

# Workshop on Biological Science Opportunities Provided by the APS Upgrade



Argonne National Laboratory,  
August 20-21, 2018

### **About Argonne National Laboratory**

Argonne is a U.S. Department of Energy laboratory managed by UChicago Argonne, LLC under contract DE-AC02-06CH11357. The Laboratory's main facility is outside Chicago, at 9700 South Cass Avenue, Argonne, Illinois 60439. For information about Argonne and its pioneering science and technology programs, see [www.anl.gov](http://www.anl.gov).

### **DOCUMENT AVAILABILITY**

**Online Access:** U.S. Department of Energy (DOE) reports produced after 1991 and a growing number of pre-1991 documents are available free at OSTI.GOV (<http://www.osti.gov/>), a service of the U.S. Dept. of Energy's Office of Scientific and Technical Information

### **Reports not in digital format may be purchased by the public from the National Technical Information Service (NTIS):**

U.S. Department of Commerce  
National Technical Information Service  
5301 Shawnee Rd  
Alexandria, VA 22312  
**[www.ntis.gov](http://www.ntis.gov)**  
Phone: (800) 553-NTIS (6847) or (703)  
605-6000 Fax: (703) 605-6900  
Email: [orders@ntis.gov](mailto:orders@ntis.gov)

### **Reports not in digital format are available to DOE and DOE contractors from the Office of Scientific and Technical Information (OSTI):**

U.S. Department of Energy  
Office of Scientific and Technical Information  
P.O. Box 62  
Oak Ridge, TN 37831-0062  
**[www.osti.gov](http://www.osti.gov)**  
Phone: (865) 576-8401  
Fax: (865) 576-5728  
Email: [reports@osti.gov](mailto:reports@osti.gov)

### **Disclaimer**

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor UChicago Argonne, LLC, nor any of their employees or officers, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of document authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof, Argonne National Laboratory, or UChicago Argonne, LLC.

# Workshop on Biological Science Opportunities Provided by the APS Upgrade

Argonne National Laboratory,  
August 20-21, 2018

## Organization Committee

Erin J. Adams, *U. Chicago*  
Paul D. Adams, *LBNL*  
Brian H. Davison, *ORNL*  
Robert F. Fischetti, *APS/ANL*

Andrzej Joachimiak, *ANL*  
Philippe Noirod, *ANL*  
Todd O. Yeates, *UCLA*

## Invited Participants

Crysten Blaby-Haas, *BNL*  
Ian Blaby, *BNL*  
Sean Crosson, *U. Chicago*  
Shi-You Ding, *MSU*  
Carrie Eckert, *UC Boulder*  
Mike Himmel, *NREL*  
James Holton, *ALS/LBNL*  
Udaya Kalluri, *ORNL*  
Eaton E. Lattman, *Hauptman-Woodward MRC*  
Wei Liu, *ASU*  
Artem Lyubimov, *SSRL*  
Michelle O'Malley, *UCSB*  
Engin Özkan, *U. Chicago*

Heather Pinkett, *Northwestern U.*  
Rama Ranganathan, *U. Chicago*  
Jose Rodriguez, *UCLA*  
Holger Sondermann, *Cornell U.*  
Timothy A. Springer, *Harvard Med. School*  
Stephen Streiffer, *APS/ANL*  
Eric Sundberg, *U Maryland*  
Michael Thompson, *UCSF*  
Michael Wall, *LANL*  
Gayle Woloschak, *Northwestern U.*  
Junko Yano, *LBNL*

## Table of Contents

A. Executive Summary .....	1
B. Current State of the Art.....	3
C. The Advanced Photon Source Upgrade.....	4
D. Addressing Scientific Grand Challenges .....	5
Dynamics and Complex Conformational States of Macromolecules .....	6
Protein and Complexes in Their Natural Operating Environment .....	7
Membrane and membrane-associated proteins and complexes .....	7
Higher-order macromolecular complexes .....	9
Extracellular enzymes and complexes .....	9
Investigating the Sequence-Structure-Function Continuum.....	11
E. Newly Emerging Technological Developments and New Tools for Extracting Valuable Information on Structural Dynamics. ....	12
Rapid Perturbation of Proteins Coupled to Rapid Time-Resolved Readouts of X-ray Diffraction Patterns .....	12
Dynamic Understanding of Diffuse Scattering .....	12
Exploring Conformational Landscapes Under Physiological Conditions .....	13
F. Technical Developments Needed to Realize the Scientific Impact .....	13
Accelerating Structure Determination of Proteins and Macromolecular Complexes.....	13
Controlling Radiation Damage .....	14
Computational Approaches to Big-Data Collection.....	15
Improving Protein Production and Crystallization .....	15
Antibodies as Crystallization Chaperones .....	17
G. Synergistic Structural and Imaging Techniques.....	17
Cryo-Electron Microscopy .....	17
X-ray Imaging Techniques .....	18
Small-Angle X-ray Scattering and Small-Angle Neutron Scattering .....	19
Neutron Imaging Techniques .....	19
Molecular Dynamic Simulations.....	20
H. Partnership with the User Community.....	20
Sample Delivery, Automation of Beamlines, and User Interface Standardization .....	20
User Training .....	21
I. References .....	21

## A. Executive Summary

Synchrotron-based macromolecular X-ray crystallography (MX) is the primary method for determining the structure of biological molecules including proteins, nucleic acids, complexes, and large assemblies at high resolution (e.g., 95% of all new Protein Data Bank [PDB] deposits in 2017). The Advanced Photon Source (APS), which is the largest and the most productive light source for MX, will be upgraded to a multi-bend achromat lattice (the Advanced Photon Source Upgrade — APS-U) that will deliver the ultimate diffraction-limited X-ray source with radically transformative MX capabilities. The specific advantages of the APS-U X-ray source include (1) increased brightness allowing delivery of intense beams with dimensions of 1- $\mu\text{m}$  or less, resulting in a 100-fold improvement in signal-to-noise, (2) increased brightness at high energy (especially for energies of 12 keV-40 keV commonly used in MX) that can be exploited to reduce primary radiation damage, and (3) low-convergence X-ray beams that will be essential for the study of large macromolecular complexes at high resolution. A workshop on “Biological Science Opportunities Provided by the APS Upgrade” was organized at Argonne National Laboratory on August 18-20, 2018, to discuss the biological science opportunities offered by the unique new capabilities of the APS-U and how the broad structural biology community could best exploit them.

The APS-U takes place in a context where advances in X-ray free-electron laser (XFEL) sources and cryo-electron microscopy (cryo-EM) have added to the constellation of techniques and propelled structural biology in new directions. Synergies exist between these different technologies and it is essential to develop enabling technologies that exploit the unique, world-leading, and synergistic characteristics of each light source to boost capabilities for producing leading-edge biological research for U.S. Department of Energy (DOE) missions in energy and environment, and National Institutes of Health (NIH) missions in human health.

Workshop attendees identified several scientific challenges requiring novel high-throughput capabilities that can bridge the gap in our understanding of the relations between sequence, structure, dynamics, function, and phenotype. Three-dimensional imaging of macromolecules with X-rays provides a crucial component for enabling a predictive understanding of sequence-to-function, knowledge that is essential for successfully engineering protein targets and pathways of direct relevance to DOE and NIH missions. Workshop attendees expressed confidence that the APS-U will provide unique capabilities to address the following challenges:

*Dynamics and Complex Conformational States of Macromolecules.* One of the greatest enduring challenges in structural biology is how to extract the most information about the dynamics and complex conformational states of macromolecules, including proteins, nucleic acids, and their complexes. We need to track protein conformational changes in real time by simultaneously recording structural and kinetic data that will allow us to understand correlations between molecular motions and chemistry. Examples include understanding vibrational and other modes of molecular motion in catalysis, how enzymes function in divergent environments, the effects of post-translational modifications, and details of molecular interactions. The low emittance of the APS-U source will be essential for providing the small, high-intensity X-ray beams needed to study the conformational dynamics of macromolecules across varied length and time scales.

*Protein and Complexes in Their Natural Operating Environment.* To perform their function, proteins and protein machines operate in environments such as the cytosol, lipidic membranes, or the extracellular milieu. Exemplars include membrane-associated protein machines such as transporters; cell surface receptors and photosystems; higher-order macromolecular complexes formed from repeating oligomeric proteins such as bacterial microcompartments and flagella; and extracellular enzymes and machines that are involved either in biomass deconstruction such as fungal glycoside hydrolases and cellulosomes or in the biosynthesis of the cell wall, i.e., the plant cellulose synthase complex. An understanding of the dynamical functioning of these systems

requires the capability to characterize the different chemical and conformational states during operation of these machines, the movements of the protein domains associated with these states, the presence of protein domains intrinsically devoid of ordered structure in the absence of a partner, and the impact of cellular components (carbohydrates, lipids, small metabolites, and metal ions) on protein structure. The low emittance of the APS-U source will be essential for providing the low-convergence, high-intensity X-ray beams required to study large protein machines and cellular components such as the nuclear pore complex.

*Sequence-Structure-Function Continuum.* Predicting the level of activity of an enzyme variant under a particular reaction condition from its primary amino acid sequence is a long-standing challenge. One major hurdle is the prohibitive cost of generating structural information for a large number of enzyme variants displaying a wide range of activity levels. We need a dramatic acceleration of structure determination concurrent with a steep reduction in cost. Ultrafast serial crystallography (SX) can produce such an acceleration by leveraging the APS-U smaller and brighter X-ray beams in combination with advanced modes of sample delivery (i.e., microcrystals), high-performance computational pipelines, and advances in protein engineering and synthetic biology. Faster and cheaper structure determination will also open new avenues to assess the functional significance of protein variants resulting from adaptive mutations or alternative mRNA splicing by allowing one to associate structural changes with functional and kinetic changes. In addition, the rapid determination of a large number of protein structures will lead to a reevaluation of how structural biology can help to make sense of the overwhelming amount of “omics” data that are being generated in order to better understand protein functions and interactions.

Workshop participants also identified a synergistic challenge, which is the need to provide strategies for integrating atomic structures with imaging technologies that together span the breadth of biological scales. The advances in MX promoted by the APS-U need to be complementary with advanced imaging techniques using X-rays, visible light, electrons, and neutrons to gain mechanistic understanding of the interactions of proteins and complexes with their operating environment. A sister workshop, “Opportunities in Biological and Environmental Research Uniquely Enabled by the APS Upgrade,” focused on APS-U X-ray imaging opportunities in biological and environmental research was also organized at Argonne in August 2018.

The technological developments and new tools needed for extracting valuable information on structural dynamics were also discussed at the biological science workshop. New data collection strategies, particularly SX and other low-dose, high-redundancy methods, will make ambient temperature (rather than cryogenic) data collection routine, providing access to structural information under more biologically relevant conditions. Rapid perturbation of proteins using methods such as electric field pulses and laser-induced temperature jump will excite intrinsic macromolecular structural dynamics and reveal the accessible excited-state structures of a protein or nucleic acid, rather than a single ground-state structure. The capture of these excited-state structures will be made possible by the high-flux, sub-nanosecond-scale X-ray pulses available at the APS-U. The short X-ray exposures also will allow the exploration of protein conformational landscapes under time scales relevant to physiological conditions, as it will become possible to conduct time-resolved serial microcrystallography, where reactions are initiated within protein crystals and protein structures are determined as a function of time. Finally, the brighter beams provided by the APS-U will result in increased observation of diffuse non-Bragg scatter, which arises from correlated motions across wide-ranging length scales in crystal specimens. It is generally accepted that diffuse scattering holds much potential information about protein dynamics, but obtaining decisive dynamic interpretations remains a frontier problem that will require investigation and provide synergy with the computational advances required to meet high-data demands.

