

Received 18 March 2021
Accepted 13 April 2021

Edited by K. Moffat, University of Chicago, USA

Keywords: structural biology; SARS-CoV-2; X-ray crystallography; cryoelectron microscopy; COVID-19.

Structural biology in the time of COVID-19: perspectives on methods and milestones

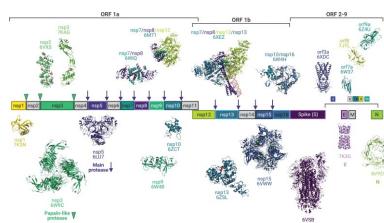
Miranda L. Lynch,^a Edward H. Snell^{a,b} and Sarah E. J. Bowman^{a,c*}

^aHauptman–Woodward Medical Research Institute, 700 Ellicott Street, Buffalo, NY 14203, USA, ^bDepartment of Materials Design and Innovation, The State University of New York at Buffalo, Buffalo, NY 14203, USA, and ^cDepartment of Biochemistry, Jacobs School of Medicine and Biomedical Sciences at The State University of New York at Buffalo, Buffalo, NY 14203, USA. *Correspondence e-mail: sbowman@hwi.buffalo.edu

The global COVID-19 pandemic caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has wreaked unprecedented havoc on global society, in terms of a huge loss of life and burden of morbidity, economic upheaval and social disruption. Yet the sheer magnitude and uniqueness of this event has also spawned a massive mobilization of effort in the scientific community to investigate the virus, to develop therapeutics and vaccines, and to understand the public health impacts. Structural biology has been at the center of these efforts, and so it is advantageous to take an opportunity to reflect on the status of structural science vis-à-vis its role in the fight against COVID-19, to register the unprecedented response and to contemplate the role of structural biology in addressing future outbreak threats. As the one-year anniversary of the World Health Organization declaration that COVID-19 is a pandemic has just passed, over 1000 structures of SARS-CoV-2 biomolecules have been deposited in the Worldwide Protein Data Bank (PDB). It is rare to obtain a snapshot of such intense effort in the structural biology arena and is of special interest as the 50th anniversary of the PDB is celebrated in 2021. It is additionally timely as it overlaps with a period that has been termed the ‘resolution revolution’ in cryoelectron microscopy (CryoEM). CryoEM has recently become capable of producing biomolecular structures at similar resolutions to those traditionally associated with macromolecular X-ray crystallography. Examining SARS-CoV-2 protein structures that have been deposited in the PDB since the virus was first identified allows a unique window into the power of structural biology and a snapshot of the advantages of the different techniques available, as well as insight into the complementarity of the structural methods.

1. Structural methods

In the weeks after the first report of the cluster of pneumonia cases caused by the novel coronavirus, the first SARS-CoV-2 genome sequence was made publicly available (Wu *et al.*, 2020). Within hours, structural biologists worldwide had pivoted to focus on solving the structures of the biomolecules that make up the virus. In a matter of just two weeks, the first three-dimensional structure of one of the critical viral proteins, the main protease, was solved using macromolecular X-ray crystallography (MX) and deposited in the PDB (PDB entry 6lu7), and was released just over a week later (Jin *et al.*, 2020). The first CryoEM structure, of the spike protein, was similarly deposited within weeks of the genome release (PDB entry 6vsb; Wrapp *et al.*, 2020). Structural biology has played crucial roles in impacting our understanding of how the viral mechanisms work, how different viral proteins interact with human host proteins and how ligands bind to the viral proteins (Baker, 2020; Subramaniam, 2020). Structural biologists have



OPEN ACCESS

investigated a substantial number of the proteins encoded by the viral genome, which have provided essential details about how the virus functions (Fig. 1).

The spike protein drives viral entry into host cells and is therefore the main vaccine-development target. It serves as an example of the structural approach, with studies revealing the spike protein in a number of conformations, in complex with the human ACE2 receptor and in complex with human antibodies (Barnes *et al.*, 2020; Shang *et al.*, 2020; Walls *et al.*, 2020; Wrapp *et al.*, 2020; Yuan *et al.*, 2020). Vaccine development has benefitted from 3D structures of the spike protein (Vogel *et al.*, 2020). Once the virus is inside the host cell, it is able to hijack the host cellular machinery directly to begin translation of viral RNA. The largest viral gene (orf1ab) comprises two overlapping reading frames that encode two polyproteins that are cleaved into 16 nonstructural proteins (nsp1–nsp16; Gordon *et al.*, 2020; V'kovski *et al.*, 2020). Two cysteine proteases, the papain-like protease encoded by a portion of nsp3 and the main protease encoded by nsp5, function to process these precursor proteins into all of the constituent nsp proteins required for directing SARS-CoV-2 replication,

transcription and viral assembly. These proteins, including the two proteases, RNA polymerase, helicase and endonuclease, are especially attractive targets for the development of therapeutics (Douangamath *et al.*, 2020; Gao *et al.*, 2020; Gordon *et al.*, 2020; Jin *et al.*, 2020; Kneller, Phillips, O'Neill *et al.*, 2020; Newman *et al.*, 2021; Shin *et al.*, 2020; Schuller *et al.*, 2021; Yin *et al.*, 2020; Zhang, Lin *et al.*, 2020). Significant progress has been made in solving the structures of the nonstructural proteins, including a substantial number of structures with potential inhibitors and fragments bound (Douangamath *et al.*, 2020; Newman *et al.*, 2021; Schuller *et al.*, 2021).

