World's physics instruments turn their focus to COVID-19

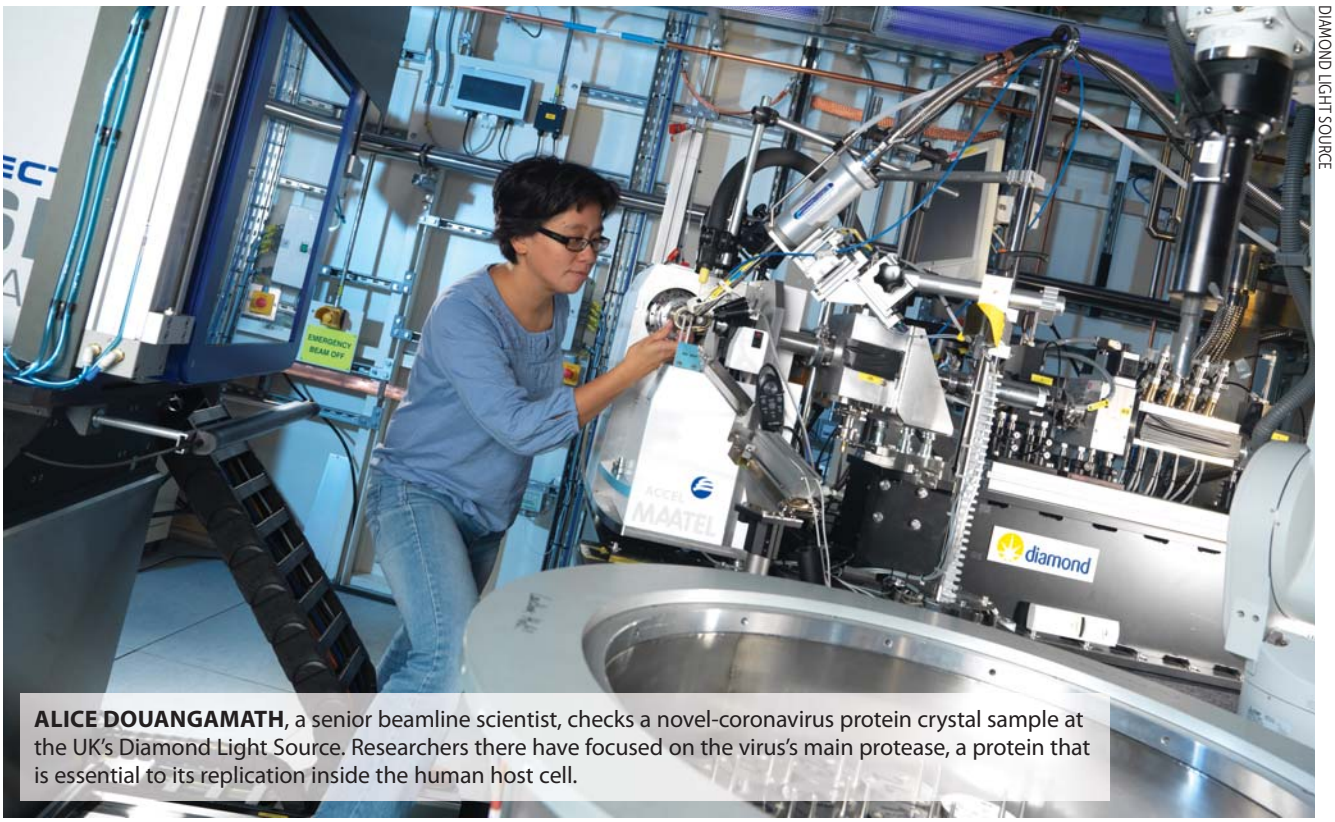
Scientists are employing x rays, electrons, and neutrons to decipher and disable the molecular machinery of the novel coronavirus.

Although most basic research has been suspended by the coronavirus pandemic, some labs remain open to engage in a furious effort to find treatments for the disease. Physicists and chemists are vital to a key part of that quest: decoding the three-dimensional structures of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) proteins and finding locations where drugs could latch on and disable the viral machinery. The virus itself is not used for those experiments, only the cloned proteins that are its principal working parts.

