Biophysical *Journal*
New and Notable

Biophysical Society

New and notable: A multiscale coarse-grained model of the SARS-CoV-2 virion

John E. Straub[1,*]
[1]Department of Chemistry, Boston University, Boston, Massachusetts

The global corona virus disease 2019 (COVID-19) pandemic has led to intense activity to develop an effective vaccine. In addition, there have been efforts to develop therapeutics to treat symptoms of the disease, reduce the severity of infection, or prevent infection entirely. Inspired by the potential of structure-based therapeutic design, research has led to detailed molecular models of key protein structural components of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virion (1).

Motivation for this approach comes from the substantial success of past efforts to create effective therapeutics for human immunodeficiency virus (HIV) and influenza using rational design approaches (2,3). An important lesson of that past work is that the most effective rational design approaches account for multiple conformational states of the protein targets and the flexibility of pharmacophores (4,5). As such, the development of atomistic models of virion structures, which enable molecular dynamics simulations that can account for and assess the multiple conformational states defining the virion structural ensemble, are needed (6).

The SARS-CoV-2 virion is composed of spike, membrane, nucleocapsid, and envelope proteins encoded by the virus's 30 kb genome. Significant progress has been made in modeling the spike proteins that mediate receptor recognition and cell membrane fusion. Other proteins play key roles in organizing RNA into ribonucleoprotein complexes, supporting the formation of ion channels in the virion membrane and supporting viral budding. To elucidate the mechanism of viral replication and infectivity, which are dependent on the collective behavior of these many components, a holistic model of the virion is required.

Cryo-electron microscopy (cryo-EM) and x-ray crystallographic studies have been successful in providing insight into the structure of key viral proteins. Because substantial regions of these proteins are left unresolved, homology modeling and structure prediction methods have been used to complete structural models of component proteins (7). Although these experimental and computational methods provide insight into individual protein components, additional work is required to complete an integrated model of the complete SARS-CoV-2 virion.

Complete models of virions are known and have been constructed for HIV-1 (8). Artistic renderings of the molecular structure of the SARS-CoV-2 virion have been widely circulated in scientific writing and popular culture (9). However, until recently we have lacked a detailed atomistic model that is based on state-of-the-art methods for translating the partial knowledge provided by experimental structural biology, homology modeling, and fold prediction into an integrated virion model.

In this issue of the *Biophysical Journal* (10), the groups of Rommie Amaro and Greg Voth report the results of an impressive effort to use state-of-the-art computational methods to develop a complete model of the virion. Employing a "bottom-up" approach, they develop a coarse-grained (CG) model of the SARS-CoV-2 virion based on available experimental structural and atomistic simulation data. All-atom models of the key structural components of the virion were developed based on information from cryo-EM, x-ray crystallographic structures, and homology models. After the construction of models of the key virion components, all-atom simulations were used to inform the construction of coarse-grained models.

A variety of coarse-graining methods pioneered in the Voth group were used to iteratively develop a CG model for the full SARS-CoV-2 virion in a hierarchical manner. The essential dynamics coarse-graining (EDCG) approach was used to develop a coarse-grained model that preserves the principal modes of motion of the system (11). After the initial coarse-graining, the intramolecular interactions within a protein unit were treated using an elastic network model (ENM)

Submitted December 28, 2020, and accepted for publication December 30, 2020.

*Correspondence: straub@bu.edu

Editor: Tamar Schlick.
https://doi.org/10.1016/j.bpj.2020.12.032

(12). In modeling nonbonded interactions, the CG potential energy function was represented as a linear combination of basis functions, in which the coefficients were determined using force-matching (FM) (13) or relative entropy minimization (REM) (14) in the multiscale coarse-graining (MS-CG) approach (15). Finally, to address large-scale conformational changes or bond cleavage, a multistate ultra-coarse-graining (UCG) approach may be effective (16).

Using the constructed virion model, CG dynamics were performed and followed by an analysis of the structure and fluctuations. It was observed that the highest variance principal component was associated with splaying motions in the S1/S2 domain of the spike protein. That motion accounted for 51% of the observed fluctuations in the system. The next principal component due to rocking motions of the S1/S2 domain accounted for 12.5% of the fluctuation. Finally, a principal component associated with twisting of the S1/S2 domain accounted for 7% of the fluctuations. The authors conclude that "these correlated modes of motion are likely informative of how the virion collectively utilizes spike proteins to explore and detect receptors" (10).

