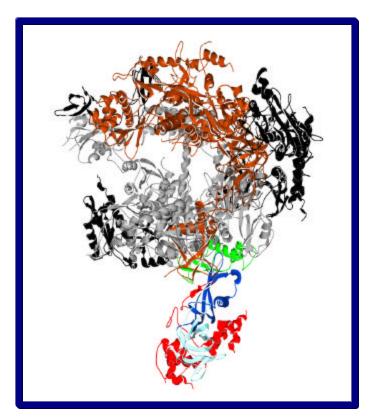# *Access Use of Synchrotron X-Radiation for Research in Structural Molecular Biology*

April 2003

**Prepared by a Subcommittee of the
Biological and Environmental Research
Advisory Committee
Dr. Janet Smith
Subcommittee Chairperson**

Cover photograph - Top view of crystallographic structure of complete 12-subunit RNA Polymerase II (pol II), courtesy of David Bushnell and Roger Kornberg of Stanford University. Published in *Proceedings of the National Academy of Sciences (USA)* **100**, 6969–6973 (2003). Structural data obtained at the Advanced Light Source (Lawrence Berkeley National Laboratory) and Stanford Synchrotron Radiation Laboratory.

Executive Summary

A Subcommittee of the Biological and Environmental Research Advisory Committee (BERAC) was established in response to the charge given in a letter from Dr. Raymond Orbach dated January 9, 2003. The Subcommittee met on February 11, 2003, at the University of Chicago to discuss the use of synchrotron X-radiation for research in structural molecular biology. The Subcommittee membership is listed at the end of this Executive Summary.

The Subcommittee considered current and future use of synchrotron X-rays for three types of experiments: X-ray crystallography, X-ray spectroscopy and small-angle scattering/diffraction. We did not discuss in detail the evolving use of synchrotron X-rays for biological microscopy. The Subcommittee also had extensive discussions on the potential use of 4th generation synchrotron sources for biological experiments.

At the end of its deliberations, the Subcommittee arrived at a strong consensus on the following major conclusions and recommendations.

1. Synchrotron X-radiation is a critical resource for structural biology. Synchrotron-based X-ray photon techniques make unique contributions to understanding biological structure and function. This long-standing trend will continue into the foreseeable future. The improved stability, reliability and rapid accessibility of synchrotron sources has had a major positive impact on structural biology experiments both in the US and worldwide. Department of Energy (DOE) synchrotron user facilities, operated by DOE-BES (Office of Basic Energy Sciences) and instrumented by DOE-BER (Office of Biological and Environmental Research) and by National Institutes of Health (NIH), play a major role in this enterprise. Further stunning advances in structural biology require the availability of X-ray sources whose properties at least match those available today. DOE-BER has a major stewardship role in developing and providing access to synchrotron resources for the US user community. This role also has direct benefits to many elements of the DOE-BER science portfolio. **We recommend that DOE-BER maintain its stewardship role, make full use of synchrotron-based structural biology capabilities in its science portfolio, particularly in the emerging Genomes to Life program, and foster the continual upgrade of beamlines at existing synchrotron sources and, to the extent possible, the upgrade of sources themselves. These steps will enhance this positive trend and keep pace with advances in biology.**

2. Crystal diffraction experiments consume more synchrotron X-ray beam time than any other biological application, and will continue to do so. The practice of crystallography will be transformed by technologies for automation, currently under vigorous development around the world. Automated sample handling and also advanced computational control and analysis methodologies will increase throughput enormously, will reduce or eliminate the need for many experimenters to travel with their samples to the synchrotron laboratory, will reduce substantially the cost and time per experiment, and will facilitate advances in structural genomics. That is, the productivity of both the structural biologists and the synchrotron beamlines will be greatly enhanced. Adequate staffing for operations and R&D and resources for continuing improvements are essential to effectively meet the changing needs of the user community over the next 20

years. **We recommend that DOE-BER, in close coordination with NIH, continue its stewardship of biological crystallography at synchrotron sources. Given the continuing strong demand for crystal structures by the biological research community, special attention must be given to new opportunities for enhancing productivity through development of automation at beamlines.**

3. Experiments in X-ray spectroscopy and in small angle X-ray scattering/diffraction from non-crystalline samples require synchrotron radiation, but involve smaller user communities than crystallography. They are, however, an important complement to crystallography and at least current levels of activity should be effectively maintained. Liaisons and cooperative developments with scientific communities outside biology would be beneficial to biological practitioners of both methods: the environmental science community for X-ray spectroscopy experiments, and the materials science/polymer community for X-ray scattering experiments. **We recommend that DOE-BER maintain adequate resources for these critical fields of structural biology and explore opportunities to further strengthen them through interdisciplinary cooperation.**

4. The role of the DOE in supporting the development of a 4th generation X-ray source is both visionary and welcomed. It is unknown exactly how 4th generation sources might benefit biological research. The Subcommittee agrees with the conclusion of the 2002 BioSync Report that higher average brightness than available from existing 3rd generation storage rings is not likely to result in "proportional increases in biological throughput" and that on balance a relatively small number of today's experiments would benefit. However, 4th generation sources would directly benefit new experiments. Of the 4th generation sources discussed by the Subcommittee, the energy-recovery linac (ERL) is considered evolutionary while the X-ray free-electron laser (X-FEL) is revolutionary. R&D and pre-construction planning on X-FELs is significantly further advanced, and efforts in Germany and the US appear to be capable of delivering such sources within 6-10 years with a reasonably low degree of risk. **Given the early state of R&D of the ERLs, and the fact that X-FELs will have greater potential for breakthrough science by biological community, we recommend that a higher priority be placed on X-FEL construction.**