The speed with which these structural investigations have occurred has been breathtaking. Simultaneously, the availability of these data has benefitted from the long history of making structural data and metadata accessible via the PDB, the worldwide data repository for macromolecular structures (Berman *et al.*, 2003). As structural biologists, we have an extensive history of open access and resource sharing; these tendencies and the platforms already established by the community for data sharing have encouraged even more rapid

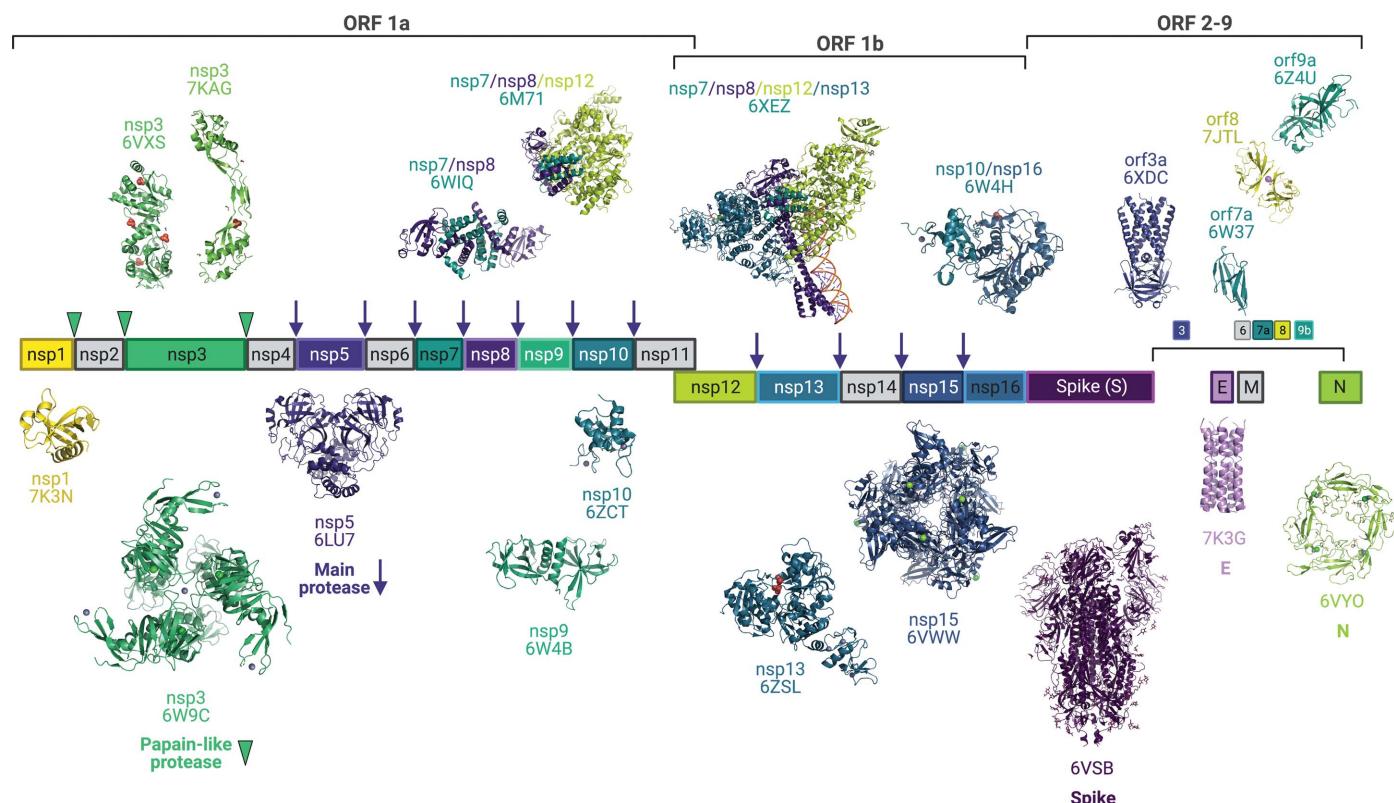


Figure 1
 Map of SARS-CoV-2 protein structures across the open reading frames of the viral genome. Each structure depicted is the first deposited structure of that particular SARS-CoV-2 protein in the PDB [PDB entries 6vxs (Michalska *et al.*, 2020), 7kag (Center for Structural Genomics of Infectious Diseases, unpublished work), 6w1q (Center for Structural Genomics of Infectious Diseases, unpublished work), 6m71 (Gao *et al.*, 2020), 6xez (Chen *et al.*, 2020), 6w4h (Rosas-Lemus *et al.*, 2020), 6xdc (Kern *et al.*, 2021), 7jtl (Flower *et al.*, 2021), 6w37 (Center for Structural Genomics of Infectious Diseases, unpublished work), 6z4u (S. D. Weeks, S. De Graef & A. Munawar, unpublished work), 7k3n (Semper *et al.*, 2021), 6w9c (Center for Structural Genomics of Infectious Diseases, unpublished work), 6lu7 (Jin *et al.*, 2020), 6w4b (Center for Structural Genomics of Infectious Diseases, unpublished work), 6zct (Rogstam *et al.*, 2020), 6zsl (J. A. Newman, Y. Yosaatmadja, A., Douangamath, C. H. Arrowsmith, F. von Delft, A. Edwards, C. Bountra & O. Gileadi, unpublished work), 6vww (Kim *et al.*, 2020), 6vsb (Wrapp *et al.*, 2020), 7k3g (Mandala *et al.*, 2020) and 6vyo (Center for Structural Genomics of Infectious Diseases, unpublished work)]. The cleavage sites of the proteases are shown by green triangles (papain-like protease) and purple arrows (main protease).

dissemination of structural details during the pandemic. Even in our resource-sharing-rich community, the level and swiftness of data sharing over the last year have been unprecedented. There have been several reviews of SARS-CoV-2 protein structures that have been solved and the insights that the structural details provide into viral mechanisms (Bárcena *et al.*, 2021; Mariano *et al.*, 2020; Wang *et al.*, 2020). While the number of structures of SARS-CoV-2 biomolecules deposited in the PDB increases weekly, the trends that we report here have remained consistent since we began tracking these details in early 2021. The goal of this perspective piece is to provide a snapshot of this instant in time, a year into the COVID-19 pandemic, as a window into the structural biology techniques being used and what these outcomes might mean for the future of structural methods.

1.1. Experimental approaches for SARS-CoV-2 biomolecules

The first SARS-CoV-2 structure deposited in the PDB (PDB entry 6lu7; Jin *et al.*, 2020), which was deposited on 26 January 2020, was of the main protease, a major therapeutics development target, and was swiftly followed by 23 more SARS-CoV-2 protein structures in February 2020. Since January 2020, a total of 9243 structures have been deposited and released in the PDB; of these, an astonishing 1038 structures have been of SARS-CoV-2 proteins. In other words, ~11% of all structures deposited in the PDB between 1 January 2020 and 10 March 2021 have been SARS-CoV-2 structures. These numbers paint a compelling picture of the dramatic pivot towards studying SARS-CoV-2 that the structural biology community has made in response to the COVID-19 pandemic. Of additional note in these numbers is the steep decline in the total number of structures that have been deposited during the course of the pandemic. Querying similar time frames over the prior two-year period yields total counts of deposited structures of 14 706 from 1 January 2019 to 11 March 2020 and 13 796 from 1 January 2018 to 13 March 2019. While 9243 structures represent a decrease of over one-third of the expected structure depositions in this time period based on recent history, it also reflects some level of productivity of structures during the pandemic. These structural depositions could comprise a combination of data collected previously that could be worked on during lockdown periods and new data enabled by remote data-collection capabilities. The long-term impacts of the worldwide lockdowns on scientific research remain to be seen in their entirety, but even at this point it is clear that there will be enduring consequences.