For the technological advances discussed above, software developments will continue to play an important role in addressing the needs for new algorithms and new approaches to data processing, analysis, and presentation. We anticipate far greater data throughputs and volumes that will enable new data collection/analysis approaches. For example, there will be a need for real-time feedback into how diffraction data are collected to achieve optimal results based on inputs from assessment of radiation damage, rapid data processing, structure solution, and analysis of conformational state dynamics. The development of new algorithms and new software platforms will be critical to keep pace with next-generation synchrotron beams and fully benefit from them.

Participants in the workshop felt strongly that X-ray beam properties will position the APS-U to take a leading role in synchrotron SX applications. To realize the scientific impact, workshop participants recommended research in several key areas: Research in crystallization that seeks innovative approaches to monitor, optimize, and accelerate crystallization is critically needed. An in-depth understanding of radiation damage, which results from interactions of high-brightness X-ray beams with biological samples, is essential for modeling damage-free states. The high brightness at high energy of the APS-U will be key to studying and minimizing radiation damage. Innovative approaches to minimize radiation damage (e.g., “painting crystals with X-rays” by producing virtual beams to match any sample size through optimization of X-ray beam size, energy, and dose) are also needed.

For the research community to benefit from these improvements, it will be necessary to make measurements on small, micron-size crystals. The production, optimization, detection, and harvesting of micro- and nanocrystals is not routine. Experimental facilities will be needed at light sources to grow these crystals routinely, coupled with X-ray-based methods to detect, image, and deliver very small crystals, including technology to enable shipment (at cryogenic and room temperatures) of crystals in their native growth media or in trays or chips for direct mounting at beamlines. The rapid delivery of such crystals to the X-ray beam represents another challenge. All of these new experimental capabilities will need to be accompanied by improved data analysis tools that are able to recover signal from weak data of micron-size crystals, and extract information about dynamics in biological systems.

Beyond the theoretical and technical developments, workshop participants recommend that protein purification and crystallization services be supported as a part of the user program accessible at the APS-U. The development of robust and scalable techniques for the expression of prokaryotic and eukaryotic proteins and complexes is needed. Antibodies could be used as chaperones to facilitate crystallization of complex, multiple domain, disulfide-linked, glycosylated, and other recalcitrant protein targets of interest to DOE or NIH. As the production of crystals is still a major bottleneck, best practices for crystal growth, harvesting, data collection, and processing need to be determined to help establish a strong user base. From the workshop, it is clear that researchers will need to be extensively trained to make use of the new experimental resources that will become available to the biology community. Furthermore, improvements should be made to automate collection of data, including sampling of a large number of crystals to find those that diffract the best, and offering rapid access. Users now screen crystals directly at the synchrotron rather than on home sources. Several recommendations are discussed in the report.

## **B. Current State of the Art**

Macromolecular X-ray crystallography (MX) is a powerful and most productive method to visualize biological macromolecules at the atomic level. Over the past 60 years, MX has been used to obtain the three-dimensional structures of individual proteins and entire functional assemblies captured in different states. Today, synchrotron-based MX is the primary method for determining

the structure of biological molecules including proteins, nucleic acids, complexes, and large assemblies at high resolution. Last year, 91% of all new Protein Data Bank (PDB) deposits were solved using X-rays, and over 96% of these were determined using synchrotron radiation. The important role for structural biology in U.S. Department of Energy (DOE) science was emphasized in a recent document - the “Grand Challenges for Biological and Environmental Research: Progress and Future Vision”<sup>1</sup>, which was prepared by the DOE Biological and Environmental Research (BER) Advisory Committee. MX serves as the foundation of biology with >129,000 structures deposited in the PDB (<https://www.rcsb.org>) and it is essential for biomedical and basic research, drug discovery, and biotechnology. Light sources have provided many stimuli in advancing MX, and the powerful synchrotron X-rays are far superior to traditional home sources for single-crystal and solution-diffraction experiments as well as for many other imaging techniques. Synchrotron light sources have broadly expanded the range for many flux- and brightness-limited experiments, and they have introduced a whole new arena of X-ray studies that provide new structural information.

Currently, there are more than 130 beamlines available for MX experiments worldwide (<http://biosync.rcsb.org>). In the U.S., several of these resources for biological research are supported by DOE/BER (<https://www.berstructuralbioportal.org/>). In recent years, X-ray free-electron laser (XFEL) sources have driven the development of new techniques and technology such as femtosecond serial crystallography (SX). A synergy exists between the storage ring and XFEL-based sources, whereby technology developed at one is implemented at the other. A wide variety of SX techniques are now available on many beamlines worldwide. Recently reported megahertz serial crystallography at the European X-Ray Free-Electron Laser Facility<sup>5,6</sup> shows impressive progress in expanding the experimental repertoire. Advances with cryo-electron microscopy (cryo-EM) have added to the constellation of techniques that can propel structural biology in directions and opportunities not anticipated previously<sup>7,8</sup>. It is essential to continue the development of enabling technologies through the construction of state-of-the-art instruments that exploit the unique, world-leading, and synergistic characteristics of each of the light sources. These instruments will boost capabilities for conducting basic biological research, thereby producing leading-edge science for DOE missions in energy and environment, and National Institutes of Health (NIH) missions in human health. When combined with future XFELs, cryo-EM, and computational resources, the Advanced Photon Source Upgrade (APS-U) will have a tremendous transformational impact on the future of structural biology and its discoveries by enabling multimodal imaging capabilities over length scales from sub-nanometers to millimeters.

### **C. The Advanced Photon Source Upgrade**

In the U.S., the APS is the largest and the most productive light source for MX. Upgrading the APS to a multi-bend achromat lattice will deliver the ultimate diffraction-limited X-ray source with radically transformative MX capabilities. Components of the APS-U are currently being designed and fabricated, and first X-rays are planned to be delivered in the middle of 2023. Compared to the present-day APS, the specific advantages of the APS-U X-ray source include (1) increased brightness that will allow beamlines to deliver intense beams with dimensions of 1-  $\mu\text{m}$  or less, resulting in a 100-fold improvement in signal-to-noise; (2) increased brightness at high energy (especially for energies of 12 keV-40 keV commonly used in MX) that can be exploited to reduce primary radiation damage due to the escape of photoelectrons that cause most of the damage; (3) low-convergence X-ray beams that will be essential for the study of large macromolecular complexes at high resolution; and (4) high-speed detectors that will allow a reduction of secondary radiation damage at room temperature by using very short exposures to potentially “out-run” free-

---

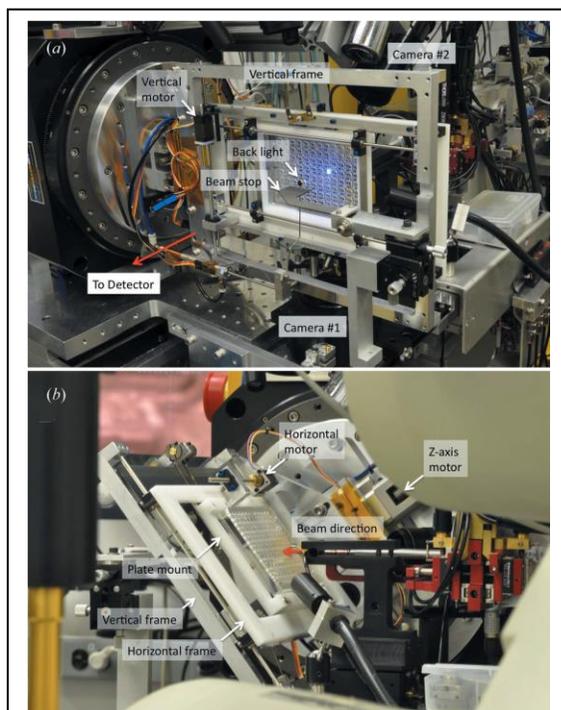
<sup>1</sup> <https://science.energy.gov/~media/ber/berac/pdf/Reports/BERAC-2017-Grand-Challenges-Report.pdf>

radical diffusion and chemistry. The source will enable new science in structural biology and also be ideal to exploit macromolecular systems *in situ* (Fig. 1) and *in cellulo*.

A workshop on “Biological Science Opportunities Provided by the APS Upgrade” was held at Argonne National Laboratory on August 18-20, 2018, sponsored by the APS-U and Computing, Environment, Life Sciences (CELS) directorates. The work-shop focused on the unique capabilities of the APS-U and how the broad structural biology community could best exploit these capabilities. Paramount was what important science challenges could be addressed, what new technologies and methods would be needed to maximize the impact of the APS-U on biological research, and how future MX methodology could leverage the synergistic advances in XFELs, cryo-EM, and other emerging technologies. Discussions also focused on whether these future structural biology resources could be used to address grand challenges in biology for predictive understanding of the relationship between the genome, structure, function, and interactions important for research relevant to DOE energy, biological, and environmental missions, and NIH human health and well-being missions.

#### D. Addressing Scientific Grand Challenges

To achieve its missions in sustainable bioenergy, systems biology, and biosystems design applications, DOE-BER science must go beyond the accumulation of genomic information to understand biological functions and engineer them. This endeavor will require novel high-throughput capabilities that can bridge the gap in our understanding of the relations between sequence, structure, dynamics, function, and phenotype. Three-dimensional imaging of macromolecules with X-rays, in particular, provides a crucial component for enabling a predictive understanding of sequence-to-function, knowledge that is essential for successfully engineering protein targets and pathways of direct relevance to DOE and NIH missions. The systems biology approaches that underpin much of BER science will be able to exploit the large number of samples and conditions that can be analyzed with smaller, brighter beams. DOE Basic Energy Sciences (BES) research needs to combine high-brightness beams with other X-ray-based measurement methods like absorption and emission spectroscopy, X-ray transmission, fluorescence microtomography, and other spectroscopic techniques, in order to simultaneously observe conformational and chemical states of catalysts and the degree of sample damage. The ability to probe the time domain of biological processes from microseconds to seconds will be essential and will be enabled by crystal perturbation techniques coupled with new sample delivery methods. NIH science will also benefit from high-throughput screening methods at synchrotrons together with the high resolution typically achieved with X-ray crystallography to enable structure-based ligand/ drug discovery and optimization.



**Fig. 1. *In situ* data collection.** A 96-well plate scanner for *in situ* data collection at sector 19 ID beamline at the APS. (a) A back view of the plate scanner mounted on the beamline goniometer at the rotation angle of  $\omega = 0^\circ$ . (b) A side front view of the plate scanner at its maximum tilt angle,  $\omega = 40^\circ$ . The total rotation range established was  $45^\circ$ , from  $-5^\circ$  to  $+40^\circ$  <sup>2</sup>.

The future challenges with X-rays will include how to maximize the capture of biologically relevant information about protein structure and dynamics. We need deeper insights into protein energetic/conformational landscapes and how they are related to function. Addressing this challenge will require the development of technical strategies that allow structural characterization under physiological conditions, that minimize radiation damage, and that capture protein conformational and chemical dynamics. Examples include snapshots of catalytic cycles, the effects of post-translational modifications, and details of molecular interactions. A synergistic challenge is to provide strategies for integrating atomic structures with imaging technologies that together span the breadth of biological scales.