As of early April, the life sciences beamlines were open for SARS-CoV-2 research at the UK's Diamond Light Source. In the US, at least 23 groups were working at the Advanced Photon Source at Argonne National Laboratory as of press time, according to Bob Fischetti, life sciences adviser at the APS. The National Synchrotron Light Source II at Brookhaven National Laboratory, SLAC's Stanford Synchrotron Radiation Lightsource (SSRL), and the Advanced Light Source at Lawrence Berkeley National Laboratory are operating with minimal staff, each keeping open several x-ray protein crystallography beamlines strictly for coronavirus research.

The BESSY II light source in Berlin closed briefly but resumed operations on 2 April for coronavirus research, which is also ongoing at the Shanghai synchrotron in China, where the first 3D structure of the main protease protein was resolved. Officials there did not respond to requests for comment. The European Synchrotron Radiation Facility in France has been closed for an upgrade, but it announced in early April that it would consider reopening beamlines on a case-by-case basis for coronavirus research.

Structures of many of the virus's 28 or 29 proteins (estimates vary on the exact number) have been resolved, both alone and in complexes with various molecules, known as ligands, that bind to them. Among those resolved structures are the main protease (M^{pro}), an enzyme that processes long viral polyproteins into shorter functional units; an endoribonuclease called Nsp15; and the spike protein that protrudes from the coronavirus



DIAMOND LIGHT SOURCE

ALICE DOUANGAMATH, a senior beamline scientist, checks a novel-coronavirus protein crystal sample at the UK's Diamond Light Source. Researchers there have focused on the virus's main protease, a protein that is essential to its replication inside the human host cell.

surface and initiates infiltration to human cells.

As of 25 March, 108 structural determinations of SARS-CoV-2 proteins, both alone and with attached compounds, had been deposited in the open-access Worldwide Protein Data Bank (PDB). At that time, 77 structures of M^{pro}, with various ligands, had been submitted by teams working at the Diamond Light Source. More structures were expected to be released in mid-April.

Several scientists caution that some of the deposited structures are not very well defined. "Fast and automated does not always mean good quality," says Andrzej Joachimiak, who heads a crystallography group at the APS. Crystallographers traditionally deposit structures in the PDB before their research has been refereed, notes John Helliwell, a retired University of Manchester biophysicist, chemist, and crystallographer.

For structural biologists, x-ray crystallography is by far the most commonly used tool for unraveling protein structures. It also produces the highest-resolution images of structures. The process typically is performed at cryogenic temperatures to limit ionizing-radiation damage to proteins. (See the article by Bob Glaeser, PHYSICS TODAY, January 2008, page 48.)

Small-angle x-ray scattering has dedicated beamlines at each of the synchrotron light sources. The technique lacks the angular resolution of crystallography, but it can be used to examine macromolecules in solution, nearer a protein's native state at room temperature. The scattering can explore how structures change over time as the virus matures. That evolutionary process could occur over hours or days, says Britt Hedman, SSRL science director.

Main protease

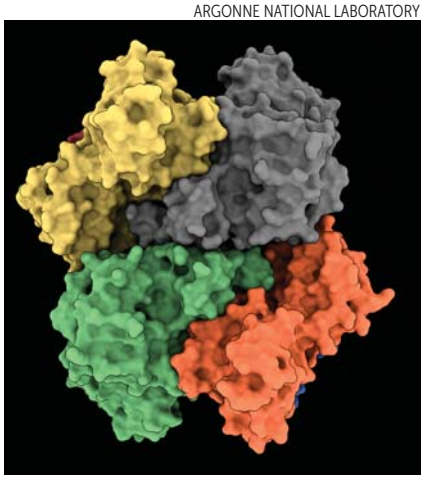
When crystal-growth conditions are known, as they are for ligand-binding analyses, x-ray crystallography is the fastest method to determine 3D structures and can produce diffraction data in fractions of a second. One of the most important SARS-CoV-2 structures solved by crystallography is M^{pro}.

