RESEARCH ARTICLE

JOURNAL OF MEDICAL VIROLOGY WILEY

A guideline for homology modeling of the proteins from newly discovered betacoronavirus, 2019 novel coronavirus (2019-nCoV)

Shengjie Dong[1] | Jiachen Sun[2] | Zhuo Mao[3] | Lu Wang[4] | Yi-Lin Lu[5] | Jiesen Li[6,7] ◉

[1]Faculty of Education and Sports, Guangdong Baiyun University, Guangzhou, China

[2]School of Biotechnology and Food Science, Tianjin University of Commerce, Tianjin, China

[3]Tianjin Key Laboratory of Low Dimensional Materials Physics and Preparing Technology, Department of Applied Physics, Institute of Advanced Materials Physics, Faculty of Science, Tianjin University, Tianjin, China

[4]School of Science, Inner Mongolia University of Science & Technology, Baotou, China

[5]College of New Energy, Bohai University, Jinzhou, China

[6]School of Environment and Chemical Engineering, Foshan University, Foshan, China

[7]Department of Research and Development, Guangzhou Ginpie Technology Co., Ltd., Guangzhou, China

Correspondence
Yi-Lin Lu College of New Energy, Bohai University, 121007 Jinzhou, China.
Email: yilinlu@tju.edu.cn

Jiesen Li School of Environment and Chemical Engineering, Foshan University, 528000 Foshan, China.
Email: 2lgy@163.com

Abstract

During an outbreak of respiratory diseases including atypical pneumonia in Wuhan, a previously unknown β-coronavirus was detected in patients. The newly discovered coronavirus is similar to some β-coronaviruses found in bats but different from previously known SARS-CoV and MERS-CoV. High sequence identities and similarities between 2019-nCoV and SARS-CoV were found. In this study, we searched the homologous templates of all nonstructural and structural proteins of 2019-nCoV. Among the nonstructural proteins, the leader protein (nsp1), the papain-like protease (nsp3), the nsp4, the 3C-like protease (nsp5), the nsp7, the nsp8, the nsp9, the nsp10, the RNA-directed RNA polymerase (nsp12), the helicase (nsp13), the guanine-N7 methyltransferase (nsp14), the uridylate-specific endoribonuclease (nsp15), the 2'-O-methyltransferase (nsp16), and the ORF7a protein could be built on the basis of homology templates. Among the structural proteins, the spike protein (S-protein), the envelope protein (E-protein), and the nucleocapsid protein (N-protein) can be constructed based on the crystal structures of the proteins from SARS-CoV. It is known that PL-Pro, 3CL-Pro, and RdRp are important targets for design antiviral drugs against 2019-nCoV. And S protein is a critical target candidate for inhibitor screening or vaccine design against 2019-nCoV because coronavirus replication is initiated by the binding of S protein to cell surface receptors. It is believed that these proteins should be useful for further structure-based virtual screening and related computer-aided drug development and vaccine design.

KEYWORDS
2019-nCoV, BLAST algorithm, CLUSTAL analysis, coronavirus, homology modeling, MERS-CoV, SARS-CoV, sequence alignment

1 | INTRODUCTION

In the past two decades, the epidemics of the two betacoronaviruses, severe acute respiratory syndrome coronavirus (SARS-CoV) and Middle East respiratory syndrome coronavirus (MERS-CoV) have caused more than 10 000 cumulative cases.[1,2] Very recently, there have been thousands of pneumonia cases in Wuhan, China. These cases of pneumonia were found to be related to a large seafood and

animal market in Wuhan, where local government agencies quickly adopted sanitation and disinfection measures. Now, it is known that these cases of pneumonia were caused by a novel betacoronavirus, that is, the 2019 novel coronavirus (2019-nCoV).[3-12]