Complexity, variability, and degeneracy are inherent to biological systems, confounding their detailed characterization. As has already been observed at XFEL sources, small, high-brightness X-ray beams at synchrotrons will enable rapid data collection and analysis of large numbers of challenging samples and new data collection modalities that can study systems currently intractable at third generation sources, leading to greater insight into complexity and variability. The ability to make measurements from many samples will also enable researchers to probe the responses of biological systems to different physiological conditions and perturbations. These developments are likely to be transformative in our ability to make measurements under more physiological conditions, provide strategies for reducing radiation damage — a major problem for current X-ray sources — and correct for thermal and or mechanical drift, a potential issue at the APS-U.

Ambient-temperature crystallography has (re-)emerged as a powerful tool to study protein dynamics<sup>9-11</sup>. We anticipate that new analytical tools and crystal manipulations will enable extraction of dynamic information from crystal structures. Beyond the crystalline environment, the ability to study biological systems in solution using scattering methods, coupled with high-brightness beams, will potentially enable studies of weakly interacting proteins and rare events. Challenges still remain in using structural methods to study complex materials, such as the plant cell wall, where other scattering or imaging methods may provide greater insight.

Below, we define several broad categories of scientific challenges discussed during the workshop.

### ***Dynamics and Complex Conformational States of Macromolecules***

One of the greatest enduring challenges in structural biology is how to extract the most information about the dynamics and complex conformational states of macromolecules, including proteins, nucleic acids, and their complexes. The traditional view of X-ray crystallography is mainly static, while the prevalence and importance of dynamics in macromolecules is universally appreciated. However, to understand biological processes we must track protein conformational changes in real time by simultaneous recording of structural and kinetic data that allow us to understand correlations between molecular motion and chemistry.

To frame the challenges in understanding protein dynamics, it is worth noting that proteins are “soft” materials with many degrees of freedom. A general hypothesis is that, embedded within this very high-dimensional space of dynamical motions are a much lower dimensional set of biologically relevant motions that define the evolutionarily selected reaction coordinate. These relevant motions could involve both local and distributed collective modes operating on time-scales from picosecond to many seconds, and because of the steep distance and geometry dependence of the fundamental forces, could comprise subtle sub-angstrom motions. Thus, we require experimental methods that can deliver both high spatial resolution and broad temporal resolution, and we require new data analytic methods to extract and model the relevant motions.

Understanding the dynamic behavior of key biological systems will advance the missions of diverse research agencies. In addition to the importance of dynamics as a fundamental concern,

a number of practically important phenomena are affected by movements and conformational transitions in macromolecules. Catalysis is an important component of DOE-related research, and understanding vibrational and other modes of molecular motion are critical. How enzymes function in divergent environments (e.g., high or low temperature, pH, salinity extremes) is a topic of potentially high interest for DOE. Regulation and systems control are likewise subjects of DOE interest; those phenomena rely extensively on dynamic properties of macromolecules. Static structural snapshots leave unanswered important questions in many systems, and so working to advance our understanding of protein dynamics remains critically important.

The workshop participants expressed confidence that the APS-U will provide unique capabilities to address these challenges.

## ***Protein and Complexes in Their Natural Operating Environment***

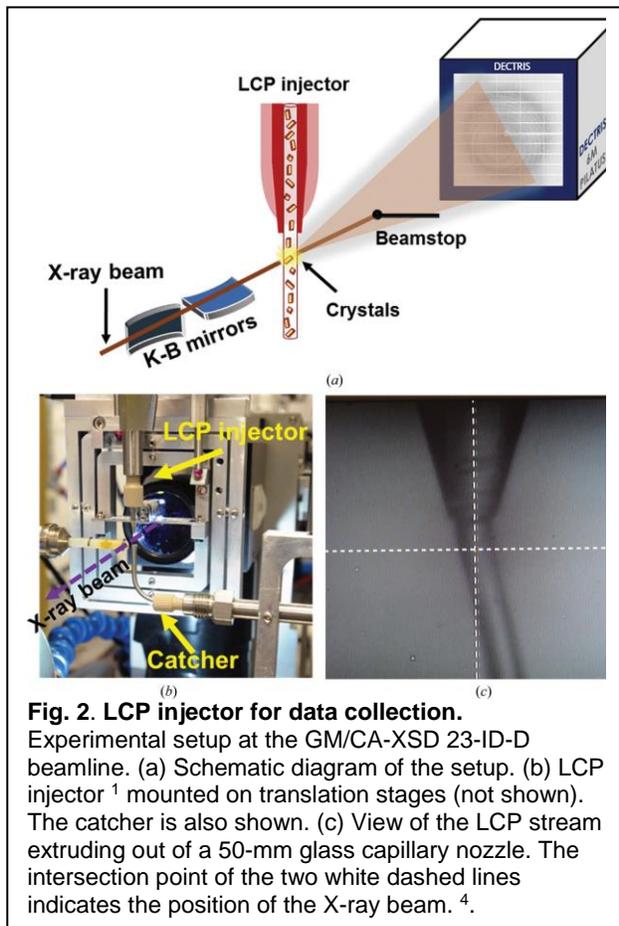
### *Membrane and membrane-associated proteins and complexes*

Membrane proteins (MPs) play important roles in cell survival and cell communication, as they function as transporters, receptors, anchors, and enzymes. Structural biology has advanced our understanding of membrane proteins like no other scientific discipline<sup>12</sup>. Several examples discussed at the workshop illustrate the important role of MX in structure-function studies of membrane proteins and highlight the future research challenges.

In plants, algae, and cyanobacteria, Photosystem II (PSII) catalyzes the light-driven splitting of water at a protein-bound  $\text{Mn}_4\text{CaO}_5$  cluster, the water-oxidizing complex<sup>13</sup>. The mechanistic understanding of water oxidation has advanced significantly over the years by using crystallography and X-ray spectroscopy at synchrotron facilities to study samples at cryogenic temperature, and more recently at XFELs at room temperature under physiological conditions<sup>14</sup>. In addition to understanding the mechanism of the water oxidation reaction, another important aspect of PSII is the assembly mechanism of the  $\text{Mn}_4\text{CaO}_5$  cluster by photo-activation. The assembly is a chaperone-free process from soluble  $\text{Mn}^{2+}$  and  $\text{Ca}^{2+}$  ions that requires light and involves several sequential steps. It harnesses the charge-separation at the reaction center and the coordination environment provided by the amino acid side chains of the protein to oxidize and organize the incoming manganese ions to form the oxo-bridged  $\text{Mn}_4\text{Ca}$  cluster capable of  $\text{H}_2\text{O}$  oxidation. Assembly of the metal cluster can be studied using crystallography and X-ray spectroscopy at synchrotron sources with apo-PSII and the partially assembled metal clusters of PSII by cryo-trapping these intermediate states<sup>15</sup>. Future experiments will require crystallography under physiological conditions to understand the dynamics and the structures of transient or short-lived intermediate states in metal cluster assembly.

In humans, G-protein coupled receptors (GPCRs) constitute the largest family of cell surface receptors that mediate numerous cell signaling pathways and the largest class of targets for therapeutic drugs. Despite their biomedical importance, fewer than ten GPCRs have been crystallized in an active state<sup>16</sup>. GPCRs interact with heterotrimeric G-proteins that are key signaling constituents involved in regulation of many growth and developmental pathways in all eukaryotes. GPCRs are also found in fungi where numerous G-proteins and GPCR receptors have been involved in regulatory pathways controlling nutrient sensing, pheromone response and mating, stress response, and pathogenesis<sup>17</sup>. In plants, G-protein membrane receptors are distinct from animal or fungal GPCRs as they have a single trans-membrane domain (instead of seven), and the G-protein signaling protein networks include novel receptors and effector proteins and display different architectures<sup>18</sup>. G-protein signaling in plants has been shown to control a wide range of developmental and physiological processes, including cell proliferation, root development, leaf and fruit development, chloroplast development, oxidative stress and hypoxia responses, pathogen responses, and hormonal regulation of seed germination. One major

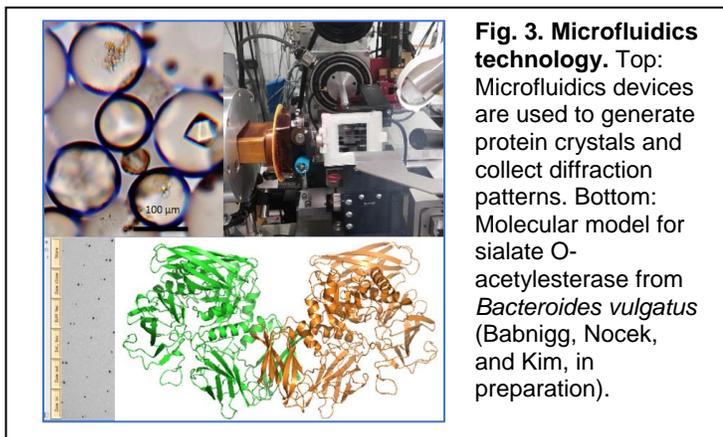
bottleneck in GPCR structural determination is that most GPCR crystals are small (5-10  $\mu\text{m}$ ) and are difficult to optimize to larger sizes suitable for traditional synchrotron radiation diffraction experiments. GPCR structure elucidation re-quires the merging of multiple diffraction data sets that could not be obtained in an efficient, high throughput manner. However, several cur-rent technology developments show promise for overcoming these hurdles, including SX at the



**Fig. 2. LCP injector for data collection.** Experimental setup at the GM/CA-XSD 23-ID-D beamline. (a) Schematic diagram of the setup. (b) LCP injector<sup>1</sup> mounted on translation stages (not shown). The catcher is also shown. (c) View of the LCP stream extruding out of a 50-mm glass capillary nozzle. The intersection point of the two white dashed lines indicates the position of the X-ray beam.<sup>4</sup>

National Institute of General Medical Sciences and National Cancer Institute (GM/CA-XSD) structural biology facility at the APS using lipidic cubic phase (LCP) as a carrier medium (Fig. 2)<sup>4</sup>. Synchrotron SX techniques will enable us to address the issues of crystal size and lack of multiple ligand-GPCR interaction modalities. Additionally, SX using a model GPCR has been performed successfully at the BioCARS facility at the APS in pink-beam Laue diffraction experiments, showing further promise for potential high-throughput crystallography owing to the nature of the Laue diffraction technique. Microfluidics technologies also promise new approaches for delivering small samples to micron-size beams (Fig. 3). Together, these advances coupled with the APS-U will provide the bright micro-beams and techniques necessary to investigate GPCR crystals at high throughput, enabling the study of human GPCR structures in multiple physiological conformations that together define mechanism/function and their complexes with potential drug targets.