The pandemic has given us an unprecedented opportunity to interrogate the contributions of different structural biology methods. This snapshot shows that a year into the pandemic, ~77% of the structural information about SARS-CoV-2 comes from MX and ~23% from CryoEM (Fig. 2). However, ~91% of our knowledge of how ligands interact with viral target proteins comes from MX. In some ways, these outcomes may reflect the availability of established approaches and pipelines. MX methods have benefitted from well established procedures, including high-throughput crystallization, high-

throughput data-collection methods and high-throughput ligand screening. Structural biology plays an important role in structure-based drug design, and the pressing need for the development of vaccines and therapeutics against SARS-CoV-2 enabled the immediate application and extension of these high-throughput methods. Indeed, work by the XChem Team at Diamond Light Source and by the Coronavirus Research Group at UCSF has exploited recently developed fragment-screening MX methods that have been instrumental in the identification of potential chemicals and fragments that bind the main protease, helicase, endonuclease and macrodomain of nsp3, even for fragments that may not be bound as tightly at the active site, may be bound at possible allosteric sites or are found with lower occupancies (Douangamath *et al.*, 2020; Newman *et al.*, 2021; Schuller *et al.*, 2021). These high-throughput fragment-screening efforts have resulted in at least 428 SARS-CoV-2 structures that have been deposited in the PDB with a wide variety of bound fragments.

To further probe the contributions of the structural biology methods MX and CryoEM in the fight against COVID-19, we examined details regarding the deposition date, resolution and molecular weight of SARS-CoV-2 structures (Fig. 3). We anticipated that the initial structural approaches would be biased towards CryoEM, in part because of the well known difficulties in finding crystallization conditions for protein targets. We were surprised to discover that the structures deposited in the first few months of the pandemic were nearly all solved using MX methods (Fig. 3, inset). In general, the lower molecular weight SARS-CoV-2 biomolecules have

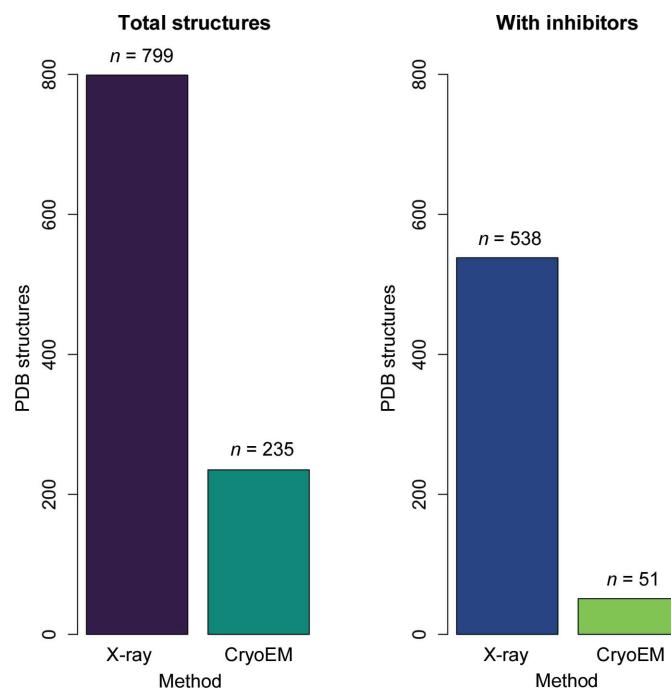


Figure 2
Bar charts of deposited and released structures of SARS-CoV-2 samples. Left: of the 1038 SARS-CoV-2 structures in the PDB as of 10 March 2021, 1034 were solved using MX and CryoEM methods, with a ratio of ~77% to ~23%. Right: when considering the subset of structures with inhibitors or fragments bound, the ratio is ~91% MX and ~9% CryoEM.

primarily been solved using MX techniques, while those at higher molecular weight have primarily been solved using CryoEM methods. The smallest structure solved via CryoEM is an outlier as a 28.3 kDa RNA structure, although its resolution is only 6.9 Å (PDB entry 6xrz; Zhang, Pintilie *et al.*, 2020). Except for this RNA structure, all structures of SARS-CoV-2 samples below 50 kDa have exclusively been solved by MX. Similarly, all structures and complexes above 380 kDa have exclusively been solved by CryoEM. There is a range of biomolecules with molecular-weight values between ~100 and ~220 kDa that are accessible by both methods. This analysis emphasizes the necessity of bringing the entire arsenal of structural techniques to bear and highlights the capabilities and complementarity of the MX and CryoEM methods.

1.2. Computational approaches for SARS-CoV-2 proteins

The physical models of the various SARS-CoV-2 proteins derived by experimental methods have been critical for unlocking the atomic-level details that have underpinned drug discovery and vaccine development. The SARS-CoV-2 pandemic has also shone a light on the many diverse contributions of *in silico* structural biology work (Mulholland & Amaro, 2020). Initial structure-prediction work using *C-I-TASSER* (Zhang, Zheng *et al.*, 2020; Zhang, 2020) and *AlphaFold* from the DeepMind consortium (Jumper *et al.*, 2020; Senior *et al.*, 2020) was able to rapidly leverage the released genomes, determine protein sequences and create computational models that appeared in early February and

March 2020. Molecular-dynamics simulations have massively extended the utility of experimental structural data, providing novel insights via the rapid analysis of *in silico* mutagenesis for many of the key viral proteins (Casalino, Dommer *et al.*, 2020; Rynkiewicz *et al.*, 2021; Sheik Amamuddy *et al.*, 2020) and the ability to model glycosylation of the spike protein (Casalino, Gaieb *et al.*, 2020; Woo *et al.*, 2020). Computational approaches have also helped to probe potential host-pathogen protein interactions (HPIs), contributing network-based and machine-learning-based assessments of putative interactions leveraging multiple HPI databases (Dey *et al.*, 2020; Messina *et al.*, 2020). One ambitious project, *Folding@Home*, has even employed crowdsourcing to access exascale computing resources for multiple SARS-CoV-2 simulation projects (Achdout *et al.*, 2020; Zimmerman & Bowman, 2021). The synergy between computational models and experimental methods, especially in the areas of protein dynamics, glycosylation and protein-interaction studies, has propelled both approaches forward.