Then, the genome sequence and amino acid sequences were originally released by several research groups from different countries.[13-15] The genome of the newly discovered CoV consists of a single, positive-stranded RNA that is about 30k nucleotides long. The overall genome organization of the newly discovered CoV is similar to that of other coronaviruses. The newly sequenced virus genome encodes the open reading frames (ORFs) common to all betacoronaviruses, containing ORF1ab that encodes several enzymatic proteins, the spike-surface glycoprotein (S protein), the small envelope protein (E protein), the matrix protein (M protein), and the nucleocapsid protein (N protein), as well as several nonstructural accessory proteins. Currently, the specific primers and probes for detection 2019-nCoV were synthesized and fabricated against the genetic targets such as ORF1ab and N gene region. According to the gene sequencing data, Wu et al, Tao et al, and Uehara et al initially resolved the primary sequences of ten critical proteins, that is, ORF1ab polyprotein, surface glycoprotein, ORF3a protein, envelope protein, membrane glycoprotein, ORF6 protein, ORF7a protein, ORF8 protein, nucleocapsid phosphoprotein, and ORF10 protein, and published them on National Center for Biotechnology Information (NCBI).

Within the past several months, with the new coronavirus pneumonia becoming more and more serious, the domestic research team is urgently tackling the problem. In recent days, relevant basic research has made rapid progress. Xu et al showed that the 2019-nCoV shared with the SARS/SARS-like coronaviruses a common ancestor that resembles the bat coronavirus HKU9-1. Their work also pointed to the important discovery that the RBD domain of the 2019-nCoV S-protein supports strong interaction with human ACE2 molecules despite its sequence diversity with SARS-CoV S-protein and thus the 2019-nCoV poses a significant public health risk for human transmission via the S-protein–ACE2 binding pathway.[16] On this basis of the structure of 3CL hydrolase (Mpro) of 2019-nCoV, the joint team integrated the strategy of virtual screening and enzymology test, focused on the drug screening for the marketed drugs, as well as the self-built "database of high-yield compounds" and "database of compounds from medicinal plants", and quickly found 30 kinds of possible treatments for 2019-nCoV drugs, active natural products, and traditional Chinese medicine. A joint research team from the structural pharmacology and pharmaceutical chemistry laboratory has been working on the current outbreak of 2019-nCoV pneumonia and their results showed that papain-like protease (PLP) encoded by coronavirus nsp3 played an important role in viral genome replication and escape from the antivirus natural immunity of the host. PLP not only has protein hydrolase activity but also has the activity of de ubiquitinase (DUB). PLP uses its protein hydrolase activity and DUB activity to evade the host's antiviral immune response and inhibit the expression of interferon through a series of molecular mechanisms. It is another enzyme besides the coronavirus 3CL hydrolase, which are important proteins necessary for human infection. To find the potential drugs to inhibit the 2019-nCoV coronavirus, the research team found that the 2019-nCoV PLP sequence has 86% amino acid homology with SARS-CoV PLP

sequence. The research team built the protein structure of 2019-nCoV PLP by using homology modeling method, defined its drug-binding pocket, and screened the active molecules from the existing drug compound library of ZINC, the database of traditional Chinese medicine and natural products established by their own laboratory by using the computer-aided structure-based virtual screening method. Then, 33 potential inhibitors of 2019-nCoV PLP were selected and expected to have possible anti-2019-nCoV activity. Several scientists speculated that Remdesivir may help fight the new coronavirus. The genetic material of the RNA virus is RNA (RNA ribonucleic acid). For the replication of the RNA virus genome, most of them need intermediate synthesis, which requires virus-specific polymerase (RNA dependent RNA polymerase, RdRp, or RNA dependent DNA polymerase, RdDp). These polymerases are ideal targets for the design of antiviral drugs. Remdesivir works by interfering with RdRp. It is a precursor drug of adenosine analog. After triphosphate in vivo, it is incorporated into the newly synthesized RNA chain of the virus as a substrate, and then the synthesis of the virus genome is interrupted.[17-21]

It can be predicted that the research results of computational biology and bioinformatics will play an important role in the process of resistance to new coronavirus.[22,23] Realization and prediction of these structural and nonstructural proteins from 2019-nCoV should be used by more scientific and technological workers, especially those engaged in drug research and development. Using the observed and achieved three-dimensional models of these structural and nonstructural proteins, the screening of potential drugs against 2019-nCoV could be carried out. In this study, using the amino acid sequences released on NCBI, we try to clarify the closest sequences and identify the most suitable template existed for homology modeling of several crucial proteins of 2019-nCoV. This study would be beneficial to the next drug screening research.