Organisms perceive their environment through membrane-associated molecular machines. Transporters (e.g., ABC importers) selectively transport small molecules across the membrane and sensors perceive environmental cues, change conformation upon ligand binding, and transduce information through the membrane to elicit specific adaptive responses within the cell. For example, adaptive responses to environmental cues that include mechanisms for adaptation and survival in prokaryotes (e.g., surface adhesion, biofilm formation), persistence and virulence of bacterial pathogens<sup>19</sup>, and immune responses to viral and bacterial invasion in eukaryotes are mediated by cyclic dinucleotides, which have emerged as important signaling molecules in most kingdoms of life<sup>20</sup>. In addition to the cyclic dinucleotides, these signaling networks are composed of a vast number of proteins, often associated with the membrane, that are involved in the intracellular production and



**Fig. 3. Microfluidics technology.** Top: Microfluidics devices are used to generate protein crystals and collect diffraction patterns. Bottom: Molecular model for sialate O-acetyltransferase from *Bacteroides vulgatus* (Babnigg, Nocek, and Kim, in preparation).

turnover of the cyclic dinucleotides, as well as effectors that translate the signals into physiological responses. Structural information is critical to understanding small-molecule transport and signal transduction. An understanding of the dynamical functioning of these systems requires the ability to characterize the different chemical and conformational states during the transport reaction or signal transduction, the movements of the transmembrane domains associated with these states, the presence of protein domains intrinsically devoid of ordered structure in the absence of a partner, and the impact of cellular components (carbohydrates, lipids, and small metabolites) on protein structure. For the signaling pathways, the environmental cues and cue-specific enzyme–receptor interactions need to be characterized to decipher how a particular cue is perceived and translated into a specific physiological response.

For all these exemplar systems, numerous technical advances enabled by the APS-U have the potential to radically transform the study of membrane-associated proteins that are highly relevant to the missions of DOE and NIH.

#### *Higher-order macromolecular complexes*

Other important complexes include the so-called “higher-order macromolecular complexes,” which are formed from repeating oligomeric proteins. One example is the microcompartments found across many bacterial phyla, which are self-assembling organelles consisting of an enzymatic core encapsulated by a selectively permeable protein shell<sup>21</sup>. Structural and genetic studies indicate that metabolic substrates enter these microcompartments by passing through the central pores in hexameric assemblies of shell proteins. It was shown that the pore of the major hexameric shell protein in a bacterial microcompartment that metabolizes 1,2-propanediol has a lower energy barrier for passage of the propanediol substrate compared to the toxic propion-aldehyde generated within the microcompartment. This mechanism limits the escape of toxic metabolic intermediates created inside the microcompartments, thus conferring a selective advantage for the host organism<sup>22</sup>.

Another example is that of the bacterial flagella, which is critical for bacterial motility, surface interaction, and pathogenesis. Flagella are extracellular filaments that grow up to 15- $\mu\text{m}$  long and spin at hundreds of revolutions per second, acting like a helical propeller. The flagellar filament has served as an enlightening system for understanding how a protein polymer composed of a single protein, flagellin (except for the cap protein at the end that acts as an assembling chaperone), switches among different states to supercoil. This supercoiling allows the rotating filament to behave as an Archimedean screw and produce thrust. Cryo-EM structures of flagellar filaments from two bacterial species yield a structural model of flagellar filament switching<sup>23</sup>. However, some parts of the cryo-EM structure show poor resolution and combinations of cryo-EM and crystal structures are required to obtain a complete model of the flagellar filament.

Thus, bacterial microcompartments and flagella exist in multiple relevant functional conformations and visualizing these states by combining X-ray crystallography and other complementary techniques is critical to our understanding of the function of these higher-order complexes and to our ability to engineer them. Advances at the APS-U will allow us to analyze smaller and heterogeneous macromolecular crystals to improve structure determination of macro-molecular complexes and assemblies.

#### *Extracellular enzymes and complexes*

Fungi are the major source of industrial enzymes used for lignocellulose degradation and conversion. It was recently discovered that anaerobic fungi from an early branching family of fungi commonly encountered in the digestive tract of ruminants and monogastric herbivores possess a large and comprehensive array of biomass-degrading enzymes<sup>24</sup>. These hydrolytic enzymes are secreted in the extracellular milieu where they degrade the lignocellulose and significantly

contribute to its decomposition into fermentable sugars. The cell-wall-degrading enzymes secreted by anaerobic fungi include both free enzymes and extracellular multi-enzyme molecular machines called “cellulosomes.” Cellulosomes are large, flexible, multiprotein complexes that tether plant biomass-degrading enzymes together for improved hydrolysis. A comprehensive set of over 300 carbohydrate-active enzymes, including conserved scaffoldin and dockerin proteins, have been identified as parts of the fungal cellulosome<sup>25</sup>. Many of the fungal carbohydrate active enzymes originate from bacteria and have acquired fungal dockerin domains that make them part of the fungal cellulosome. These dockerin domains can be grafted to exogenous hydrolytic enzymes to recruit them to the cellulosome. Structural studies will be required to understand the mode of action of the cellulosome at the interface with the lignocellulose. The APS-U will be the ideal X-ray source to rapidly determine structures of enzymes and assemblies from anaerobic fungi and to image at high resolution this interface.

Engineering enzymes that deconstruct biomass is important for the success of the biofuels industry, but understanding their mechanisms of action at the molecular level still represents outstanding fundamental challenges. Most enzymes critical to biomass deconstruction require precise conformational freedom to function optimally. For example, glycoside hydrolases (GH) are a class of enzymes with flexible semi-disordered regions that are necessary for cellulase/hemicellulase function. Rational engineering of these multi-modular enzymes requires insight into the structure and function of the individual domains, as well as the entire enzyme complexed with cellulose<sup>26</sup>. GHs can be mono-functional enzymes (most fungal enzymes), multi-functional enzymes (most bacterial enzymes), or highly complexed enzymes (bacterial or fungal cellulosomes). Fungal GHs are often glycosylated or otherwise post-translationally modified. For example, the glycosylation of *Trichoderma reesei* Cel7A GH at multiple sites is important for increased linker flexibility, thermostability, resistance to proteases, and binding to cellulose<sup>27</sup>.

Importantly, understanding the mechanisms contributing to chemical and enzymatic solubilization of plant cell walls also requires an understanding of the structurally complex plant lignocellulosic biopolymers, which are organized over a wide range of length scales from tissue to cellular to polymeric architectures. For example, correlative imaging in real time was used to assess the nanometer-scale changes in cell wall structure upon pretreatment and subsequent digestion by two commercially relevant cellulase systems<sup>28</sup>. This revealed that the small, non-complexed fungal cellulases deconstruct cell walls using mechanisms that differ considerably from those of the larger multi-enzyme cellulosomes. Microscopic tools, such as atomic force microscopy, stimulated Raman scattering microscopy, and fluorescence microscopy have been applied to investigations of cell wall structure and chemistry from a native wall to a wall treated by thermal chemical pretreatment and enzymatic hydrolysis<sup>29</sup>. Therefore, imaging tools with improved spatial resolution will steadily enhance our fundamental understanding of plant cell wall function and recalcitrance.

The biosynthesis of cellulose chains and their assembly into microfibrils within the cell walls of land plants is mediated by the cellulose synthase complex, a large, multi-subunit transmembrane protein complex. Low-resolution structural analyses using small-angle neutron scattering, small-angle X-ray scattering, and computational modeling revealed that catalytic domains assemble into stable trimer complexes<sup>30</sup>. Further structural studies combining a variety of diffraction and imaging methods will be required to understand the architecture of the cellulose synthase complex and the number of glucan chains that comprise the cellulose microfibrils.

Overall, advanced imaging techniques using X-rays, visible light, electrons, and neutrons will be synergistic with advances in molecular crystallography promoted by the APS-U to gain a mechanistic understanding of the interactions with complex polymeric substrates, and of catalysis in these systems.

## *Investigating the Sequence-Structure-Function Continuum*

Predicting from primary amino acid sequence the level of activity of enzyme variants under particular reaction conditions (e.g., pH, temperature) is currently very challenging. One major hurdle is the lack of available structural information for a large number of enzyme variants displaying a wide range of activity levels because such studies would have prohibitive costs. Still, accurate prediction of activity levels from the enzyme primary sequence is a highly desirable goal as it would dramatically accelerate the rational engineering of enzymes to optimize their activities for specific bioprocesses and assays, with potential applications in environment, bioenergy, and biomedicine. The smaller and brighter APS-U X-ray beams will transform structural biology by enabling the development of new approaches, such as SX, which will dramatically accelerate structure determination while reducing cost. These advances, in combination with ongoing progress in synthetic biology, offer a unique opportunity to investigate protein structure-function relationships at the scale needed to gain a predictive understanding of the structure-function relationship.

Protein engineering is now possible at a very large scale using synthetic biology approaches such as CRISPR-Enabled Trackable Genome Engineering (CREATE)<sup>31</sup>. Libraries of mutants can be generated and subjected to high-throughput functional screens. From these screens, selected variants exhibiting a range of phenotypic outputs (or a specific desired phenotype) can be further evaluated by biochemical assays (e.g., with substrates) to validate and quantify activity, and then be submitted for structure determination. Structural information coupled with activity will guide further protein engineering and molecular modeling, leading to an iterative Design-Build-Test-Learn process that has the potential to accelerate structure-function studies for hundreds to thousands of protein variants. This type of approach requires access to a facility for high-throughput protein purification and characterization. Importantly, there is significant potential to integrate more with other user facilities — in particular, the DOE Joint Genome Initiative's DNA Synthesis Science program, and the Environmental Molecular Sciences Laboratory more generally, which suggests the possibility of incorporation of synchrotrons into the DOE "Facilities Integrating Collaborations for User Science" (FICUS) program.

In higher eukaryotes, when a particular protein or family is implicated functionally in a phenotype at the cellular or whole organism level, understanding the sequence-structure-function relationship becomes critical for assessing the significance of variations that may be observed in the population under relevant conditions. For example, a genome-wide association study of over 540 *Populus trichocarpa* trees revealed the genomic bases of adaptive variation across a wide latitudinal range<sup>32</sup>. These variations can alter the protein-encoding genes through single, non-synonymous nucleotide substitutions and insertion/deletions resulting in protein domain acquisition/loss. They can also result from the natural expression of diverse mRNA splicing variants with and without exon-encoded protein domains<sup>33</sup>. Understanding how these changes affect protein structure is necessary for assessing how functions such as ligand binding and interactions with protein partners or with other macromolecules may be affected under relevant conditions (e.g., pH, metals, small metabolites).

Faster and cheaper structure determination will also lead to reevaluation of how structural biology can help to make sense of the overwhelming amount of (meta)genomic, transcriptomic, proteomic, and genetic data that are being generated to better understand protein function and interactions. Considering current efforts to sequence organisms and identify useful activities from underrepresented branches of the tree of life (e.g., bacteria, protists, anaerobic fungi and archaea that are emerging from human and soil microbiome studies), functional genomics (i.e., the development of new genetic approaches in these systems) and structural genomics efforts will be needed to characterize and engineer these novel activities.

## **E. Newly Emerging Technological Developments and New Tools for Extracting Valuable Information on Structural Dynamics.**

New data collection strategies, particularly SX and other low-dose, high-redundancy methods, will make ambient temperature (rather than cryogenic) data collection routine. Ambient temperature structures generally can provide greater information about protein conformational heterogeneity and yield insight into conformational ensembles under physiological conditions<sup>34,35,36</sup>, which demonstrates greater internal consistency for data collected at room temperature vs. cryo-temperatures. Three main opportunities can be articulated and are briefly described below.