2. Materials and methods

The PDB structures in Fig. 1 were generated with *PyMOL* (version 2.4.0; Schrödinger) curated by SBGrid (Morin *et al.*, 2013). Fig. 1 was created using the *BioRender* software. All PDB data in Figs. 2 and 3 were extracted from the PDB on 10 March 2021 (Burley *et al.*, 2020). These data include a total of 1038 SARS-CoV-2 structures that have been deposited and released. Four structures were solved using NMR spectroscopy and are not included in Figs. 2 and 3. The two structures solved by joint neutron and X-ray diffraction methods [PDB entries 7jun (Kneller, Phillips, O'Neill *et al.*, 2020) and 7lb7 (Kneller *et al.*, 2021)] were included in the MX totals. We note that the scatter plot in Fig. 3 does not show structures at >8 Å resolution or with molecular weight >1000 kDa (a total of 16 structures are not shown in the scatter plot). Bar plots and scatter plots were generated using the *R* statistical software (R Core Team, 2020) and the *viridisLite* palette package (Garnier, 2018).

3. Discussion: rumors of the death of X-ray crystallography have been greatly exaggerated

The structural biology community has come together in a tour de force of collegiality and cooperation to rapidly address the COVID-19 pandemic. The past year has seen an impressive leveraging of remote capabilities from crystallization screening supported by our laboratory at Hauptman-Woodward

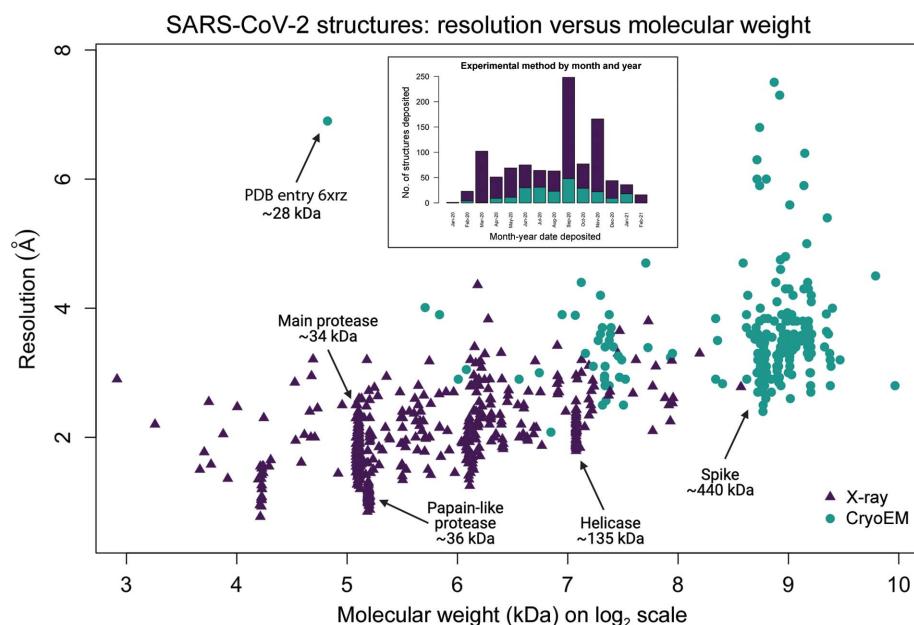


Figure 3

SARS-CoV-2 structures deposited in the PDB. The scatter plot displays resolution versus molecular weight for the deposited structures solved using MX (purple triangles) or CryoEM (teal circles) methods. Proteins with a number of structures deposited are noted; for instance, there are 280 structures of the main protease, over 200 of which are bound to inhibitors or fragments. The outlier RNA CryoEM structure (PDB entry 6xrz) is also noted. Inset: bar plots showing the number of structures deposited and the experimental method used by month and year from January 2020 to February 2021.

Medical Research Institute (Iketani *et al.*, 2021; Kneller, Phillips, O'Neill *et al.*, 2020; Kneller, Phillips, Weiss *et al.*, 2020; Ye *et al.*, 2020) to the remarkable fragment screening and data collection that has occurred with limited staffing at synchrotron beamlines and CryoEM facilities around the world. The availability of open-source materials has had a critical impact on our ability to solve structures of SARS-CoV-2 biomolecules, from the initial distribution of the genome and data hosting on GISAID to structural depositions in the PDB. Indeed, structures were often being deposited and released in the PDB before even preprint manuscripts were released. We have a more refined picture of the individual capabilities and complementarity of our structure-determination methods. The pandemic has revealed that even in the CryoEM resolution revolution (Kühlbrandt, 2014), where it has become capable of producing biomolecular structures at similar resolutions to those traditionally associated with MX (Nakane *et al.*, 2020; Zhang, Pintilie *et al.*, 2020), crystallography excels in the ability to yield high-resolution structures of low-molecular-weight targets and to rapidly screen fragments and inhibitors for structural knowledge. The specific contributions of MX and CryoEM to structures of SARS-CoV-2 biomolecules emphasize how the methods complement each other and have been required in the fight against COVID-19. The high level of information sharing has also empowered the deployment of several online resources, including collaborative efforts to refine initial structural models (Brzezinski *et al.*, 2021; Croll *et al.*, 2020). Computational approaches have enabled the study of glycosylation and emerging variants using MD simulations. Integration of all of these efforts, and the synergy between experimental structural work and computational models, has led to fundamental advancements in pinning down the mechanistic details of the SARS-CoV-2 virus. The fight is ongoing; unsolved SARS-CoV-2 protein structures and emerging variants remain that require the continued attention and skills of the structural biology community.