When the viral RNA enters the human cell, it hijacks the host's protein factories—the ribosomes—to make two long polyproteins that contain the components needed for viral replication. Among them are two proteases, which cut the polyproteins into individual proteins. M^{pro} makes most of the cuts. It's 96% identical to that of the SARS coronavirus of 2003, says Andrey Kovalevsky, a senior scientist at Oak Ridge National Laboratory, and it's "absolutely essential for the virus to reproduce."

The APS has 16 beamlines devoted to protein crystallography, though not all are currently in use. With other light sources trained on M^{pro}, teams at the APS turned their focus to some of the virus's other nonstructural proteins (NSPs). A group led by Joachimiak at the APS announced the decoding of the Nsp15 structure on 2 March. The functions of many NSPs are not as clearly understood as those of proteases, but Nsp15, which is 89% identical to its counterpart in the SARS 2003 virus, may increase SARS-CoV-2 virulence by interfering with the body's immune response, says Joachimiak, whose group also has determined structures of four other NSPs. Research on the SARS 2003 Nsp15 protein showed that its inhibition could slow viral replication, although no drugs have sprung from those findings.

Groups from Northwestern and Purdue Universities, Scripps Research Institute, and Walter Reed Army Institute of

ARGONNE NATIONAL LABORATORY



RESEARCHERS AT ARGONNE NATIONAL LABORATORY'S Advanced Photon Source solved the structure of a severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) protein known as Nsp15, an endoribonuclease. It consists of six identical protein chains, four of which are visible here. Nsp15 cleaves specific regions of RNA and is believed to interfere with human antiviral defenses. From preliminary results, other research groups have identified compounds that may deactivate the protein by disrupting its assembly.

DIAMOND LIGHT SOURCE

AN AERIAL VIEW OF THE DIAMOND LIGHT SOURCE near Oxford, UK. Researchers have used the synchrotron to determine dozens of three-dimensional structures of the main coronavirus protease in combination with various ligands that may inhibit the protein's function. That information could be used in the design of new antiviral drugs.

Research are using the APS to determine structures of three other NSP proteins and the spike protein. As of 1 April, scientists using the APS had found structures of six proteins. They also found seven more structures of those proteins in combination with potential antiviral compounds or antibodies. The lab issued a notice of availability on 25 March for SARS-CoV-2 research proposals to its 15 000-member user community. Since then, 25 groups have either collected data or requested beamtime to do so, says Fischetti.

The Shanghai synchrotron's determination of the M^{pro} structure was published in the PDB on 4 February. A second determination, made at BESSY II by a University of Lübeck team led by Rolf Hilgenfeld, was posted to the bioRxiv preprint server on 20 February. A third structure followed a week later from Diamond crystallographers. Each replicated and refined earlier work. Several teams are examining the raw diffraction data that Diamond crystallographers have made available via Zenodo, a CERN-operated data repository. "Every group in the world that has the means to look at this protein will be trying to do that," says David Owen, a structural biology postdoc at Diamond.

From 0 to 1000 in 10 days

Diamond was the first to take the next step, the rapid screening of compounds for binding with M^{pro}. In a process known as fragment screening, crystals are paired in solution with hundreds or thousands of small organic compounds with a molecular weight of 200–300 daltons. The combined structures are then examined in a high-throughput process that screens for insightful interactions. "In 10 days we went from having nothing at all to over 1000 crystal structures, and over 60 molecules were found to bind to the protein so far," says Owen.

As do most light sources, Diamond offers an automated crystallography system that allows experimenters to work and obtain their data from home. An operator need only load the crystals once a day. Diamond can produce hundreds of crystal–compound combinations in a day and expose them at a rate of 30 per hour.

Through screening, "you can build a picture of what portions of a drug are going to work because you look at fragments of the compound and see whether individually the fragments bind to the protein binding site," says Helliwell. "It's proved effective to build from fragments upward to find new drugs. The synthetic chemist can take the information and string [fragments] together by chemical organic synthesis."