This collaborative effort has established a valuable computational resource for the community of researchers focused on studies of SARS-CoV-2 that includes atomistic trajectory and experimental structural data deposited in the National Science Foundation Molecular Sciences Software Institute (MolSSI) as they become publicly available (https://doi.org/10.34974/q8ya-wh69). This resource will serve researchers as a platform for incorporating computational and experimental data.

Given the success of past efforts to develop therapeutics for HIV and influenza using rational design approaches, we can be hopeful that similar approaches will be effective in the development of COVID-19 treatment and prevention strategies. A natural starting point will be multiscale modeling studies of the SARS-CoV-2 virion made possible by the creation of this resource. Those simulations stand to provide insight into structure-function relationships, as well as the generation of structural ensembles for use in rational design approaches accounting for multiple conformational states of this complex molecular assembly.

REFERENCES

1. Amaro, R. E., and A. J. Mulholland. 2020. A community letter regarding sharing biomolecular simulation data for COVID-19. *J. Chem. Inf. Model.* 60:2653–2656.
2. Kaldor, S. W., V. J. Kalish, …, J. H. Tatlock. 1997. Viracept (nelfinavir mesylate, AG1343): a potent, orally bioavailable inhibitor of HIV-1 protease. *J. Med. Chem.* 40:3979–3985.
3. von Itzstein, M., W. Y. Wu, …, C. R. Penn. 1993. Rational design of potent sialidase-based inhibitors of influenza virus replication. *Nature.* 363:418–423.
4. Carlson, H. A., K. M. Masukawa, and J. A. McCammon. 1999. Method for including the dynamic fluctuations of a protein in computer-aided drug design. *J. Phys. Chem. A.* 103:10213–10219.
5. Carlson, H. A., K. M. Masukawa, …, J. A. McCammon. 2000. Developing a dynamic pharmacophore model for HIV-1 integrase. *J. Med. Chem.* 43:2100–2114.
6. Amaro, R. E., J. Baudry, …, J. C. Smith. 2018. Ensemble docking in drug discovery. *Biophys. J.* 114:2271–2278.
7. Heo, L., and M. Feig. 2020. Modeling of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) proteins by machine learning and physics-based refinement. *bioRxiv* https://doi.org/10.1101/2020.03.25.008904.
8. Johnson, G. T., D. S. Goodsell, …, A. J. Olson. 2014. 3D molecular models of whole HIV-1 virions generated with cellPACK. *Faraday Discuss.* 169:23–44.
9. Hays, V. F. 2020. Inside the coronavirus. *Scientific American.* 323:32–37.
10. Yu, A., A. J. Pak, …, G. A. Voth. 2021. A multiscale coarse-grained model of the SARS-CoV-2 virion. *Biophys. J.* 120:1097–1104.
11. Zhang, Z., L. Lu, …, G. A. Voth. 2008. A systematic methodology for defining coarse-grained sites in large biomolecules. *Biophys. J.* 95:5073–5083.
12. Lyman, E., J. Pfaendtner, and G. A. Voth. 2008. Systematic multiscale parameterization of heterogeneous elastic network models of proteins. *Biophys. J.* 95:4183–4192.
13. Noid, W. G., J. -W. Chu, …, H. C. Andersen. 2008. The multiscale coarse-graining method. I. A rigorous bridge between atomistic and coarse-grained models. *J. Chem. Phys.* 128:244114.
14. Shell, M. S. 2008. The relative entropy is fundamental to multiscale and inverse thermodynamic problems. *J. Chem. Phys.* 129:144108.
15. Izvekov, S., and G. A. Voth. 2005. A multiscale coarse-graining method for biomolecular systems. *J. Phys. Chem. B.* 109:2469–2473.
16. Dama, J. F., A. V. Sinitskiy, …, G. A. Voth. 2013. The theory of ultra-coarse-graining. 1. General principles. *J. Chem. Theory Comput.* 9:2466–2480.