A major factor at the base of the amazing work that has been accomplished so far was the prior research on other coronaviruses, such as SARS and MERS. Previous research empowered the remarkable response time and underscores the critical foundation of information that made it possible to mobilize in the face of the SARS-CoV-2 outbreak. We are therefore at a unique time not just to reflect on what has been done, but also to contemplate how the community can rapidly face new threats when they emerge. We find ourselves confronted by a set of compelling questions. What have we learned from the pandemic and the collective responses of scientific communities to this worldwide crisis? What features of the pipelines that we have been developing empower sharing and cross-fertilization across disciplines? What other health crises could benefit from the approaches that we have developed? We believe that this is a critical time for thinking about how to continue to build strategies for constructing frameworks for response to emerging threats.

The structural biology response to the COVID-19 pandemic has allowed the world to see the virus that has been attacking them. It has aided vaccine and therapeutic development and

has shown foresight in the investment in structural capabilities, MX and CryoEM. With the intense focus on a single area in a short time period, it has nicely illustrated the complementary strengths of approaches. MX is used for the rapid study of smaller components and the high-resolution study of many potential therapeutic ligand complexes, while CryoEM has shown its strength in larger components. We do not doubt that the currently small overlap of methods illustrated in Fig. 3 will grow over time, as CryoEM continues to develop and new MX capabilities become available. We are well placed to fight the next pandemic, but will be even better placed in the future. Robotic approaches will improve, computational techniques will become more powerful and developments in the MX arena will allow studies at physiological temperatures. When combined, structural biology methods provide us with a complete picture to fight a virus; when techniques are used alone, important areas may be missed.

Acknowledgements

We apologize to those authors whose work we were unable to mention.

Funding information

The following funding is acknowledged: National Science Foundation, Directorate for Biological Sciences (grant No. 2029885 to Miranda L. Lynch; grant No. 2029943 to Sarah E. J. Bowman); National Science Foundation, BioXFEL Science and Technology Center (award No. 1231306 to Edward H. Snell); National Institutes of Health, National Institute of General Medical Sciences (grant No. GM 124135).

References

Achdout, H., Aimon, A., Bar-David, E., Barr, H., Ben-Shmuel, A., Bennett, J., Bobby, M. L., Brun, J., Sarma, B., Calmiano, M., Carbery, A., Cattermole, E., Chodera, J. D., Clyde, A., Coffland, J. E., Cohen, G., Cole, J., Contini, A., Cox, L., Cvitkovic, M., Dias, A., Douangamath, A., Duberstein, S., Dudgeon, T., Dunnett, L., Eastman, P. K., Erez, N., Fairhead, M., Fearon, D., Fedorov, O., Ferla, M., Foster, H., Foster, R., Gabizon, R., Gehrtz, P., Gileadi, C., Giroud, C., Glass, W. G., Glen, R., Glinert, I., Gorichko, M., Gorrie-Stone, T., Griffen, E. J., Heer, J., Hill, M., Horrell, S., Hurley, M. F. D., Israely, T., Jajack, A., Jnoff, E., John, T., Kantsadi, A. L., Kenny, P. W., Kiappes, J. L., Koekemoer, L., Kovar, B., Krojer, T., Lee, A. A., Lefker, B. A., Levy, H., London, N., Lukacik, P., Macdonald, H. B., MacLean, B., Malla, T. R., Matviiuk, T., McCorkindale, W., Melamed, S., Michurin, O., Mikolajek, H., Morris, A., Morris, G. M., Morwitzer, M. J., Moustakas, D., Neto, J. B., Oleinikovas, V., Overheul, G. J., Owen, D., Pai, R., Pan, J., Paran, N., Perry, B., Pingle, M., Pinjari, J., Politi, B., Powell, A., Psenak, V., Puni, R., Rangel, V. L., Reddi, R. N., Reid, S. P., Resnick, E., Robinson, M. C., Robinson, R. P., Rufa, D., Schofield, C., Shaikh, A., Shi, J., Shurrush, K., Sittner, A., Skyner, R., Smalley, A., Smilova, M. D., Spencer, J., Strain-Damerell, C., Swamy, V., Tamir, H., Tennant, R., Thompson, A., Thompson, W., Tomasio, S., Tumber, A., Vakonakis, I., van Rij, R. P., Varghese, F. S., Vaschetto, M., Vitner, E. B., Voelz, V., von Delft, A., von Delft, F., Walsh, M., Ward, W., Weatherall, C., Weiss, S., Wild, C. F., Wittmann, M., Wright, N., Yahalom-Ronen, Y., Zaidmann, D., Zidane, H. & Zitzmann, N. (2020). *bioRxiv*, 2020.10.29.339317.

Baker, E. N. (2020). *IUCrJ*, **7**, 366–367.

Bárcena, M., Barnes, C. O., Beck, M., Bjorkman, P. J., Canard, B., Gao, G. F., Gao, Y., Hilgenfeld, R., Hummer, G., Patwardhan, A., Santoni, G., Saphire, E. O., Schaffitzel, C., Schendel, S. L., Smith, J. L., Thorn, A., Veesler, D., Zhang, P. & Zhou, Q. (2021). *Nat. Struct. Mol. Biol.* **28**, 2–7.

Barnes, C. O., Jette, C. A., Abernathy, M. E., Dam, K. A., Esswein, S. R., Gristick, H. B., Malyutin, A. G., Sharaf, N. G., Huey-Tubman, K. E., Lee, Y. E., Robbiani, D. F., Nussenzweig, M. C., West, A. P. & Bjorkman, P. J. (2020). *Nature*, **588**, 682–687.

Berman, H., Henrick, K. & Nakamura, H. (2003). *Nat. Struct. Mol. Biol.* **10**, 980.

Brzezinski, D., Kowiel, M., Cooper, D. R., Cymborowski, M., Grabowski, M., Włodawer, A., Dauter, Z., Shabalin, I. G., Gilski, M., Rupp, B., Jaskolski, M. & Minor, W. (2021). *Protein Sci.* **30**, 115–124.

