.

26. L. Wang, L. Li, E. Alexov, pKa predictions for proteins, RNAs, and DNAs with the Gaussian dielectric function using DelPhi pKa, *Proteins*, **83** (2015), 2186–2197.

27. L. Wang, M. Zhang, E. Alexov, DelPhiPKa web server: predicting pKa of proteins, RNAs and DNAs, *Bioinformatics*, **32** (2015), 614–615.

28. K. Talley, E. Alexov, On the pH-optimum of activity and stability of proteins, *Proteins*, **78** (2010), 2699–2706.

29. R. C. Mitra, Z. Zhang, E. Alexov, In silico modeling of pH-optimum of protein–protein binding, *Proteins*, **79** (2011), 925–936.

30. A. V. Onufriev, E. Alexov, Protonation and pK changes in protein-ligand binding, *Q. Rev. Biophys.*, **46** (2013), 181–209.

31. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, et al., UCSF Chimera—a visualization system for exploratory research and analysis, *J. Comput. Chem.*, **25** (2004), 1605–1612.

32. Y. Xian, Y. Xie, S. M. Silva, C. B. Karki, W. Qiu, L. Li, Structure manipulation tool structureMan: A structure manipulation tool to study large scale biomolecular interactions, *Front. Mol. Biosci.*, **7** (2020), 476.

33. W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, *J. Mol. Graph*, **14** (1996), 33–38.

34. F. Sievers, A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, et al., Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega, *Mol. Syst. Biol.*, **7** (2011), 539.

35. Y. Xian, C. B. Karki, S. M. Silva, L. Li, C. Xiao, The roles of electrostatic interactions in capsid assembly mechanisms of giant viruses. *Int. J. Mol. Sci.*, **20** (2019), 1876.