Each major light source hosts industrial users. Pharmaceutical and biotech companies perform proprietary research, but also benefit from openly available results from academic users. "We have provided drug companies with a lot of structural information," says Owen. "We've been able to put out the structure of [M^{pro}] to a very high resolution, better than 1.3 Å. That provides really precise information about the location of each atom in the structure and is very useful to people designing drugs." Among other things, chemists will be looking at the strength and tightness of the binding attraction between the protease and the molecules they themselves have built.

On 31 March, Diamond announced an initiative with Exscientia, a UK artificial-intelligence-driven drug discovery company, and Calibr, the drug-development arm of Scripps Research, to screen existing, clinically approved drugs against several SARS-CoV-2 protein targets. If inhibiting action against the pathogen is discovered, approved drugs could accel-

erate the path to a treatment, although human trials will still be needed.

Exscientia will have access to Scripps's collection of 15 000 approved and human-tested drugs, Martin Redhead, head of quantitative pharmacology at Exscientia, said in a statement. The protein targets to be screened include M^{pro} and the spike protein.

Other pharmaceutical companies are working with Diamond, but a spokesperson declined to name them.

On 2 April, BESSY II began fragment screening against M^{pro}, and scientists were analyzing the data at press time. "We start with 1.5 million commercially available compounds that you can buy in quantities of at least 1 milligram," says Manfred Weiss, director of the macromolecular research group at BESSY II. "We filter the compounds according to the size, solubility, and some rules that medicinal chemists have." Some 1200 compounds met those criteria. A typical fragment screening produces 10 to 12 compounds that bind and can serve as starting points for drug development, Weiss says. He adds that the terminology is misleading because the molecules aren't actually fragments of anything.

Industry has begun fragment screening, says Lisa Keefe, executive director of IMCA-CAT, an industry-funded nonprofit association that, unlike academic users, pays to operate beamlines at the APS. Formed in 1992, IMCA-CAT restricts membership to industrial users, including Merck, Pfizer, Bristol-Myers Squibb, Novartis, and Abbvie. Nonmember companies can arrange to use the IMCA-CAT beamlines. Firms can screen their large proprietary collections of fragments and don't have to make the results public. Keefe won't say which companies are screening or what proteins they are targeting. But she notes that in the current crisis "there has been some reaching out between members" to compare findings.

Weiss laments the inability of most academic researchers to gain access to industry's compound libraries. But Aled Edwards, director of the Structural Genomics Consortium (SGC), a public–private international nonprofit organization, says the differences between public and proprietary collections of compounds are not that great. "There's no evidence to suggest that one is massively better than the other," he says.

Electrons and neutrons

Apart from crystallography, several other techniques are used for protein structure determination. For molecules that resist crystallization, such as large protein complexes, researchers can turn to cryoelectron microscopy (cryo-EM). (See PHYSICS TODAY, December 2017, page 22.) A team led by Jason McLellan of the University of Texas at Austin used the technique to determine the structure of the spike protein; the results were published in *Science* on 13 March. The spike protein is considered a prime candidate for a vaccine target. However, because the base of the protein is anchored to the hydrophobic viral membrane while the rest of the protein is hydrophilic, the full-length spike protein is hard to crystallize.

Stanford University has kept one of its six cryo-EM instruments at SLAC open for SARS-CoV-2 research. Diamond has five cryo-EM instruments available, one of which is reserved for industrial use.

With the cryo-EM technique, as in x-ray crystallography, myriad individual molecules contribute to a structural determination. Whereas a crystal's molecules are identically arrayed, the molecules embedded in vitreous ice in cryo-EM are randomly oriented, and their fuzzy, individual 2D projection images are assembled computationally into a single, clear 3D image.

In favorable conditions, where molecules are well-preserved on an electron microscope support grid, it takes about 10 hours to collect enough data to generate a structure with atomic detail, says Wah Chiu, who heads SSRL's cryo-EM and bioimaging division. More difficult projects could require 48 hours or longer. Cryo-EM can be used to image a thin region of vitrified cells and generate a 3D tomogram from which maturing virus particles can be identified.