2 | MATERIALS, METHODS, AND PROCEDURES

2.1 | Computational tools

The Basic Local Alignment Search Tool (BLAST) could find regions of local similarity between sequences.[24-29] The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance of matches. The maximum number of aligned sequences to display was set to 100. The expected number of chance matches in a random model was set to 10. The length of the seed that initiates an alignment was set to 6. SWISS-MODEL is a fully automated protein structure homology-modeling server, accessible via the ExPASy web server or from the program DeepView (Swiss Pdb-Viewer).[30-33] The purpose of this server is to make protein modeling accessible to all life science researchers worldwide. Building a homology model comprises four main steps: identification of structural template(s), alignment of the target sequence and template structure(s), model-building, and model quality evaluation. Clustal Omega is the latest addition to

the Clustal family.[34,35] Clustal Omega is a multiple sequence alignment program that uses seeded guide trees and HMM profile-profile techniques to generate alignments between three or more sequences. It produces biologically meaningful multiple sequence alignments of divergent sequences. In this study, using three popular online tools, that is, BLAST, SWISS-MODEL, and Clustal Omega, we performed sequence alignment investigations on these 10 primary sequences of 2019-nCoV.

2.2 | Experimental procedures

The structures of the proteins from ORF1ab currently published on RCSB Protein Data Bank (PDB) and yet recommended by Protein Information Resource (PIR) were summarized in Table S1. To construct the proteins from ORF1ab, we distinguished the sequences related to the crystallized structures of the proteins from ORF1ab. After careful evaluation and screening, we compared the sequence of 2GDT, 2GRI, 2ACF, 2JZF, 2FE8, 2K87, 1Q2W, 2AHM, 6NUR, 1QZ8, 5C8S, 6JYT, 2H85, and 2XYQ with the whole sequence of 2019-nCoV ORF1ab using Clustal Omega. After performing sequence alignment, we identified the sequences associated with these PDB files. Then, we searched for the template proteins of these specific sequences using SWISS-MODEL. For proteins from 2019-nCoV other than ORF1ab, we directly searched for the template proteins of their sequences using SWISS-MODEL.

3 | RESULTS AND DISCUSSION

3.1 | Sequence alignment of ORF1ab among 2019-nCoV, SARS-CoV, and MERS-CoV

According to BLAST analysis, the sequence identity of ORF1ab protein between 2019-nCoV and SARS-CoV is more than 90% with the query cover of about 100% while the sequence identity of ORF1ab protein between 2019-nCoV and MERS-CoV is less than 60% with the query cover of about 98%, which proves that the 2019-nCoV shared a better sequence homology toward the sequences of SARS-CoV than that of MERS-CoV, which is consistent with previous research results.[8] Using the Clustal Omega, the sequence alignment of 2019-nCoV and SARS-CoV, 2019-nCoV and MERS-CoV was carried out (see Figures S1 and S2).

3.2 | Homology modeling of the proteins encoded by ORF1ab

Coronavirus family members are important pathogens of many domestic animals, pets, including human beings, causing a variety of acute and chronic diseases. Coronavirus has a large genome RNA, 27 to 32 kb in length, encoding structural proteins, nonstructural proteins, and some auxiliary proteins. The nonstructural protein of coronavirus is hydrolyzed from the virus-encoded polyprotein 1ab. These nonstructural proteins are important functional proteins of coronavirus, which are involved in the regulation of viral genome RNA replication and subgenomic RNA transcription, or involved in genome replication and transcription as transcription/replication complexes, and involved in infection of the host.

It is known that the Replicase polyprotein 1ab (alternative name: ORF1ab polyprotein) could be cleaved into the following 15 chains: (1) Host translation inhibitor nsp1 (alternative names: Leader protein, nsp1), (2) Nonstructural protein 2 (alternative names: p65 homolog, nsp2), (3) Papain-like proteinase (alternative names: PL-Pro, Nonstructural protein 3, nsp3), (4) Nonstructural protein 4 (alternative name: nsp4), (5) 3C-like proteinase (alternative names: 3CL-Pro, nsp5), (6) Nonstructural protein 6 (alternative name: nsp6), (7) Nonstructural protein 7 (alternative name: nsp7), (8) Nonstructural protein 8 (alternative name: nsp8), (9) Nonstructural protein 9 (alternative name: nsp9), (10) Nonstructural protein 10 (alternative names: Growth factor-like peptide, GFL, nsp10), (11) RNA-directed RNA polymerase (alternative names: RdRp, nsp12), (12) Helicase (alternative names: Hel, nsp13), (13) Guanine-N7 methyltransferase (alternative names: ExoN, nsp14), (14) Uridylate-specific endoribonuclease (alternative names: NendoU, nsp15), and (15) 2′-O-methyltransferase (alternative name: nsp16).