Burley, S. K., Bhikadiya, C., Bi, C., Bittrich, S., Chen, L., Crichlow, G. V., Christie, C. H., Dalenberg, K., Di Costanzo, L., Duarte, J. M., Dutta, S., Feng, Z., Ganesan, S., Goodsell, D. S., Ghosh, S., Green, R. K., Guranović, V., Guzenko, D., Hudson, B. P., Lawson, C., Liang, Y., Lowe, R., Namkoong, H., Peisach, E., Persikova, I., Randle, C., Rose, A., Rose, A., Sali, A., Segura, J., Sekharan, M., Shao, C., Tao, Y., Voigt, M., Westbrook, J., Young, J. Y., Zardecki, C. & Zhuravleva, M. (2020). *Nucleic Acids Res.* **49**, D437–D451.

Casalino, L., Dommer, A., Gaieb, Z., Barros, E. P., Sztain, T., Ahn, S.-H., Trifan, A., Brace, A., Bogetti, A., Ma, H., Lee, H., Turilli, M., Khalid, S., Chong, L., Simmerling, C., Hardy, D. J., Maia, J. D. C., Phillips, J. C., Kurth, T., Stern, A., Huang, L., McCalpin, J., Tatineni, M., Gibbs, T., Stone, J. E., Jha, S., Ramanathan, A. & Amaro, R. E. (2020). *bioRxiv*, 2020.11.19.390187.

Casalino, L., Gaieb, Z., Goldsmith, J. A., Hjorth, C. K., Dommer, A. C., Harbison, A. M., Fogarty, C. A., Barros, E. P., Taylor, B. C. & McLellan, J. S. (2020). *ACS Cent. Sci.* **6**, 1722–1734.

Chen, J., Malone, B., Llewellyn, E., Grasso, M., Shelton, P. M., Olinares, P. D. B., Maruthi, K., Eng, E. T., Vatandaslar, H., Chait, B. T., Kapoor, T. M., Darst, S. A. & Campbell, E. A. (2020). *Cell*, **182**, 1560–1573.

Croll, T., Diederichs, K., Fischer, F., Fyfe, C., Gao, Y., Horrell, S., Joseph, A. P., Kandler, L., Kippes, O., Kirsten, F., Müller, K., Nolte, K., Payne, A., Reeves, M., Richardson, J., Santoni, G., Stäb, S., Tronrud, D., von Soosten, L., Williams, C. & Thorn, A. (2020). *bioRxiv*, 2020.10.07.307546.

Dey, L., Chakraborty, S. & Mukhopadhyay, A. (2020). *Biomed. J.* **43**, 438–450.

Douangamath, A., Fearon, D., Gehrtz, P., Krojer, T., Lukacik, P., Owen, C. D., Resnick, E., Strain-Damerell, C., Aimon, A., Ábrányi-Balogh, P., Brandão-Neto, J., Carbery, A., Davison, G., Dias, A., Downes, T. D., Dunnnett, L., Fairhead, M., Firth, J. D., Jones, S. P., Keeley, A., Keserü, G. M., Klein, H. F., Martin, M. P., Noble, M. E. M., O'Brien, P., Powell, A., Reddi, R. N., Skyner, R., Snee, M., Waring, M. J., Wild, C., London, N., von Delft, F. & Walsh, M. A. (2020). *Nat. Commun.* **11**, 5047.

Flower, T. G., Buffalo, C. Z., Hooy, R. M., Allaire, M., Ren, X. & Hurley, J. H. (2021). *Proc. Natl Acad. Sci. USA*, **118**, e2021785118.

Gao, Y., Yan, L., Huang, Y., Liu, F., Zhao, Y., Cao, L., Wang, T., Sun, Q., Ming, Z., Zhang, L., Ge, J., Zheng, L., Zhang, Y., Wang, H., Zhu, Y., Zhu, C., Hu, T., Hua, T., Zhang, B., Yang, X., Li, J., Yang, H., Liu, Z., Xu, W., Guddat, L. W., Wang, Q., Lou, Z. & Rao, Z. (2020). *Science*, **368**, 779–782.

Garnier, S. (2018). *viridisLite: Colorblind-Friendly Color Maps (Lite Version)*. <https://cran.r-project.org/package=viridisLite>.

Gordon, D. E., Jang, G. M., Bouhaddou, M., Xu, J., Obernier, K., White, K. M., O'Meara, M. J., Rezelj, V. V., Guo, J. Z., Swaney, D. L., Tummino, T. A., Hüttenhain, R., Kaake, R. M., Richards, A. L., Tutuncuoglu, B., Foussard, H., Batra, J., Haas, K., Modak, M., Kim, M., Haas, P., Polacco, B. J., Braberg, H., Fabius, J. M., Eckhardt, M., Soucheray, M., Bennett, M. J., Cakir, M., McGregor, M. J., Li, Q., Meyer, B., Roesch, F., Vallet, T., Mac Kain, A., Miorin, L., Moreno, E., Naing, Z. Z. C., Zhou, Y., Peng, S., Shi, Y., Zhang, Z., Shen, W., Kirby, I. T., Melnyk, J. E., Chorba, J. S., Lou, K., Dai, S. A., Barrio-Hernandez, I., Memon, D., Hernandez-Armenta, C., Lyu, J., Mathy, C. J. P., Perica, T., Pilla, K. B., Ganesan, S. J., Saltzberg, D. J., Rakesh, R., Liu, X., Rosenthal, S. B., Calviello, L., Venkataraman, S., Liboy-Lugo, J., Lin, Y., Huang, X.-P., Liu, Y., Wankowicz, S. A., Bohn, M., Safari, M., Ugur, F. S., Koh, C., Savar, N. S., Tran, Q. D., Shengjuler, D., Fletcher, S. J., O'Neal, M. C., Cai, Y., Chang, J. C. J., Broadhurst, D. J., Klippsten, S., Sharp, P. P., Wenzell, N. A., Kuzuoglu-Ozturk, D., Wang, H.-Y., Trenker, R., Young, J. M., Caverio, D. A., Hiatt, J., Roth, T. L., Rathore, U., Subramanian, A., Noack, J., Hubert, M., Stroud, R. M., Frankel, A. D., Rosenberg, O. S., Verba, K. A., Agard, D. A., Ott, M., Emerman, M., Jura, N., von Zastrow, M., Verdin, E., Ashworth, A., Schwartz, O., d'Enfert, C., Mukherjee, S., Jacobson, M., Malik, H. S., Fujimori, D. G., Ideker, T., Craik, C. S., Floor, S. N., Fraser, J. S., Gross, J. D., Sali, A., Roth, B. L., Ruggero, D., Taunton, J., Kortemme, T., Beltrao, P., Vignuzzi, M., García-Sastre, A., Shokat, K. M., Shoichet, B. K. & Krogan, N. J. (2020). *Nature*, **583**, 459–468.