In addition to studying noncrystallizing molecules, cryo-EM, in combination with advanced image processing, can obtain from a single sample multiple conformations of proteins. That approach offers insight into the dynamical properties of protein molecules under different biochemical and functional conditions. An emerging use of cryo-EM is to uncover the overall architecture of cellular structures *in situ* under normal and pathological conditions, Chiu says.

Neutron crystallography allows investigators to elucidate the positions of

hydrogen atoms, which are invisible to crystallography, in 3D protein structures. That advantage is important because the strongest binding sites on a protein involve hydrogen bonds. In x-ray crystallography, the photons scatter off the electron charge clouds of the protein molecule's constituent atoms. As the lightest element, hydrogen scatters x rays extremely weakly. So x-ray crystallographers often must infer the location of hydrogen atoms in the structure, especially in the ionizable amino acids that are often involved in enzyme reactions. In contrast, neutrons interact with atomic nuclei to provide a hydrogen signature comparable to those of the protein's nitrogen, oxygen, and carbon atoms.

Matthew Blakeley, a chemist and crystallographer at France's Institut Laue-Langevin (ILL) neutron source, notes that x rays produce structurally damaging free radicals when they pass through crystals. To limit the radicals' diffusion, x-ray data collections are typically performed at 100 K. Neutrons don't damage protein samples, so crystallography with neutrons can be performed at room temperature. Neutron beam intensity, however, is several orders of magnitude less than is achievable with x rays, and structural data gathering can take 7 to 10 days, says Kovalevsky. That makes neutron techniques less suited for fragment screening. It also requires much larger crystals, which are difficult to grow.

Oak Ridge National Laboratory reopened its Spallation Neutron Source on 7 April and gave priority to SARS-CoV-2 research. The lab's other neutron source, the High Flux Isotope Reactor, was scheduled to resume operations on 27 April, and coronavirus research is to be prioritized on two beamlines there. At the outset of the pandemic, both facilities had been shut down for routine scheduled maintenance. The ILL is closed, and no reopening date has been set.

Although x-ray free-electron lasers (XFELs) could be highly useful for SARS-CoV-2 research, both the Linac Coherent Light Source at SLAC and the European XFEL in Germany were offline as of early April. The LCLS has been shut down for an upgrade, but a SLAC spokesperson says a restart could occur soon after California's shelter-in-place order is lifted. European XFEL group leader Bernd Ebeling said in an email that beam time will be devoted to coronavirus research once operation becomes possible. It was unclear at press time whether Japan's SACLA, the planet's other XFEL, was conducting coronavirus research.

XFELs offer two advantages over synchrotrons, says Helliwell: They can determine structures at room temperature, and their femtosecond pulse length and high intensity allow the collection of diffraction data before radiation damages the protein. Radiation damage can change the protein structure and thus affect its interpretation.

Because of those advantages, XFELs can obtain structural information from submicrometer crystals. Using the LCLS, a team led by Lars Redecke of the University of Hamburg in 2012 reported the first new biological structure solved with an XFEL.

Follow-through needed

Despite the frantic pace of research, many scientists are skeptical that drugs or vaccines to fight the virus will come in time to ease the current pandemic. Human trials could be a year or more away, even for approved drugs or antibodies found to provide some action against the coronavirus; the timeline for newly identified medicinal compounds is even longer. "Neither synchrotrons nor cryo-EM will likely make a big impact on this pandemic," says Edwards of the SGC, "but both will be really powerful tools in getting science ready to stop the next one."

Stephen Streiffer, APS director, disagrees; he argues that any existing compound approved for human use that's found to be therapeutically useful would be moved into clinical use immediately. He notes that current public health strategy is aimed at slowing the spread of infection until a vaccine can be deployed.

Joachimiak laments the lack of follow-through on drug development efforts from SARS and MERS (Middle East Respiratory Syndrome, which appeared in 2012). "NIH spent $700 million in the past 20 years to study SARS and MERS. And we have no antiviral, no antibodies, no vaccine for SARS and MERS," he says. Drugs developed in response to the earlier outbreaks might have worked against the current pathogen, he notes.

David Kramer