Here, we summarized the main findings mentioned in this section on the structure prediction of the proteins encoded by ORF1ab using the homology modeling method in Table 1. Specific measures and experimental details to obtain these valuable results are disclosed in the supplementary documents (also see Figures S3-S62). According to the recent reports mentioned in chapter 1, the PL-Pro, 3CL-Pro, and RdRp should be important targets for design antiviral drugs against 2019-nCoV. In addition, the cartoons of the proteins modeled in this section are shown in Figure 1. We found that the main color of these ribbons is composed of blue, indicating that the structures of these proteins are stretched to a considerable extent.

3.3 | Homology modeling of the other proteins from 2019-nCoV

The diameter of coronavirus is 80 to 120 nm. Under the electron microscope, the surface of virus particles has a club like protrusion, which is composed of spike protein (S protein). The envelope of the virus consists of membrane glycoprotein (M protein), which is embedded in the envelope of the virus through three transmembrane domains. In addition, a small amount of small transmembrane protein envelope protein (E protein) also appeared in the envelope. Finally, the nucleocapsid protein (N protein) binds to the RNA genome in the form of beads, forming a helically symmetric nucleocapsid.

Here, we explored the rationality and feasibility of the homology modeling of other proteins. According to BLAST search results, it is found that SARS spike glycoprotein is a good template for

TABLE 1 Initial reference templates, final recommended templates, and sequence identities information for homology modeling of the ORF1ab related proteins

Target description	Initial reference template	Final recommended template	Sequence identity
Leader protein nsp1	2GDT_A	2HSX_A	86.09%
Papain-like protease nsp3	2GRI_A	2IDY_A	79.28%
Papain-like protease nsp3	2ACF_A	2ACF_A	73.37%
Papain-like protease nsp3	2JZF_A	2WCT_A	79.33%
Papain-like protease nsp3	2KAF_A	2KAF_A	72.31%
Papain-like protease nsp3	2FE8_A	5TI6_A	82.86%
Papain-like protease nsp3	2K87_A	2K87_A	81.74%
Nonstructural protein 4	3VCB_A	3VCB_A	60.23%
3C-like protease nsp5	1Q2W_A	5B6O_A	94.53%
Nonstructural protein 7	2AHM_A	2AHM_A	98.80%
Nonstructural protein 8	2AHM_G	2AHM_G	97.47%
Nonstructural protein 9	1QZ8_A	1UW7_A	97.35%
Nonstructural protein 10	5C8S_A	5C8S_A	98.50%
RNA-directed RNA polymerase nsp12	6NUR_A	6NUR_A	96.35%
Helicase nsp13	6JYT_A	6JYT_A	99.83%
Guanine-N7 methyltransferase nsp14	5C8S_B	5C8S_B	95.07%
Uridylate-specific endoribonuclease nsp15	2H85_A	2H85_A	88.15%
2'-O-methyltransferase nsp16	2XYQ_A	3R24_A	93.24%

constructing 2019-nCoV surface glycoprotein. We also found that MERS spike glycoprotein is similar to 2019-nCoV surface glycoprotein. But the homology of SARS S protein (PDB ID: 6ACC_A) and 2019-nCoV S protein is about 76%, and the query coverage rate is 95%; the homology of MERS S protein (PDB ID: 5X59) and 2019-nCoV S protein is about 34%, and the query coverage rate is 74% (see Figures S63 and S64). Using SWISS-MODEL, 6ACD_C can be used as a potential template for modeling 2019-nCoV S protein (see Figure S65). Then, the homology modeling of S protein could be realized (see Figure S66). Next, we try to perform the homology modeling of ORF3a protein. Nevertheless, although 2019-nCoV ORF3a protein and SARS-CoV ORF3a protein are highly homologous (see Figure S67), the crystal structure of this protein is unknown up to now (see Figure S68). Fortunately, the crystal structure of the template protein for modeling 2019-nCoV E protein is realized experimentally (see Figure S69). Using SWISS-MODEL, the template was found successfully (see Figure S70). The sequence identity is more than 90% (see Figure S71). Regretfully, the crystal structure of the template protein for modeling membrane glycoprotein and ORF6 protein is deficient (see Figures S72-S75). Hopefully, the structure of the template protein of ORF7a protein has been identified experimentally (see Figure S76). Using SWISS-MODEL, the template was found successfully (see Figure S77). The sequence identity is more than 90% (see Figure S78). Currently, there is no suitable template protein for the homology modeling of ORF8 protein (see Figures S79 and S80).

