

Article

Structurally Observed Electrostatic Features of the COVID-19 Coronavirus-Related Experimental Structures inside Protein Data Bank: A Brief Update

Wei Li^{ID}*

¹ Institute of Special Environment Medicine, Nantong University, No. 9, Seyuan Road, Nantong City, Jiangsu Province, P. R. China

* Correspondence: wli148@aucklanduni.ac.nz

Abstract: Since the Coronavirus disease (COVID-19) outbreak at the end of 2019, the past two month has seen an acceleration both in and outside China in the R&D of the diagnostics, vaccines and therapeutics for this novel coronavirus. As one of the molecular forces that determine protein structure, electrostatic effects dominate many aspects of protein behaviour and biological function. Thus, incorporating currently available experimental structures related to COVID-19, this article reports a simple python-based analysis tool and a \LaTeX -based editing tool to extract and summarize the electrostatic features from experimentally determined structures, to strengthen our understanding of COVID-19's structure and function and to facilitate machine-learning and structure-based computational design of its neutralizing antibodies and/or small molecule(s) as potential therapeutic candidates. Finally, this article puts forward a brief update of the structurally observed electrostatic features of the COVID-19 coronavirus.

Keywords: COVID-19; Electrostatic feature; Salt bridging network; Structural update

1. Introduction

On 31 December 2019, WHO was informed of a cluster of cases of pneumonia of unknown cause detected in Wuhan City, Hubei Province of China. The coronavirus disease (COVID-2019) was identified as the causative virus by Chinese authorities on January 7, 2020. The recent emergence of the COVID-2019 coronavirus disease is putting the whole world, instead of China alone, on alert [1–4]. Since the COVID-19 coronavirus disease outbreak, the past two months has seen a series of studies in the R&D of the diagnostics, vaccines and therapeutics for COVID-19 [5–8]. From a structural point of view, this article aims to investigate the electrostatic features embedded in experimentally determined COVID-19 coronavirus-related structures.

2. Materials and Methods

As of March 4, 2020, the Protein Data Bank (PDB) [9] hosts only four COVID-19 coronavirus-related experimental structures, as listed in Table 1 below,

No.	PDB ID	Structure Title
1	6LU7	The crystal structure of covid-19 main protease in complex with an inhibitor n3
2	6LVN	Structure of the 2019-ncov hr2 domain
3	6LXT	Structure of post fusion core of 2019-ncov s2 subunit
4	6VSB	Prefusion 2019-ncov spike glycoprotein with a single receptor-binding domain up

Table 1. Experimentally determined COVID-19 Coronavirus-related structures inside Protein Data Bank (PDB [9]) as of March 4, 2020. All the structures were retrieved from the PDB [9] website with a text search **covid-19**.

Among the four COVID-19 Coronavirus-related structures, three were determined using X-ray crystallography and one using electron microscopy (PDB ID:6VSB), as listed in Table 2.

PDB ID	Molecular IDs	Chain IDs	Experiments
6LU7	1,2	A,C	X-ray crystallography
6LVN	1	A, B, C, D	X-ray crystallography
6LXT	1	A, B, C, D, E, F	X-ray crystallography
6VSB	1	A, B, C	Electron microscopy

Table 2. Molecular and chain IDs of the four COVID-19 Coronavirus-related structures inside PDB as of March 4, 2020.

After the four structures were accessed and downloaded directly from the PDB website [9], a comprehensive set of salt bridging analysis was carried out as described in [10] previously. Nonetheless, no hydrogen bonding network analysis was carried out here because usually hydrogen atoms are missing from experimentally determined biomolecular structures solved using X-ray crystallography.

3. Results

With the electrostatic analysis as described previously in [10], this article puts forward a comprehensive set of electrostatic interaction features for the four COVID-19 coronavirus-related structures as of March 4, 2020, all included as a large set of separate tables in the supplementary file **supplementary.pdf**, including specifically:

1. all salt bridges formed within all four COVID-19 coronavirus-related structures as of March 4, 2020.
2. all interfacial salt bridges formed within all four COVID-19 coronavirus-related structures as of March 4, 2020.
3. all salt bridges formed within all four COVID-19 coronavirus-related structures as of March 4, 2020, PDB ID-specifically for all four COVID-19 coronavirus-related structures.
4. all interfacial salt bridges formed within all four COVID-19 coronavirus-related structures as of March 4, 2020, PDB ID-specifically for all four COVID-19 coronavirus-related structures.

No side chain or main chain hydrogen bond was structurally identified for the four COVID-19 coronavirus-related structures (Tables 2 and 1), as shown in supplementary file **supplementary.pdf**.