### ***Rapid Perturbation of Proteins Coupled to Rapid Time-Resolved Readouts of X-ray Diffraction Patterns***

Two new techniques for excitation of intrinsic macromolecular structural dynamics are under development for rapid perturbation: electric field pulses<sup>37</sup> and laser-based temperature jump methods<sup>34</sup>. The first drives atomic displacements vectorially, according to the magnitude and orientation of the applied electric field and the distributions of charges and dipoles. The second causes coordinate shifts and fluctuations dictated by the specific redistribution of conformational states that occurs when the thermal energy of the system is increased. Other methods of perturbation could also emerge. For example, time-resolved studies of systems activated by rapid mixing (i.e., of enzyme and substrate) are also under way. The essence of these new approaches is to reveal the accessible excited-state structures of a protein or nucleic acid rather than a single ground-state structure. Very specific discoveries will be possible on each system studied using these new methods, and much will be learned more broadly about macromolecules from the body of collected studies. The coming increases in synchrotron beam flux make these short time-scale exposures possible. Many of the key challenges moving forward in this area concern engineering elements — e.g., how to apply high-voltage and temperature fluctuations to crystals during or while timed to X-ray exposure without physical movement or destruction of the sample. In many of the fluctuation type experiments, room (or other non-cryogenic) temperatures will be important. Shorter X-ray exposures should generally reinvigorate non-cryogenic structural studies, while also calling for further advances in crystal handling. From a practical community engagement point-of-view, the BioCARS facility at APS Sector 14 is gearing up to provide the hardware and expertise through the general user program to carry out these technically challenging experiments.

### ***Dynamic Understanding of Diffuse Scattering***

Many protein crystals show prominent non-Bragg scattering that arises from correlated motions across wide-ranging length scales in crystal specimens. It is generally accepted that diffuse scattering holds much potential information about protein dynamics, but obtaining decisive dynamic interpretations remains largely a frontier problem<sup>38,39</sup>. Ideas and programs for interpreting dynamic information from diffuse scattering data could be advanced substantially with the increased brightness and throughput possible with much brighter beams. Additional sample mounting and delivery technologies will be required to reduce the background in X-ray exposures during data collection; this can be achieved either using larger crystals or eliminating scattering from extraneous material, e.g., employing laser ablation or low-background, barrier-free fixed target mounts. Use of these technologies with the smaller, brighter beams of the APS-U will result in an increased number of high-quality diffuse scattering datasets. The new datasets will greatly accelerate development of data-driven approaches for refining specific models of correlated motions using the diffuse signal. In addition, as the accuracy of the models increases, the combination of diffuse scattering with crystalline molecular dynamics (MD) simulations will create opportunities for improving currently employed MD force fields.

### ***Exploring Conformational Landscapes Under Physiological Conditions***

Recent findings emphasize that individual crystal specimens of substantial size may be non-uniform, with somewhat different unit cells and subtle conformational differences exhibited within a single specimen<sup>40</sup>. Likewise, as bright beams and serial data collection strategies move towards collecting diffraction from very large numbers of very small crystals, it will be important to recognize the likelihood that somewhat different conformations may be encountered in different microcrystals. Recent software developments allow for systematic analysis of differences and grouping of samples at the diffraction data-merging stage between many specimens or oscillation sweeps. Within the scope of SX, different sample delivery modes may be more suitable for analyzing such systems. For example, the ability to collect substantial numbers of reflections from individual microcrystals in fixed-target formats could be an advantage. With the availability of bright, sub-nanosecond-scale X-ray pulses it becomes possible to conduct time-resolved serial microcrystallography, where reactions are initiated with protein crystals and protein structures are determined as a function of time. The sample delivery methods and improvements planned through the APS Upgrade and the next phase of BioCARS will help facilitate these studies. These improvements emphasize the theme of multiple structures from single-crystal forms or single-crystal specimens.

Software developments will continue to play an important role in the advances discussed above as we anticipate greater data throughputs and the special nature of data collection/analysis associated with jumps in temperature and electric fields. The advent of these methods has motivated new algorithms and new software platforms, and these developments will need to continue in order to keep pace with and fully benefit from next-generation synchrotron beams.

### **F. Technical Developments Needed to Realize the Scientific Impact**

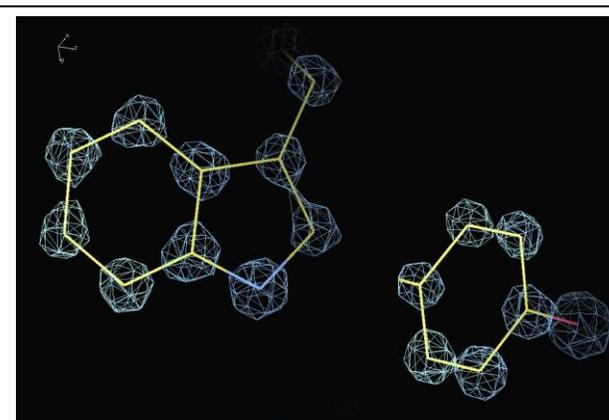
For the research community to benefit from improvements to the APS X-ray source, a number of technical challenges must be addressed. There will be a need for more complex experimental environments that often combine different measurements methods. For many experiments, it will be necessary to make measurements on small crystals that are micro- or nano-millimeter in size. The production, optimization, detection, and harvesting of micro- and nanocrystals is currently not routine. Experimental facilities at beamlines will be needed to grow these crystals routinely, coupled with X-ray-based methods to detect very small crystals and/or technology enabling their shipment (at cryogenic and room temperatures). Delivery of such crystals to the X-ray beam represent another challenge. All of these new experimental capabilities must be accompanied by improved data analysis tools that are able to keep pace with the very large volumes of data being generated, recover signal from weak data, and extract information about dynamics in biological systems. Beyond the technical developments, it is clear that researchers will need to be trained to make use of the new experimental resources. Workshop participants discussed several priority areas where technical advances are required in order to realize the scientific impact promised by the APS-U.

### ***Accelerating Structure Determination of Proteins and Macromolecular Complexes***

During the workshop, there was a strong emphasis on the APS-U X-ray beam properties that will allow it to become a major contributor to SX applications at near-physiological temperatures. In fact, it was suggested that SX may become the default data collection method throughout the field within a few years, and SX should become a major MX application at the APS-U. Structures determined at near-physiological temperatures provide conditions that are much more biologically revealing and avoid artifacts associated with crystal cryo-cooling. With more than a dozen MX stations operating at the upgraded APS and available to users 24 hours/day, 200 days/year, this light source can make SX much more accessible to the broader biology community. Mono-

chromatic and pink beams will permit data collection from small crystals (1 to 10  $\mu\text{m}$ ) while preserving the capability to collect data from larger crystals. SX using pump-probe dynamic experiments can provide high-resolution information about protein dynamics and catalysis, and could enable the study of transient states at the atomic level. Emerging technologies such as mix-and-inject or tape-drives will allow studies of chemical reactions including enzyme single-turnover kinetics. The use of small or microcrystals for time-resolved enzymology allows rapid diffusion through the entire crystal volume on the time scale of chemical transformations from nanoseconds to seconds. This research will be highly complementary to studies conducted at XFELs.

Participants at the workshop felt strongly that the continued development of advanced crystal sample delivery methods at APS-U MX beamlines was essential to best utilize the new source size. Microcrystals will allow static, dynamic, and time-resolved macromolecular studies at unmatched speed and scale, at a wide range of temperatures (10 to 350 K), providing high-resolution and atomic level insight (Fig. 4). Such approaches can access protein conformational ensembles and help us to better understand protein dynamics that confer important function in a living cell. The APS-U integrated suite of experimental and computational capabilities will examine biological molecules at sub-nanometer to millimeter length scales, and on time scales from microseconds to seconds over a wide range of physiologically relevant temperatures and pressures. This will position the APS-U as a powerful, one-of-a-kind tool for structural biology.



**Fig. 4. Atomic resolution structure.** Section of the electron density map (residues W305 and Y19) determined at atomic resolution (1.02 Å, PDB ID 5JH8) at 100 K of chitinase from *Chromobacterium violaceum*. This bioprocessing enzyme hydrolyzes chitin – a polymer of N-acetylglucosamine. The map was contoured at  $4\sigma$  level. The structure was determined at the SBC-CAT. (C. Chang and K. Michalska, SBC-CAT)

### Controlling Radiation Damage

The interaction of high-brightness X-ray beams with biological samples results in radiation damage that can provide useful information but can be detrimental to the sample under study. The workshop participants maintained that an in-depth understanding of these processes is essential for modeling damage-free states. There are many approaches to minimizing radiation damage<sup>41,42</sup>. An example that was presented “painted the sample with APS-U X-rays” by producing virtual beams to match any sample size through optimization of X-ray beam size, energy, and dose. Use of many crystals exposed with a low dose and computationally summed to reconstruct reciprocal space and obtain complete data set without radiation damage may provide significant advantages. Data collection approaches should account for dose contrast effects, which are reduced by dose flattening, dose slicing through rastering the sample in three dimensions, properties of the detector, and other advances that can reduce X-ray damage and maximize signal-to-noise. These approaches call for a new paradigm of sample preparation and delivery for APS-U X-ray beams such as creating new standards including liquid and viscous jets, silicon nitride chips and microfluidics devices. Other methods include laser-reshaping the crystal sample into spheres to reduce the need for absorption corrections, removing crystal satellites, or simply allowing better alignment with the beam size. Further development will be possible with X-ray painting approaches, which could complement simultaneous developments in nanobeams for X-ray imaging through ptychography.

### ***Computational Approaches to Big-Data Collection***

Rapid advances in X-ray diffraction at the APS-U source to study complex biological samples and dynamic processes at high resolution and dramatically faster speeds will generate an enormous amount of data because of the high throughput and complexity of the experiments. The data collected for a single experiment can generate multiple terabytes per hour, further complicating the analysis pipelines and expanding the computational requirements. More efficient approaches are needed to sustain scientific progress in dynamic and SX experiments with the ultimate goal of analyzing data in real time and providing feedback to the experimenters. With SX, the rapid collection of large data sets from millions of crystals represents a new challenge for data transfer, storage, processing, and analysis. The APS-U will need to confront the “Big Data Challenge” and deliver efficient and user-friendly solutions. Interaction with XFELs and cryo-EM user communities may be practical as these fields already must cope with a large amount of data. Computational methods, in particular molecular dynamics/mechanics simulation, are also complementary as novel X-ray-based methods may inform force-field development. Vice versa, computational methods will facilitate the analysis of dynamic systems studied using X-rays.

### ***Improving Protein Production and Crystallization***

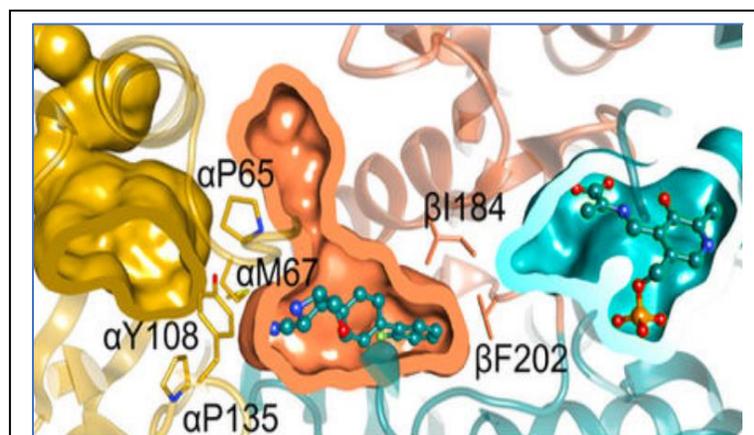
Although the smaller and brighter APS-U X-ray beams combined with improved computational pipelines will enable the generation of 3D structures from less than perfect samples, the expression and crystallization of target proteins and complexes will remain a primary limitation. Isolation of many pure proteins and protein complexes either from natural sources or by recombinant DNA methods is still very challenging. In addition, growth of diffraction-quality crystals is a bottleneck in MX. Although structural genomics programs have advanced many methods for crystal growth, new, more effective, and higher throughput approaches are needed to meet the future needs of MX at the APS-U. Approaches such as SX will require a large number of crystals and also different kinds of crystals. It is expected that to understand biologically relevant interactions and mechanisms, crystals of many variants of proteins and complexes will be needed, including those containing proteins trapped in different conformational and mechanistic states. Crystals of mutants and orthologs with altered properties and new activities will help in understanding the evolution of protein function. Structure-based drug discovery will also demand large numbers of crystals for drug targets in complex with hundreds of small ligands (drug candidates) and antibodies. Finally, dynamic and time-resolved studies will require millions of micro-crystals for SX.