Then, we tried to model the structure of the N protein. We found that the N protein would be divided into two parts and the two parts need to be modeled independently (see Figure S81). The sequence identity is more than 90% (see Figures S82 and S83). ORF10 protein is a short peptide, and there is no template available at present.

Here, we summarized the main findings mentioned in this section on the structure prediction of the other proteins from 2019-nCoV using the homology modeling method in Table 2. Coronavirus replication is initiated by the binding of S protein to cell surface receptors. S protein consists of two functional subunits. S1 (globule) is used for receptor binding and S2 (stem) is used for membrane fusion. The specific interaction between S1 and homologous receptors can trigger the conformational change of the S2 subunit, which leads to the fusion of virus envelope and cell membrane and release of nucleocapsid into the cytoplasm. Receptor binding largely determines the host range and histotropism of coronavirus. Thus, the S protein should be a critical candidate for drug screening or vaccine design against 2019-nCoV.[36,37] In addition, the cartoons of the proteins modeled in this section are shown in Figure 2. We found that the main color of these ribbons is composed of red, demonstrating that there are some squeezing and collision between the atoms in these structures to some degree. Consequently, before these models can be used as targets for computer-aided drug design, conformation optimization based on molecular mechanics or molecular dynamics should be performed sufficiently.
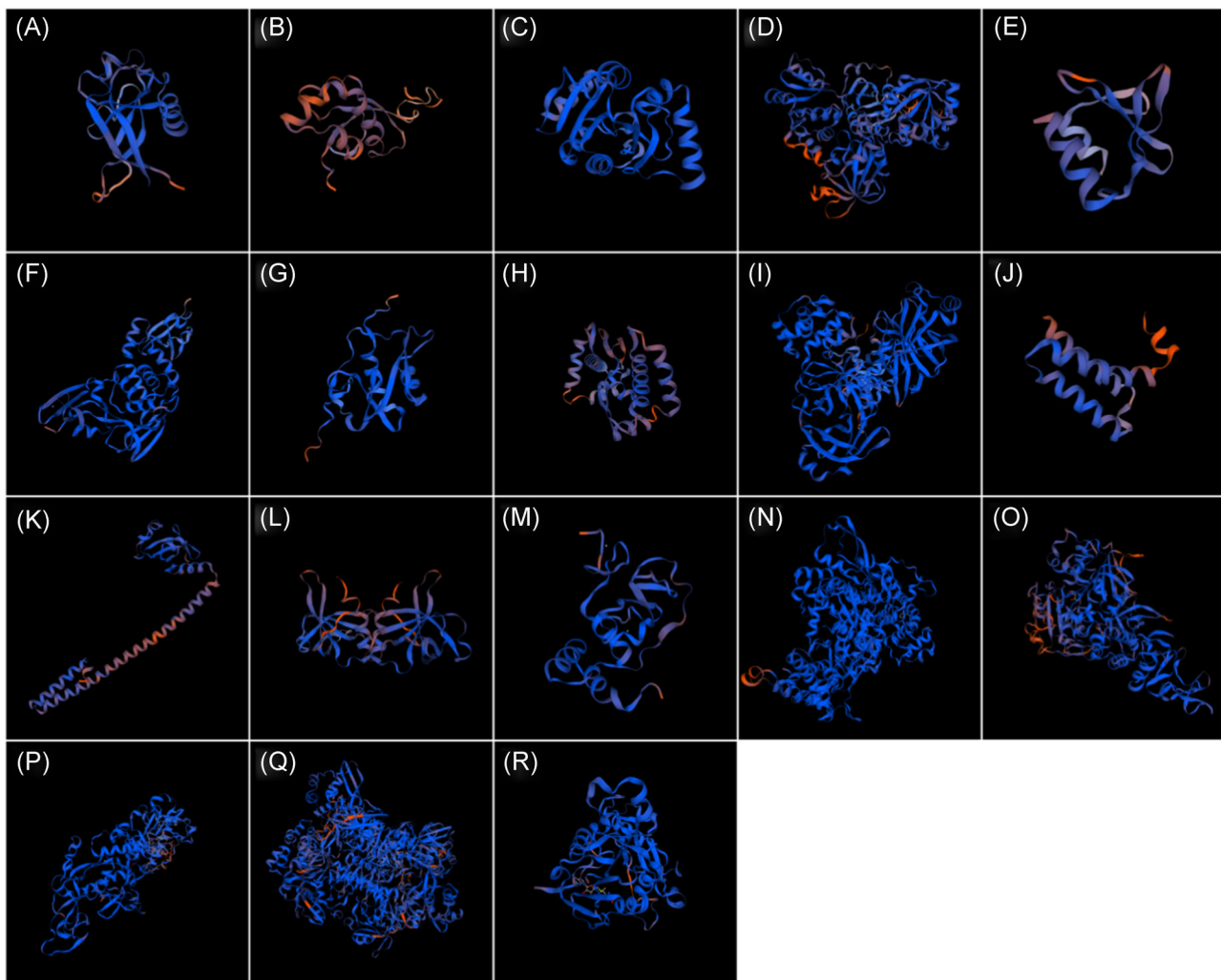
FIGURE 1 Cartoon displayed models of ORF1ab related proteins constructed based on (A) 2HSX_A, (B) 2IDY_A, (C) 2ACF_A, (D) 2WCT_A, (E) 2KAF_A, (F) 5TI6_A, (G) 2K87_A, (H) 3VCB_A, (I) 5B6O_A, (J) 2AHM_A, (K) 2AHM_G, (L) 1UW7_A, (M) 5C8S_A, (N) 6NUR_A, (O) 6JYT_A, (P) 5C8S_B, (Q) 2H85_A, and (R) 3R24_A