4. Conclusion and Discussion

For the first time, this article reports a comprehensive set of electrostatic features sucked out of the currently (as of March 4, 2020) available COVID-19 coronavirus-related structures inside PDB [9] in both PDF format (supplementary file **supplementary.pdf**) and also L^AT_EX format, i.e., a series of machine-readable -importable and -analyzable .tex files zipped in the supplementary file **scan.zip**. In supplementary file **scan.zip**, a simple python-based analysis tool and a L^AT_EX-based editing tool [11] were included to extract and summarize the electrostatic features from experimentally determined structures, where a set of PDB files (representing experimentally determined protein structures) are to be plugged into a set of python scripts included in **cmd.py**. Afterwards, **cmd.py** is to be executed with a simple command (**python cmd.py**) on a Linux machine terminal to produce the final supplementary file **supplementary.pdf**, summarizing the structurally observed electrostatic features of the four COVID-19 coronavirus-related experimental structures as of March 4, 2020.

Finally, along with the python-based analysis tool and a L^AT_EX-based editing tool [11], the structurally observed electrostatic features of the four COVID-19 coronavirus-related experimental structures constitute a preliminary starting point pointing towards a clear, coherent and comprehensive map of COVID-19's structure and function and also machine-learning and structure-based computational design of neutralizing antibodies [11] and/or small molecule(s) as potential therapeutic candidates against future outbreaks of the COVID-19 coronavirus diseases.

Author Contributions: Conceptualization, W.L.; methodology, W.L.; software, W.L.; validation, W.L.; formal analysis, W.L.; investigation, W.L.; resources, W.L.; data duration, W.L.; writing—original draft preparation, W.L.; writing—review and editing, W.L.; visualization, W.L.; supervision, W.L.; project administration, W.L.; funding acquisition, not applicable.

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflict of interest.

1. Chan, J.F.W.; Kok, K.H.; Zhu, Z.; Chu, H.; To, K.K.W.; Yuan, S.; Yuen, K.Y. Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. *Emerging Microbes & Infections* **2020**, *9*, 221–236.
2. Chan, J.F.W.; Yuan, S.; Kok, K.H.; To, K.K.W.; Chu, H.; Yang, J.; Xing, F.; Liu, J.; Yip, C.C.Y.; Poon, R.W.S.; Tsoi, H.W.; Lo, S.K.F.; Chan, K.H.; Poon, V.K.M.; Chan, W.M.; Ip, J.D.; Cai, J.P.; Cheng, V.C.C.; Chen, H.; Hui, C.K.M.; Yuen, K.Y. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *The Lancet* **2020**, *395*, 514–523.
3. Chen, N.; Zhou, M.; Dong, X.; Qu, J.; Gong, F.; Han, Y.; Qiu, Y.; Wang, J.; Liu, Y.; Wei, Y.; Xia, J.; Yu, T.; Zhang, X.; Zhang, L. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *The Lancet* **2020**, *395*, 507–513.
4. Wu, F.; Zhao, S.; Yu, B.; Chen, Y.M.; Wang, W.; Song, Z.G.; Hu, Y.; Tao, Z.W.; Tian, J.H.; Pei, Y.Y.; Yuan, M.L.; Zhang, Y.L.; Dai, F.H.; Liu, Y.; Wang, Q.M.; Zheng, J.J.; Xu, L.; Holmes, E.C.; Zhang, Y.Z. A new coronavirus associated with human respiratory disease in China. *Nature* **2020**.
5. Angeletti, S.; Benvenuto, D.; Bianchi, M.; Giovanetti, M.; Pascarella, S.; Ciccozzi, M. COVID-2019: The role of the nsp2 and nsp3 in its pathogenesis. *Journal of Medical Virology* **2020**.
6. Xu, X.W.; Wu, X.X.; Jiang, X.G.; Xu, K.J.; Ying, L.J.; Ma, C.L.; Li, S.B.; Wang, H.Y.; Zhang, S.; Gao, H.N.; Sheng, J.F.; Cai, H.L.; Qiu, Y.Q.; Li, L.J. Clinical findings in a group of patients infected with the 2019 novel coronavirus (SARS-Cov-2) outside of Wuhan, China: retrospective case series. *BMJ* **2020**, p. m606.
7. Lippi, G.; Plebani, M. Laboratory abnormalities in patients with COVID-2019 infection. *Clinical Chemistry and Laboratory Medicine (CCLM)* **2020**, *0*.
8. Hoffmann, M.; Kleine-Weber, H.; Krüger, N.; Müller, M.; Drosten, C.; Pöhlmann, S. The novel coronavirus 2019 (2019-nCoV) uses the SARS-coronavirus receptor ACE2 and the cellular protease TMPRSS2 for entry into target cells **2020**.
9. Berman, H.; Henrick, K.; Nakamura, H. Announcing the worldwide Protein Data Bank. *Nature Structural & Molecular Biology* **2003**, *10*, 980–980.
10. Li, W. How do SMA-linked mutations of *SMN1* lead to structural/functional deficiency of the SMA protein? *PLOS ONE* **2017**, *12*, e0178519.
11. Li, W. Extracting the Interfacial Electrostatic Features from Experimentally Determined Antigen and/or Antibody-Related Structures inside Protein Data Bank for Machine Learning-Based Antibody Design **2020**.