Iketani, S., Forouhar, F., Liu, H., Hong, S. J., Lin, F.-Y., Nair, M. S., Zask, A., Huang, Y., Xing, L., Stockwell, B. R., Chavez, A. & Ho, D. D. (2021). *Nat. Commun.* **12**, 2016.

Jin, Z., Du, X., Xu, Y., Deng, Y., Liu, M., Zhao, Y., Zhang, B., Li, X., Zhang, L., Peng, C., Duan, Y., Yu, J., Wang, L., Yang, K., Liu, F., Jiang, R., Yang, X., You, T., Liu, X., Yang, X., Bai, F., Liu, H., Liu, X., Guddat, L. W., Xu, W., Xiao, G., Qin, C., Shi, Z., Jiang, H., Rao, Z. & Yang, H. (2020). *Nature*, **582**, 289–293.

Jumper, J., Tunyasuvunakool, K., Kohli, P., Hassabis, D. & The AlphaFold Team (2020). *Computational Predictions of Protein Structures Associated with COVID-19*. <https://deepmind.com/research/open-source/computational-predictions-of-protein-structures-associated-with-COVID-19>.

Kern, D. M., Sorum, B., Mali, S. S., Hoel, C. M., Sridharan, S., Remis, J. P., Toso, D. B., Kotecha, A., Bautista, D. M. & Brohawn, S. G. (2021). *bioRxiv*, 2020.06.17.156554.

Kim, Y., Jedrzejczak, R., Maltseva, N. I., Wilamowski, M., Endres, M., Godzik, A., Michalska, K. & Joachimiak, A. (2020). *Protein Sci.* **29**, 1596–1605.

Kneller, D. W., Phillips, G., O'Neill, H. M., Jedrzejczak, R., Stols, L., Langan, P., Joachimiak, A., Coates, L. & Kovalevsky, A. (2020). *Nat. Commun.* **11**, 1–6.

Kneller, D. W., Phillips, G., Weiss, K. L., Pant, S., Zhang, Q., O'Neill, H. M., Coates, L. & Kovalevsky, A. (2020). *J. Biol. Chem.* **295**, 17365–17373.

Kneller, D. W., Phillips, G., Weiss, K. L., Zhang, Q., Coates, L. & Kovalevsky, A. (2021). *J. Med. Chem.*, <https://doi.org/10.1021/acs.jmedchem.1c00058>.

Kühlbrandt, W. (2014). *Science*, **343**, 1443–1444.

Mandala, V. S., McKay, M. J., Shcherbakov, A. A., Dregni, A. J., Kolocuris, A. & Hong, M. (2020). *Nat. Struct. Mol. Biol.* **27**, 1202–1208.

Mariano, G., Farthing, R. J., Lale-Farjat, S. L. M. & Bergeron, J. R. C. (2020). *Front. Mol. Biosci.* **7**, 605236.

Messina, F., Giombini, E., Agrati, C., Vairo, F., Ascoli Bartoli, T., Al Moghazi, S., Piacentini, M., Locatelli, F., Kobinger, G., Maeurer, M., Zumla, A., Capobianchi, M. R., Lauria, F. N., Ippolito, G. & COVID 19 INMI Network Medicine for IDs Study Group (2020). *J. Transl. Med.* **18**, 233.

Michalska, K., Kim, Y., Jedrzejczak, R., Maltseva, N. I., Stols, L., Endres, M. & Joachimiak, A. (2020). *IUCrJ*, **7**, 814–824.

Morin, A., Eisenbraun, B., Key, J., Sanschagrin, P. C., Timony, M. A., Ottaviano, M. & Sliz, P. (2013). *eLife*, **2**, e01456.

Mulholland, A. J. & Amaro, R. E. (2020). *J. Chem. Inf. Model.* **60**, 5724–5726.

Nakane, T., Kotecha, A., Sente, A., McMullan, G., Masiulis, S., Brown, P. M., Grigoras, I. T., Malinauskaitė, L., Malinauskas, T., Miehling, J., Uchański, T., Yu, L., Karia, D., Pechnikova, E. V., de Jong, E., Keizer, J., Bischoff, M., McCormack, J., Tiemeijer, P., Hardwick,

S. W., Chirgadze, D. Y., Murshudov, G., Aricescu, A. R. & Scheres, S. H. W. (2020). *Nature*, **587**, 152–156.

Newman, J. A., Douangamath, A., Yazdani, S., Yosaatmadja, Y., Aimon, A., Brandao-Neto, J., Dunnett, L., Gorrie-Stone, T., Skyner, R., Fearon, D., Schapira, M., von Delft, F. & Gileadi, O. (2021). *bioRxiv*, 2021.03.15.435326.

R Core Team (2020). *The R Project for Statistical Computing*. <https://www.r-project.org/>.

Rogstam, A., Nyblom, M., Christensen, S., Sele, C., Talibov, V. O., Lindvall, T., Rasmussen, A. A., André, I., Fisher, Z., Knecht, W. & Kozielski, F. (2020). *Int. J. Mol. Sci.* **21**, 7375.

Rosas-Lemus, M., Minasov, G., Shuvalova, L., Inniss, N. L., Kiryukhina, O., Brunzelle, J. & Satchell, K. J. F. (2020). *Sci. Signal.* **13**, eabe1202.

Rynkiewicz, P., Babbitt, G. A., Cui, F. O., Hudson, A. & Lynch, M. L. (2021). *bioRxiv*, 2020.09.11.293258.

Schuller, M., Correy, G. J., Gahbauer, S., Fearon, D., Wu, T., Díaz, R. E., Young, I. D., Martins, L. C., Smith, D. H., Schulze-Gahmen, U., Owens, T. W., Deshpande, I., Merz, G. E., Thwin, A. C., Biel, J. T., Peters, J. K., Moritz, M., Herrera, N., Kratochvil, H. T., Aimon, A., Bennett, J. M., Brandao Neto, J., Cohen, A. E., Dias, A., Douangamath, A., Dunnett, L., Fedorov, O., Ferla, M. P., Fuchs, M., Gorrie-Stone, T. J., Holton, J. M., Johnson, M. G., Krojer, T., Meigs, G., Powell, A. J., Rack, J. G. M., Rangel, V. L., Russi, S., Skyner, R. E., Smith, C. A., Soares, A. S., Wierman, J. L., Zhu, K., Jura, N., Ashworth, A., Irwin, J., Thompson, M. C., Gestwicki, J. E., von Delft, F., Shoichet, B. K., Fraser, J. S. & Ahel, I. (2021). *Sci. Adv.* **7**, eabf87112.