4 | SUMMARY AND CONCLUSIONS

COVID-2019 pneumonia continues to spread on a global scale.[38-42] When a frightening new virus appears, in addition to developing new drugs and vaccines, scientists will also consider trying existing antiviral drugs, which may lead to good results. During the outbreak of SARS in 2003, scientists used protease inhibitors to reduce the viral

TABLE 2 Recommended templates and sequence identities of S protein, E protein, ORF7a protein, and N protein

Target description	Recommended template	Sequence identity
S protein	6ACD_A	76.47%
E protein	5X29_A	91.38%
ORF7a	1YO4_A	91.57%
N protein	1SSK_A	92.37%
N protein	2JW8_A	95.76%

load of patients, and protease inhibitors were also proved effective for MERS. Today's new coronaviruses, like SARS and MERS viruses, are all coronaviruses. Therefore, scientists speculate that protease inhibitors can reduce a load of new coronaviruses in patients. Protease inhibitor is a kind of drug targeting protease to treat AIDS. At present, the National Health Commission has recommended that anti-AIDS drugs can be included in the trial. Chinese scientists are testing the anti-HIV drug program to see if it is effective against the new coronavirus. Several research groups from different countries reported their sequencing results and showed that this newly sequenced virus genome encodes the open reading frames (ORFs) common to all β-coronaviruses, including ORF1ab that encodes many enzymatic proteins, the spike-surface glycoprotein (S-protein), the small envelope protein (E-protein), the matrix protein (M-protein), and the nucleocapsid protein (N-protein), as well as several nonstructural proteins.[43] 2019-nCoV shared a better sequence homology toward the sequences of SARS-CoV than that of MERS-CoV. We demonstrated that it is likely that they share considerable sequence
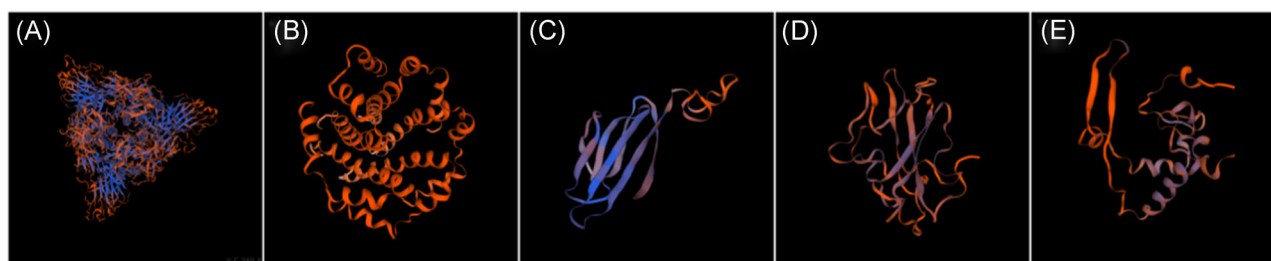
FIGURE 2 Cartoon displayed models of proteins constructed based on (A) 6ACD_A, (B) 5X29_A, (C) 1YO4_A, (D) 1SSK_A, and (E) 2JW8_A