Workshop participants called for renewed crystallization research that seeks truly innovative approaches to monitor, optimize, and accelerate protein crystallization. One example of promising approaches for higher throughput crystallization is droplet-based microfluidics. Small droplets can help reduce convective effects and create a diffusion-controlled environment in which well-defined concentration and temperature gradients can be generated. This technique provides more uniform crystals with fine control over crystal size, reduces the quantity of protein sample needed for an experiment, and can be adapted to efficiently deliver crystals to the X-ray beam (Fig. 3). An emerging technology involves the design of new scaffolds for crystal engineering. Programming self-assembled designer DNA crystals with various lattices and functions is one example of nanofabrication using nucleic acids<sup>43</sup>, but other scaffolds will need to be investigated. Engineered stabilizing antibody Fab fragments have been exploited with success as crystallization chaperones to facilitate the structure determination and enhance diffraction resolution of challenging, previously recalcitrant macromolecules and may serve as a universal strategy to enhance the success rates (see the section below). Lipid nanodiscs have been successfully used to enhance crystallization of membrane proteins. This technology, when combined with stabilizing antibody Fab fragments, can stabilize transient functional states of

membrane proteins in native-like lipid environments, allowing for their structural and biochemical characterization<sup>44</sup>. Chemical surface modifications of common crystallization platforms may also provide new avenues for crystallization enhancements. Finally, deep-learning computational approaches have been used for sequence-based protein crystallization prediction. These advanced methods can identify proteins that can produce diffraction quality crystals without the need to manually engineer additional biochemical and structural features. For example, the DeepCrystal learning framework can exploit frequently occurring k-mers and sets of k-mers from the protein sequences to distinguish proteins that will result in diffraction quality crystals<sup>45</sup>. The development of such emerging methods must be actively pursued in order to accelerate structure determination and fully benefit from next-generation synchrotron beams.

As the production of crystals will evolve, best practices for crystal growth, harvesting, data collection, and processing will need to be determined to help establish a strong user base. Experimental facilities at beamlines should have capabilities to grow microcrystals routinely, X-ray-based methods to detect microcrystals, and/or technology enabling shipment of microcrystals (cryo-protected or ambient-temperature), and should be spearheading development of best practices. A proposed way to make these methods more accessible to potential users is through publications detailing methods, versioned pre-prints, or other public repositories that are potentially effective communication channels. (User training at synchrotrons through workshops is another time-honored approach.)

In the biomedical and microbiological sciences, many highly challenging proteins are directly in line with the goals of NIH, particularly with targets that could be used in immunotherapy and/or



**Fig. 5. Allosteric site drug discovery.** Crystal structure of Mtb tryptophan synthase with novel allosteric BRD4592 inhibitor. The figure shows pockets in the protein structure. The alpha subunit pocket is in yellow, the beta-subunit with bound aminoacrylate-PLP cofactor adduct in cyan and inhibitor-binding pocket on alpha-beta interface is in orange. Key residues at the pocket-pocket interface are labelled<sup>3</sup>.

vaccine/drug discovery and development (Fig. 5). In environmental sciences and bioenergy, eukaryotic proteins such as those involved in photosynthesis, resilience to drought, nutrient deficiency, and lignocellulosic break-down are particularly recalcitrant to traditional expression and crystallization methods due to complex folding and post-translational modifications that are difficult to replicate using standard expression pipelines. Expanding the user base to include plant and fungal scientists from within BER-funded Bioenergy Research Centers, Scientific Focus Areas and including single principal investigator projects, will require the development of robust and scalable techniques for the expression of

eukaryotic proteins. This will require the development of alternate expression platforms (i.e., other than *E. coli* and yeast) for plant and fungal gene constructs that allow the post-translational modifications of eukaryotic proteins. Workshop participants recommended that protein purification and crystallization services be supported as a part of the user program at the APS-U.

In some special cases, protein crystals can be grown *in cellulo* and data collected with microbeams. Such “*in vivo*” structures with more realistic cellular environments can provide important functional information<sup>46</sup>. High-resolution crystal structures of proteins or protein complexes containing unknown, potentially heterogeneous ligands or post-translational modifications can be

determined and ligands can be identified from these structures. The APS-U capabilities with highly bright beams will clearly enable such research.

### **Antibodies as Crystallization Chaperones**

Antibody fragments, including two-chain fragment antigen binding (Fab) and camelid single-domain antibody fragments (“nanobodies”) have emerged as especially useful tools in protein structural biology, facilitating studies of conformationally dynamic proteins such as GPCRs and integrins<sup>47</sup>. Recently, an *in vitro* platform for nanobody discovery based on yeast surface display was developed for identifying nanobodies and was utilized to discover conformationally selective nanobodies to two distinct human GPCRs<sup>48</sup>. (To facilitate broad deployment of this platform, the library and associated protocols are freely available for nonprofit research.) This new technology has the potential to facilitate the use of antibodies outside of the biomedical research area into plant science and microbiome science where antibodies could be used as crystallization-facilitating chaperones of more complex, multiple domain, disulfide-linked, glycosylated, and recalcitrant crystallization targets. Fabs also can be produced by yeast display, are more diverse in the types of epitopes they recognize, and their larger size has advantages in some applications including blocking function and negative stain electron microscopy (EM). Antibodies also account for about half of new Federal Drug Administration drug approvals each year. Antibodies facilitate protein purification, functional state locks, and correlations between function and structures. Antibodies can block biochemical or cellular functions and can make important contributions that are orthogonal to structure determination to help understand protein function and interactions within complex biological pathways in humans and model organisms, and in mostly unexplored organisms. They offer an approach to rethinking/redeploying structural genomics. Antibodies are also extremely useful in negative stain EM, which can be carried out on proteins with flexible domains, and in cryo-EM. They could greatly contribute, for example, to work on complex systems such as cellulases and the cellulosome, which featured prominently in this workshop.

## **G. Synergistic Structural and Imaging Techniques**

Workshop participants also identified a synergistic challenge, namely the need to provide strategies for integrating atomic structures with imaging technologies that together span the breadth of biological scales. The advances in MX promoted by the APS-U need to be complementary with advanced imaging techniques using X-rays, visible light, electrons, and neutrons to gain a mechanistic understanding of the interactions of proteins and complexes with their operating environment. A sister workshop focused on “Opportunities in Biological and Environmental Research Uniquely Enabled by the APS Upgrade” was held at Argonne in late August 2018.

### **Cryo-Electron Microscopy**

With the advent of modern cryo-EM methods, some fraction of structure determination may shift to this approach, especially for large complexes and membrane proteins that remain challenging targets for single-crystal crystallography. Cryo-EM benefits from the ability to determine the structures of macromolecules without crystallizing them, although preparation of good samples can be challenging. The technique has been developing for more than 40 years, but recent technological breakthroughs further enabled routine structure determinations using the single-particle reconstruction approach. Conventional single-particle imaging is facilitated by spatial and temporal immobilization of radiation-sensitive material in vitreous ice under cryogenic conditions, in concert with dose-limited imaging and sensitive detectors and post experimental computational processing. The resulting data are used to reconstruct three-dimensional images of structures, sometimes in alternative conformations. This methodology has been utilized successfully to elucidate macromolecular systems that are not amenable to crystallization and analysis by

traditional MX methods. Cryo-EM shows a distinct advantage for imaging larger, potentially more disordered assemblies and offers insight into conformational and energy landscapes of macromolecules. In principle the “new” cryo-EM allows one to bridge the gap between atomic and cellular scale imaging of many biological systems and is profoundly changing structural biology by facilitating important discoveries. However, crystallography retains several unique advantages that include the power to deliver high-resolution dynamic information as a function of time, temperature, pressure, and other perturbations, and the ability for experimental high-throughput via emerging X-ray crystallography methods (SX, time-resolved, variable temperature). These methods will be enabled by upgraded synchrotron sources such as the APS-U and will provide unique experimental avenues to probe the dynamics of biological systems.

Ultimately, the future of structural biology will depend on how the individual strengths of the MX and cryo-EM approaches are utilized in both unique and complementary ways to attack problems at the biological frontiers. Synchrotron beamlines also serve as a model for efficient automated data collection and for accessibility by the user community. Continued development of algorithms, software, and hardware will allow them to continue serving as an example for other complementary techniques. For example, structural biology techniques other than MX, notably cryo-EM, are poised to benefit from the ongoing development of software for rapid data collection and analysis<sup>7,8</sup>. Developments in analytical methods will benefit both MX and cryo-EM fields. In both techniques, a core problem is the same: How can we model alternative conformational states from observed data that contains contributions from them all?

### ***X-ray Imaging Techniques***

Synchrotrons are perfectly positioned to provide synergistic state-of-the-art tools for users, from characterizing metal-binding proteins at the atomic level to imaging the trafficking and compartmentalization of metal ions at the cellular and subcellular levels to quantifying ligand environments and oxidation states. Therefore, building synergy between structural beamlines and the bio-imaging beamlines, such as the Bionanoprobe at the APS, is critical for transforming our understanding of the biology of metal ions and their interactions with other essential biological processes, such as light-capture, carbon and nitrogen metabolism, innate immunity, and neurological diseases. The resulting information would impact a broad range of fields including engineering for improved bioenergy crop resilience, understanding ecosystem interfaces, improving human health, and studying disease<sup>49</sup>. Trace metals play important roles in biological function, and X-ray fluorescence microscopy (XFM) provides a way to quantitatively image their distribution within cells. Synchrotron-based XFM utilizing advanced X-ray focusing optics, offers highly suitable capabilities for performing trace element studies of biological samples: The penetrating power of X-rays allows one to image many-micron-thick biological samples — including whole cells — in a way that visible light or electron microscopes cannot. Moreover, the sensitivity of XFM is several orders of magnitude better than electron-based techniques. The capability of imaging cryo-protected samples in both two and three dimensions with sub-50-nm resolution in various X-ray modes has greatly advanced a broad range of scientific studies<sup>50</sup>.