Semper, C., Watanabe, N. & Savchenko, A. (2021). *iScience*, **24**, 101903.

Senior, A. W., Evans, R., Jumper, J., Kirkpatrick, J., Sifre, L., Green, T., Qin, C., Žídek, A., Nelson, A. W., Bridgland, A., Penedones, H., Petersen, S., Simonyan, K., Crossan, S., Kohli, P., Jones, D. T., Silver, D., Kavukcuoglu, K. & Hassabis, D. (2020). *Nature*, **577**, 706–710.

Shang, J., Ye, G., Shi, K., Wan, Y., Luo, C., Aihara, H., Geng, Q., Auerbach, A. & Li, F. (2020). *Nature*, **581**, 221–224.

Sheik Amamuddy, O., Verkhivker, G. M. & Tastan Bishop, O. (2020). *J. Chem. Inf. Model.* **60**, 5080–5102.

Shin, D., Mukherjee, R., Grewe, D., Bojkova, D., Baek, K., Bhattacharya, A., Schulz, L., Widera, M., Mehdipour, A. R., Tascher, G., Geurink, P. P., Wilhelm, A., van der Heden van Noort, G. J., Ovaa, H., Müller, S., Knobeloch, K.-P., Rajalingam, K., Schulman, B. A., Cinatl, J., Hummer, G., Ciesek, S. & Dikic, I. (2020). *Nature*, **587**, 657–662.

Subramanian, S. (2020). *IUCrJ*, **7**, 575–576.

V'kovski, P., Kratzel, A., Steiner, S., Stalder, H. & Thiel, V. (2021). *Nat. Rev. Microbiol.* **19**, 155–170.

Vogel, A. B., Kanevsky, I., Che, Y., Swanson, K. A., Muik, A., Vormehr, M., Kranz, L. M., Walzer, K. C., Hein, S., Güler, A., Loschko, J., Maddur, M. S., Tompkins, K., Cole, J., Lui, B. G., Ziegenhals, T., Plaschke, A., Eisel, D., Dany, S. C., Fesser, S., Erbar, S., Bates, F., Schneider, D., Jesionek, B., Sänger, B., Wallisch, A.-K., Feuchter, Y., Junginger, H., Krumm, S. A., Heinen, A. P., Adams-Quack, P., Schlereth, J., Kröner, C., Hall-Ursone, S., Brasky, K., Griffor, M. C., Han, S., Lees, J. A., Mashalidis, E. H., Sahasrabudhe, P. V., Tan, C. Y., Pavliakova, D., Singh, G., Fontes-Garfias, C., Pride, M., Scully, I. L., Ciolino, T., Obregon, J., Gazi, M., Carrion, R., Alfson, K. J., Kalina, W. V., Kaushal, D., Shi, P.-Y., Klamp, T., Rosenbaum, C., Kuhn, A. N., Türeci, Ö., Dörmitzer, P. R., Jansen, K. U. & Sahin, U. (2020). *bioRxiv*, 2020.09.08.280818.

Walls, A. C., Park, Y.-J., Tortorici, M. A., Wall, A., McGuire, A. T. & Veesler, D. (2020). *Cell*, **181**, 281–292.

Wang, M., Zhao, R., Gao, L., Gao, X., Wang, D. & Cao, J. (2020). *Front. Cell. Infect. Microbiol.* **10**, 587269.

Woo, H., Park, S.-J., Choi, Y. K., Park, T., Tanveer, M., Cao, Y., Kern, N. R., Lee, J., Yeom, M. S., Croll, T. I., Seok, C. & Im, W. (2020). *J. Phys. Chem. B*, **124**, 7128–7137.

Wrapp, D., Wang, N., Corbett, K. S., Goldsmith, J. A., Hsieh, C.-L., Abiona, O., Graham, B. S. & McLellan, J. S. (2020). *Science*, **367**, 1260–1263.

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. (2020). *Nature*, **579**, 265–269.

Ye, G., Gallant, J. P., Massey, C., Shi, K., Tai, W., Zheng, J., Odle, A. E., Vickers, M. A., Shang, J., Wan, Y., Drelich, A., Kempaiah, K. R., Tat, V., Perlman, S., Du, L., Tseng, C.-T., Aihara, H., LeBeau, A. M. & Li, F. (2020). *bioRxiv*, 2020.11.17.386532.

Yin, W., Mao, C., Luan, X., Shen, D.-D., Shen, Q., Su, H., Wang, X., Zhou, F., Zhao, W., Gao, M., Chang, S., Xie, Y.-C., Tian, G., Jiang, H.-W., Tao, S.-C., Shen, J., Jiang, Y., Jiang, H., Xu, Y., Zhang, S., Zhang, Y. & Xu, H. E. (2020). *Science*, **368**, 1499–1504.

Yuan, M., Wu, N. C., Zhu, X., Lee, C.-C. D., So, R. T. Y., Lv, H., Mok, C. K. P. & Wilson, I. A. (2020). *Science*, **368**, 630–633.

Zhang, C., Zheng, W., Huang, X., Bell, E. W., Zhou, X. & Zhang, Y. (2020). *J. Proteome Res.* **19**, 1351–1360.

Zhang, K., Pintilie, G. D., Li, S., Schmid, M. F. & Chiu, W. (2020). *Cell Res.* **30**, 1136–1139.

Zhang, L., Lin, D., Sun, X., Curth, U., Drosten, C., Sauerhering, L., Becker, S., Rox, K. & Hilgenfeld, R. (2020). *Science*, **368**, 409–412.

Zhang, Y. (2020). *Modeling of the SARS-CoV-2 Genome Using I-TASSER*. <https://zhanglab.ccmb.med.umich.edu/COVID-19/>.

Zimmerman, M. I. & Bowman, G. (2021). *Biophys. J.* **120**, 299a.