consistent with the SARS or SARS-like coronaviruses, which is in accordance with the previously published work.[44] High sequence identities and similarities between 2019-nCoV and SARS-CoV were found. We searched the homologous templates of all structural and nonstructural proteins of 2019-nCoV. These protein models should be useful for molecular docking investigations and molecular dynamics simulations. Moreover, these resulting structures would be meaningful for further structure-based virtual screening and related computer-aided drug design.

CONFLICT OF INTERESTS

The authors declare that there are no conflict of interests.

AUTHOR CONTRIBUTIONS

SD and YLL designed the project. SD and JL prepared the manuscript. SD and JS performed the calculations. SD and ZM analyzed the data. SD and LW discussed the results.

ORCID

Jiesen Li http://orcid.org/0000-0002-5230-2874

REFERENCES

1. Drosten C, Günther S, Preiser W, et al. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *New Engl J Med*. 2003;348:1967-1976.
2. Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus AD, Fouchier RA. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *New Engl J Med*. 2012;367:1814-1820.
3. Huang CL, Wang YM, Li XW, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet*. 2020; 395(10223):497-506.
4. Chan JFW, Yuan SF, Kok KH, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet*. 2020;395(10223): 514-523.
5. Wang C, Horby PW, Hayden FG, Gao GF. A novel coronavirus outbreak of global health concern. *Lancet*. 2020;395(10223):470-473.
6. Lu RJ, Zhao X, Li J, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet*. 2020;395(10224):565-574.
7. Chen NS, Zhou M, Dong X, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet*. 2020;395(10223):507-513.
8. Zhu N, Zhang DY, Wang WL, et al. A novel coronavirus from patients with pneumonia in China, 2019. *New Engl J Med*. 2020;382:727-733.
9. Rubin EJ, Baden LR, Morrissey S, Campion EW. Medical journals and the 2019-nCoV outbreak. *New Engl J Med*. 2020;382:866.
10. Munster VJ, Koopmans M, van Doremalen N, van Riel D, de Wit E. A novel coronavirus emerging in China–key questions for impact assessment. *New Engl J Med*. 2020;382:692-694.
11. Perlman S. Another decade, another coronavirus. *New Engl J Med*. 2020;382:760-762.
12. Li Q, Guan XH, Wu P, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. *New Engl J Med*. 2020;382:1199-1207. https://doi.org/10.1056/NEJMoa2001316
13. https://www.ncbi.nlm.nih.gov/nuccore/MN908947
14. https://www.ncbi.nlm.nih.gov/nuccore/MN988713
15. https://www.ncbi.nlm.nih.gov/nuccore/MN994468
16. Xu XT, Chen P, Wang JF, et al. Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission. *Sci China Life Sci*. 2020;63:457-460.
17. Cohen J. Can an anti-HIV combination or other existing drugs outwit the new coronavirus? *Science*. 2020. https://doi.org/10.1126/science.abb0659
18. Holshue ML, DeBolt C, Lindquist S, et al. First case of 2019 novel coronavirus in the United States. *New Engl J Med*. 2020;382:929-936.
19. Warren TK, Jordan R, Lo MK, et al. Therapeutic efficacy of the small molecule GS-5734 against Ebola virus in rhesus monkeys. *Nature*. 2016;531:381-385.
20. Sheahan TP, Sims AC, Graham RL, et al. Broad-spectrum antiviral GS-5734 inhibits both epidemic and zoonotic coronaviruses. *Sci Transl Med*. 2017;9(396):eaal3653.
21. Lo MK, Feldmann F, Gary JM, et al. Remdesivir (GS-5734) protects African green monkeys from Nipah virus challenge. *Sci Transl Med*. 2019;11(494):eaau9242.
22. Angeletti S, Benvenuto D, Bianchi M, Giovanetti M, Pascarella S, Ciccozzi M. COVID-2019: the role of the nsp2 and nsp3 in its pathogenesis. *J Med Virol*. 2020:1-5. https://doi.org/10.1002/jmv.25719
23. Liu Z, Xiao X, Wei X, et al. Composition and divergence of coronavirus spike proteins and host ACE2 receptors predict potential intermediate hosts of SARS-CoV-2. *J Med Virol*. 2020:1-7. https://doi.org/10.1002/jmv.25726
24. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990;215(3):403-410.
25. Altschul SF, Madden TL, Schäffer AA, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 1997;25(17):3389-3402.