Cells and biological materials are large objects in comparison to the size of internal components such as organelles and proteins. An understanding of the functions of these nanoscale elements is key to elucidating cellular function. X-ray microscopy can be used to image whole, un-sectioned cells in their native hydrated state. It complements the higher resolution of electron microscopy for sub-micrometer-thick specimens and the molecule-specific imaging capabilities of fluorescence-light microscopy. However, radiation damage is one of the most severe challenges in the field. A valuable alternative is to cryo-protect samples. Fast, continuous X-ray scanning of cryo-protected hydrated cells can be employed for simultaneous sub-20-nm-resolution ptychographic transmission imaging with high contrast, and sub-100-nm-resolution de-convolved

X-ray fluorescence imaging of diffusible and bound ions at native concentrations without the need to add specific labels. By working with cells that have been flash cryo-protected without the use of chemical fixatives, and imaging them under cryogenic conditions, it is possible to obtain images with well-preserved structural and chemical composition and sufficient stability against radiation damage to obtain multiple images with no observable change<sup>51,52</sup>. Importantly, future progress in X-ray imaging techniques (e.g., XFM and ptychography) enabled by the vastly improved brightness and coherence of the APS-U X-ray beams, and how these advances will allow investigation of cell chemistry, cell structure, and biogeochemistry in biological and environmental research were assessed during the sister workshop “Opportunities in Biological and Environmental Research Uniquely Enabled by the APS Upgrade.”

Of note, the high-brightness APS-U X-ray beams will enable faster data collection using soft X-rays for native-element phasing of diffraction data at atomic resolution with sulfur elemental contrast. Such approaches (e.g., sulfur single-wavelength anomalous diffraction (S-SAD)) will enhance structural determination of proteins and domains that do not have adequate molecular replacement candidates, addressing a problem that expands as the “dark matter” of new genomes are characterized (including the microbiome and for species relevant to the DOE mission in the fields of biomass and energy).

### ***Small-Angle X-ray Scattering and Small-Angle Neutron Scattering***

Beyond the crystalline environment, scattering methods such as small-angle X-ray scattering (SAXS) and small-angle neutron scattering (SANS) will remain powerful tools for studying biomolecular structure and dynamics in solution, albeit at lower resolution. The ability to study biological systems in solution using scattering methods, coupled with high brightness beams, will potentially enable studies of weakly interacting proteins and rare events.

SAXS is an established technique that provides low-resolution structural information on macromolecules. It is used to study the size and shape of macromolecules in solution and is capable of yielding structural information on systems with size and conformational polydispersity, including highly flexible objects. When utilizing high-flux synchrotron facilities, time-resolved SAXS allows analysis of kinetic processes over time ranges from sub-millisecond to hours and can investigate structural changes related to processes such as protein folding and conformational transformations<sup>53</sup>. Time-resolved SAXS was used to investigate the assembly of sulfonated cellulose nanocrystals dispersed in a shrinking levitating aqueous drop using time-resolved measurements<sup>54</sup>. Time-resolved SAXS was used to study the R-to-T transition in carbon-monoxymyoglobin (HbCO) following a temperature jump. These studies provide unambiguous evidence for more than one R state, and unveil the time scale for interconversion between them. This structural heterogeneity must be accounted for to properly explain the kinetic heterogeneity observed in time-resolved spectroscopic studies following photolysis of HbCO<sup>55</sup>.

SANS takes advantage of the very different neutron scattering cross-sections of hydrogen (H) and deuterium (D) atoms, making it possible to selectively highlight different components within a complex system in solution by selective isotope incorporation. SANS provides unique information about complexes of biomolecules in hierarchical structures<sup>56</sup> and heterogeneous complexes, which is highly complementary to the structural restraints provided by MX, nuclear magnetic resonance imaging, and cryo-EM. Systems studied with SANS include multi-protein or protein–DNA/RNA complexes and solubilized membrane proteins.

### ***Neutron Imaging Techniques***

Neutron crystallography provides highly complementary structural information. Because neutrons interact with nuclei, it is possible to observe lighter elements such as H and D, enabling studies

of H bonding networks and protonation states of catalytic residues. Moreover, neutrons do not cause radiation damage. Other complementary techniques that utilize neutrons include inelastic and quasi-elastic neutron scattering, which provides information about vibrational modes, molecular motions, and diffusive properties of biomolecules and their hydration on the pico-second-to-nanosecond timescale, and neutron spin echo spectroscopy that probes slower motions in proteins. Neutron imaging includes neutron radiography and computed tomography, which take advantage of H/D contrast and the nondestructive nature of neutrons. An example is provided by neutron and atomic-resolution X-ray crystallographic studies of the pseudo-Michaelis ternary complex of *E. coli* DHFR with folate and NADP(+) that revealed the existence of the keto tautomer and suggested that tautomerization is unnecessary for catalysis. A modified catalytic mechanism was proposed that involves stabilization of the keto tautomer of the substrate involving protonation by water that gains access to the active site<sup>57</sup>.

### ***Molecular Dynamic Simulations***

These new directions will intersect in important ways with other structural techniques. A key synergistic method is MD simulations. Structure snapshots in time-resolved X-ray diffraction can serve as steering targets or validation checkpoints in MD simulations, revealing putative dynamical mechanisms after induced structural perturbations. As the number of high-quality diffuse scattering data sets increases, MD approaches will grow in importance for interpreting the data in terms of conformational ensembles. The combination of MD simulations with time-resolved X-ray diffraction and diffuse scattering creates opportunities for improving currently employed force fields. Multiconformer X-ray results will also connect to dynamical information from nuclear magnetic resonance and conformational heterogeneity seen in high-resolution EM structures. For example, the metallo  $\beta$ -lactamase NDM-1 is an extremely promiscuous enzyme that hydrolyzes and inactivates nearly all known  $\beta$ -lactam-based antibiotics with startling efficiency. NDM-1 can utilize different metal cofactors and seems to exploit an alternative mechanism based on the reaction conditions. A model of the enzyme's catalytic pathway was built using a combination of experimental (X-ray crystallography, biochemical assays) and theoretical (numerical simulation) methods to examine substrate binding, metal binding, and catalytic mechanism of the enzyme<sup>58</sup>. The NDM-1 enzyme interacts with the substrate solely through zinc, or other metals, bound in the active site, explaining the observed lack of specificity against a broad range of  $\beta$ -lactam antibiotic agents. The zinc ions also serve to activate a water molecule that hydrolyzes the  $\beta$ -lactam ring through a proton shuttle.

## **H. Partnership with the User Community**

How do we meet the future needs of the user community? In order to leverage the technological advances associated with the APS-U to maximize scientific impact, an expansion of the current APS user base will be needed. Workshop participants identified several key areas that will need investments in order to engage and expand the APS-U user base.

### ***Sample Delivery, Automation of Beamlines, and User Interface Standardization***

Workshop participants from academia expressed a desire to make advances in automation and user interface standardization in parallel to the APS-U. Sample preparation technologies that take advantage of the new beam capabilities — mostly for the detection and growth of microcrystals and high-throughput mounting strategies of such crystals — have to be developed and disseminated throughout the structural biology community.

Purpose built, user friendly beamlines are needed, in particular those providing multiple sample delivery technologies and measurement methods. User support before, during, and after data collection will be needed to equip users with the necessary APS-U-specific knowledge and tools

for efficient use of beam time and to maximize biological insight gained at the beamlines. A proposal to facilitate access to first time users is to associate them with an “ambassador” from the facility who would provide expertise on sample preparation, selection of workflow, and data interpretation.

Data collection and transfer may pose additional burdens on the field, as well as processing data that is collected using these newer approaches. Will there be easy-to-use, open-source software available for use by non-aficionados? Will there be data depositories for academic labs to store much larger raw datasets that will become possible through the use of SX and other new methods? Workshop participants suggested that because of the APS leadership position in MX in the U.S. there should be an APS-U institutional commitment to synchrotron MX users similar to that at the Cornell High Energy Synchrotron Source at Cornell University and the Stanford Synchrotron Radiation Lightsource at Stanford University. After the upgrade, the APS should become the leader in serial crystallography and time-resolved enzymology, and should encourage a paradigm switch in serial MX, including the development of a suite of crystal delivery methods. In parallel, the APS-U should organize and host conferences, seek grants, and become the most visible player in advanced MX.

### **User Training**

As new capabilities are added to the APS-U portfolio, active involvement of the user community will be important in realizing the full potential of these capabilities. Users will design sophisticated experiments, develop accessory instrumentation and software, and provide samples and person-power. Users will also provide feedback and recommendations for further APS-U improvements and expansion. Predicting the course of coming scientific advances in advance and in detail is difficult. Therefore, there should be sufficient room and flexibility for changes and needed modifications. As new capabilities are expanded into a broader user base, scientific progress will accelerate and the APS-U impact will grow rapidly.

Translation of the protocols to researchers will be of paramount importance but challenging considering that many current researchers do remote collection and are not on site. Ideas for implementation are:

- Online tutorials that can introduce the technology with contact information of experts in the technique that can be used for troubleshooting.
- Incorporation of these newer techniques in workshops and community based resources, like The Collaborative Computational Project Number 4, so that students and other attendees can develop sufficient technical expertise to utilize these techniques. The crystallography course at Cold Spring Harbor is another idea. Videos of these sessions could be posted online for others to view and learn from. The NIH training grants recently awarded to institutions for the teaching of single-molecule electron microscopy can also serve as models for training at or by the APS.
- Encouraging researchers to collect data on-site, but with a clear understanding of how they need to prepare and ship samples so they have the highest likelihood of collecting useful data.

### **I. References**

1. Weierstall, U. *et al.* Lipidic cubic phase injector facilitates membrane protein serial femtosecond crystallography. *Nat Commun* **5**, 3309, doi:10.1038/ncomms4309 (2014).
2. Michalska, K. *et al.* In situ X-ray data collection and structure phasing of protein crystals at Structural Biology Center 19-ID. *J Synchrotron Radiat* **22**, 1386-1395, doi:10.1107/S1600577515016598 (2015).

3. Wellington, S. *et al.* A small-molecule allosteric inhibitor of Mycobacterium tuberculosis tryptophan synthase. *Nat Chem Biol* **13**, 943-950, doi:10.1038/nchembio.2420 (2017).
4. Martin-Garcia, J. M. *et al.* Serial millisecond crystallography of membrane and soluble protein microcrystals using synchrotron radiation. *IUCrJ* **4**, 439-454, doi:10.1107/S205225251700570X (2017).
5. Wiedorn, M. O. *et al.* Megahertz serial crystallography. *Nat Commun* **9**, 4025, doi:10.1038/s41467-018-06156-7 (2018).
6. Grunbein, M. L. *et al.* Megahertz data collection from protein microcrystals at an X-ray free-electron laser. *Nat Commun* **9**, 3487, doi:10.1038/s41467-018-05953-4 (2018).
7. Cheng, R. K. Y., Abela, R. & Hennig, M. X-ray free electron laser: opportunities for drug discovery. *Essays Biochem* **61**, 529-542, doi:10.1042/EBC20170031 (2017).
8. Shoemaker, S. C. & Ando, N. X-rays in the Cryo-Electron Microscopy Era: Structural Biology's Dynamic Future. *Biochemistry* **57**, 277-285, doi:10.1021/acs.biochem.7b01031 (2018).
9. Mizohata, E., Nakane, T., Fukuda, Y., Nango, E. & Iwata, S. Serial femtosecond crystallography at the SACLA: breakthrough to dynamic structural biology. *Biophys Rev* **10**, 209-218, doi:10.1007/s12551-017-0344-9 (2018).
10. Weinert, T. *et al.* Serial millisecond crystallography for routine room-temperature structure determination at synchrotrons. *Nat Commun* **8**, 542, doi:10.1038/s41467-017-00630-4 (2017).
11. Stauch, B. & Cherezov, V. Serial Femtosecond Crystallography of G Protein-Coupled Receptors. *Annu Rev Biophys* **47**, 377-397, doi:10.1146/annurev-biophys-070317-033239 (2018).
12. Seeger, M. A. Membrane transporter research in times of countless structures. *Biochim Biophys Acta Biomembr* **1860**, 804-808, doi:10.1016/j.bbamem.2017.08.009 (2018).
13. Zhang, M. *et al.* Structural insights into the light-driven auto-assembly process of the water-oxidizing Mn<sub>4</sub>CaO<sub>5</sub>-cluster in photosystem II. *Elife* **6**, doi:10.7554/eLife.26933 (2017).
14. Young, I. D. *et al.* Structure of photosystem II and substrate binding at room temperature. *Nature* **540**, 453-457, doi:10.1038/nature20161 (2016).
15. Schuth, N. *et al.* Kalpha X-ray Emission Spectroscopy on the Photosynthetic Oxygen-Evolving Complex Supports Manganese Oxidation and Water Binding in the S<sub>3</sub> State. *Inorg Chem* **57**, 10424-10430, doi:10.1021/acs.inorgchem.8b01674 (2018).
16. Santos, R. *et al.* A comprehensive map of molecular drug targets. *Nat Rev Drug Discov* **16**, 19-34, doi:10.1038/nrd.2016.230 (2017).
17. Brown, N. A., Schrevens, S., van Dijck, P. & Goldman, G. H. Fungal G-protein-coupled receptors: mediators of pathogenesis and targets for disease control. *Nat Microbiol* **3**, 402-414, doi:10.1038/s41564-018-0127-5 (2018).
18. Pandey, S. Heterotrimeric G-protein regulatory circuits in plants: Conserved and novel mechanisms. *Plant Signal Behav* **12**, e1325983, doi:10.1080/15592324.2017.1325983 (2017).
19. Hardison, R. L. *et al.* Microevolution in response to transient heme-iron restriction enhances intracellular bacterial community development and persistence. *PLoS Pathog* **14**, e1007355, doi:10.1371/journal.ppat.1007355 (2018).
20. Krasteva, P. V. & Sondermann, H. Versatile modes of cellular regulation via cyclic dinucleotides. *Nat Chem Biol* **13**, 350-359, doi:10.1038/nchembio.2337 (2017).
21. Kerfeld, C. A., Aussignargues, C., Zarzycki, J., Cai, F. & Sutter, M. Bacterial microcompartments. *Nat Rev Microbiol* **16**, 277-290, doi:10.1038/nrmicro.2018.10 (2018).
22. Park, J., Chun, S., Bobik, T. A., Houk, K. N. & Yeates, T. O. Molecular Dynamics Simulations of Selective Metabolite Transport across the Propanediol Bacterial Microcompartment Shell. *J Phys Chem B* **121**, 8149-8154, doi:10.1021/acs.jpcc.7b07232 (2017).
23. Wang, F. *et al.* A structural model of flagellar filament switching across multiple bacterial species. *Nat Commun* **8**, 960, doi:10.1038/s41467-017-01075-5 (2017).

24. Solomon, K. V. *et al.* Early-branching gut fungi possess a large, comprehensive array of biomass-degrading enzymes. *Science* **351**, 1192-1195, doi:10.1126/science.aad1431 (2016).
25. Haitjema, C. H. *et al.* A parts list for fungal cellulosomes revealed by comparative genomics. *Nat Microbiol* **2**, 17087, doi:10.1038/nmicrobiol.2017.87 (2017).
26. Taylor, L. E., 2nd *et al.* Engineering enhanced cellobiohydrolase activity. *Nat Commun* **9**, 1186, doi:10.1038/s41467-018-03501-8 (2018).
27. Amore, A. *et al.* Distinct roles of N- and O-glycans in cellulase activity and stability. *Proc Natl Acad Sci U S A* **114**, 13667-13672, doi:10.1073/pnas.1714249114 (2017).
28. Ding, S. Y. *et al.* How does plant cell wall nanoscale architecture correlate with enzymatic digestibility? *Science* **338**, 1055-1060, doi:10.1126/science.1227491 (2012).
29. Zeng, Y., Himmel, M. E. & Ding, S. Y. Visualizing chemical functionality in plant cell walls. *Biotechnol Biofuels* **10**, 263, doi:10.1186/s13068-017-0953-3 (2017).
30. Vandavasi, V. G. *et al.* A Structural Study of CESA1 Catalytic Domain of Arabidopsis Cellulose Synthesis Complex: Evidence for CESA Trimers. *Plant Physiol* **170**, 123-135, doi:10.1104/pp.15.01356 (2016).
31. Garst, A. D. *et al.* Genome-wide mapping of mutations at single-nucleotide resolution for protein, metabolic and genome engineering. *Nat Biotechnol* **35**, 48-55, doi:10.1038/nbt.3718 (2017).
32. Evans, L. M. *et al.* Population genomics of *Populus trichocarpa* identifies signatures of selection and adaptive trait associations. *Nat Genet* **46**, 1089-1096, doi:10.1038/ng.3075 (2014).
33. Xie, M. *et al.* A 5-Enolpyruvylshikimate 3-Phosphate Synthase Functions as a Transcriptional Repressor in *Populus*. *Plant Cell* **30**, 1645-1660, doi:10.1105/tpc.18.00168 (2018).
34. Fraser, J. S. & Jackson, C. J. Mining electron density for functionally relevant protein polyesterism in crystal structures. *Cell Mol Life Sci* **68**, 1829-1841, doi:10.1007/s00018-010-0611-4 (2011).
35. Keedy, D. A. *et al.* An expanded allosteric network in PTP1B by multitemperature crystallography, fragment screening, and covalent tethering. *Elife* **7**, doi:10.7554/eLife.36307 (2018).
36. Keedy, D. A. *et al.* Mapping the conformational landscape of a dynamic enzyme by multitemperature and XFEL crystallography. *Elife* **4**, doi:10.7554/eLife.07574 (2015).
37. Hekstra, D. R. *et al.* Electric-field-stimulated protein mechanics. *Nature* **540**, 400-405, doi:10.1038/nature20571 (2016).
38. Van Benschoten, A. H. *et al.* Measuring and modeling diffuse scattering in protein X-ray crystallography. *Proc Natl Acad Sci U S A* **113**, 4069-4074, doi:10.1073/pnas.1524048113 (2016).
39. Wall, M. E., Wolff, A. M. & Fraser, J. S. Bringing diffuse X-ray scattering into focus. *Curr Opin Struct Biol* **50**, 109-116, doi:10.1016/j.sbi.2018.01.009 (2018).
40. Thompson, M. C., Cascio, D. & Yeates, T. O. Microfocus diffraction from different regions of a protein crystal: structural variations and unit-cell polymorphism. *Acta Crystallogr D Struct Biol* **74**, 411-421, doi:10.1107/S2059798318003479 (2018).
41. Finprock, Y. Z. *et al.* Mitigation of X-ray damage in macromolecular crystallography by submicrometre line focusing. *Acta Crystallogr D Biol Crystallogr* **69**, 1463-1469, doi:10.1107/S0907444913009335 (2013).
42. Sanishvili, R. *et al.* Radiation damage in protein crystals is reduced with a micron-sized X-ray beam. *Proc Natl Acad Sci U S A* **108**, 6127-6132, doi:10.1073/pnas.1017701108 (2011).
43. Zhang, F., Simmons, C. R., Gates, J., Liu, Y. & Yan, H. Self-Assembly of a 3D DNA Crystal Structure with Rationally Designed Six-Fold Symmetry. *Angew Chem Int Ed Engl* **57**, 12504-12507, doi:10.1002/anie.201807223 (2018).

44. Dominik, P. K. *et al.* Conformational Chaperones for Structural Studies of Membrane Proteins Using Antibody Phage Display with Nanodiscs. *Structure* **24**, 300-309, doi:10.1016/j.str.2015.11.014 (2016).
45. Elbasir, A. *et al.* DeepCrystal: A Deep Learning Framework for Sequence-based Protein Crystallization Prediction. *Bioinformatics*, doi:10.1093/bioinformatics/bty953 (2018).
46. Banerjee, S. *et al.* Structure of a heterogeneous, glycosylated, lipid-bound, in vivo-grown protein crystal at atomic resolution from the viviparous cockroach *Diploptera punctata*. *IUCrJ* **3**, 282-293, doi:10.1107/S2052252516008903 (2016).
47. Springer, T. A. & Dustin, M. L. Integrin inside-out signaling and the immunological synapse. *Curr Opin Cell Biol* **24**, 107-115, doi:10.1016/j.ceb.2011.10.004 (2012).
48. McMahon, C. *et al.* Yeast surface display platform for rapid discovery of conformationally selective nanobodies. *Nat Struct Mol Biol* **25**, 289-296, doi:10.1038/s41594-018-0028-6 (2018).
49. Jin, Q. *et al.* Preserving elemental content in adherent mammalian cells for analysis by synchrotron-based x-ray fluorescence microscopy. *J Microsc* **265**, 81-93, doi:10.1111/jmi.12466 (2017).
50. Chen, S. *et al.* The Bionanoprobe: Synchrotron-based Hard X-ray Fluorescence Microscopy for 2D/3D Trace Element Mapping. *Micros Today* **23**, 26-29, doi:10.1017/S1551929515000401 (2015).
51. Deng, J. *et al.* X-ray ptychographic and fluorescence microscopy of frozen-hydrated cells using continuous scanning. *Sci Rep* **7**, 445, doi:10.1038/s41598-017-00569-y (2017).
52. Hemonnot, C. Y. J. & Koster, S. Imaging of Biological Materials and Cells by X-ray Scattering and Diffraction. *ACS Nano* **11**, 8542-8559, doi:10.1021/acsnano.7b03447 (2017).
53. Tuukkanen, A. T., Spilotros, A. & Svergun, D. I. Progress in small-angle scattering from biological solutions at high-brilliance synchrotrons. *IUCrJ* **4**, 518-528, doi:10.1107/S2052252517008740 (2017).
54. Liu, Y. *et al.* Assembly of cellulose nanocrystals in a levitating drop probed by time-resolved small angle X-ray scattering. *Nanoscale* **10**, 18113-18118, doi:10.1039/c8nr05598j (2018).
55. Cho, H. S. *et al.* Dynamics of Quaternary Structure Transitions in R-State Carbonmonoxyhemoglobin Unveiled in Time-Resolved X-ray Scattering Patterns Following a Temperature Jump. *J Phys Chem B*, doi:10.1021/acs.jpcc.8b07414 (2018).
56. Mahieu, E. & Gabel, F. Biological small-angle neutron scattering: recent results and development. *Acta Crystallogr D Struct Biol* **74**, 715-726, doi:10.1107/S2059798318005016 (2018).
57. Wan, Q. *et al.* Toward resolving the catalytic mechanism of dihydrofolate reductase using neutron and ultrahigh-resolution X-ray crystallography. *Proc Natl Acad Sci U S A* **111**, 18225-18230, doi:10.1073/pnas.1415856111 (2014).
58. Kim, Y. *et al.* NDM-1, the ultimate promiscuous enzyme: substrate recognition and catalytic mechanism. *FASEB J* **27**, 1917-1927, doi:10.1096/fj.12-224014 (2013).





## Photon Sciences

Argonne National Laboratory  
9700 South Cass Avenue, Bldg. 401  
Argonne, IL 60439

[www.anl.gov](http://www.anl.gov)



Argonne National Laboratory is a U.S. Department of Energy  
laboratory managed by UChicago Argonne, LLC