26. Boratyn GM, Schaffer AA, Agarwala R, Altschul SF, Lipman DJ, Madden TL. Domain enhanced lookup time accelerated BLAST. *Biol Direct.* 2012;7:12.

27. McGinnis S, Madden TL. BLAST: at the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Res.* 2004;32:W20-W25.

28. Ye J, McGinnis S, Madden TL. BLAST: improvements for better sequence analysis. *Nucleic Acids Res.* 2006;34:W6-W9.

29. Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL. NCBI BLAST: a better web interface. *Nucleic Acids Res.* 2008;36:W5-W9.

30. Guex N, Peitsch MC, Schwede T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. *Electrophoresis.* 2009;30:S162-S173.

31. Benkert P, Biasini M, Schwede T. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics.* 2011;27(3):343-350.

32. Bienert S, Waterhouse A, de Beer TAP, et al. The SWISS-MODEL Repository—new features and functionality. *Nucleic Acids Res.* 2017;45:D313-D319.

33. Waterhouse A, Bertoni M, Bienert S, et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 2018;46:W296-W303.

34. Sievers F, Wilm A, Dineen DG, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol.* 2011;7:539.

35. Sievers F, Higgins DG. Clustal Omega for making accurate alignments of many protein sequences. *Protein Sci.* 2018;27(1):135-145.

36. Wrapp D, Wang N, Corbett KS, et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science.* 2020;367:1260-1263. https://doi.org/10.1126/science.abb2507

37. Yan R, Zhang Y, Li Y, Xia L, Guo Y, Zhou Q. Structural basis for the recognition of the SARS-CoV-2 by full-length human ACE2. *Science.* 2020. https://doi.org/10.1126/science.abb2762

38. Chinazzi M, Davis JT, Ajelli M, et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science.* 2020. https://doi.org/10.1126/science.aba9757

39. Paules CI, Marston HD, Fauci AS. Coronavirus infections—more than just the common cold. *JAMA.* 2020;323(8):707-708.

40. Adams JG, Walls RM. Supporting the health care workforce during the COVID-19 global epidemic. *JAMA.* 2020. https://doi.org/10.1001/jama.2020.3972

41. Parodi SM, Liu VX. From containment to mitigation of COVID-19 in the US. *JAMA.* 2020. https://doi.org/10.1001/jama.2020.3882

42. Grasselli G, Pesenti A, Cecconi M. Critical care utilization for the COVID-19 outbreak in Lombardy, Italy: early experience and forecast during an emergency response. *JAMA.* 2020. https://doi.org/10.1001/jama.2020.4031

43. Wu F, Zhao S, Yu B, et al. A new coronavirus associated with human respiratory disease in China. *Nature.* 2020;579:265-269. https://doi.org/10.1038/s41586-020-2008-3

44. Zhou P, Yang XL, Wang XG, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature.* 2020;579:270-273. https://doi.org/10.1038/s41586-020-2012-7

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.