5. The planned 4th generation X-ray free-electron laser (X-FEL) sources were discussed in light of two experiments. First, the X-FEL should enable time-resolved studies of biochemical processes on the fastest biologically relevant time scales. Although the number of investigators pursuing such studies is likely to be small initially, the studies would be truly fundamental and the probability for success is high. Second, the X-FEL may be a productive source for imaging non-periodic materials. The potential importance of obtaining nanometer (nm)-scale structures for non-crystalline biological molecules is enormous, although it is unknown whether a single pulse from an X-FEL would produce a detectable scattering signal from a molecule such as a protein to a resolution exceeding that currently available from other imaging methods such as electron cryo-microscopy. **We recommend that DOE play a leading role in investigating the feasibility of the imaging experiment, and that DOE-BER encourage pursuit of both time-resolved and imaging experiments.** If the early imaging experiments show evidence of working, the substantial technical problems should be addressed systematically and then means

sought to evolve the method as a new approach for high-resolution imaging of non-crystalline materials.

6. Biological diffraction experiments on micro-crystals may be well suited to the energy-recovery linac (ERL) with its at least ten-fold higher average brightness than existing 3rd generation sources. The demand for this experiment is significant if potentially serious crystal damage issues can be resolved, and if the ERL proves stable and reliable. **We recommend that DOE-BER monitor ERL developments and be prepared to take advantage of this new source if appropriate in the future.**

7. Considerable scientific and cost benefit will continue to be gained by intra- and inter-agency, and inter-disciplinary, partnerships in the application of bright synchrotron X-ray sources to structural biology. **DOE-BER should strengthen its coordination with DOE-BES in its role as the steward of large scale, multidisciplinary research facilities. DOE-BER should also continue to strengthen its coordination with NIH National Center for Research Resources (NCRR) and National Institute of General Medical Sciences (NIGMS) in the effective development and operation of beamlines. We strongly recommend strategic planning by all involved agencies under the leadership of an organization such as Office of Science and Technology Policy (OSTP).**

Subcommittee members

Janet L. Smith, Chair (BERAC member)
Purdue University

Philip Anfinrud
National Institute of Diabetes & Digestive & Kidney Diseases – NIH

Jonathan Greer (BERAC member)
Abbott Laboratories

Keith Hodgson (Chair, BERAC)
Stanford University

Eaton Lattman (not present for discussion)
Johns Hopkins University

Keith Moffat
University of Chicago

James Penner-Hahn
University of Michigan

Tom Terwilliger
Los Alamos National Laboratory

Edgar Weckert
Hamburg Synchrotron Radiation Laboratory at
  Deutsches Elektronen-Synchrotron

X-Ray Crystallography

The Future of Macromolecular Crystallography at Synchrotron X-ray Sources

The growing impact of synchrotron radiation on macromolecular crystallography has been nothing short of stunning. Today the vast majority of new crystal structures deposited in the Protein Data Bank (PDB) were solved using data from synchrotron sources. This is in striking contrast to 10 years ago when only about one-quarter of new structures were solved with synchrotron data. The rapidly growing demand for crystal structures in the biological community and the increasing dependence on synchrotron radiation have severely strained the US ability to provide suitable synchrotron facilities for macromolecular crystallography. The past year has seen a large number of beamlines under construction and nearing completion at several synchrotron X-ray sources (Hodgson-Lattman Report). This development promises a range of new sites with macromolecular crystallography capabilities. Nevertheless, it is anticipated that the demand for synchrotron beamtime for macromolecular crystallography will continue to increase in the future as it has over the past years (Biosync, 2002) because the success rate of diffraction experiments using synchrotron radiation continues to improve. It is important therefore, in planning for the next 10-20 years, to consider how additional capacity for crystallography can be achieved.

The subcommittee considered three general areas to achieve increased efficiency and capacity.

1. Automation and more efficient beamline utilization. Over the past several years, a number of advances have been made in automation of beamline processes. These now need to be deployed, together with software improvements, to dramatically increase beamline efficiency, productivity and cost effectiveness, and to enable structural studies of increasingly challenging problems such as complex molecular machines. These include:

- *Crystal mounting robots*: Such robots mount and dismount crystals, maintain the crystals at liquid-nitrogen temperature, and align the crystals to the X-ray beam unassisted. Robots eliminate the down time necessary for manual loading of a new crystal on the goniometer. In crystal-screening mode, where typically only one or two diffraction images are recorded per sample, a dramatic reduction in human effort will be realized, along with an at least ten-fold improvement in throughput. High-throughput crystal screening is proving to be one of the most important tools in obtaining high-resolution structures for challenging problems such as integral membrane proteins and large, multiprotein complexes. For data collection, where full datasets are now being collected in 15-60 minutes, robots can provide a several-fold increase in throughput. Commercial robots to mount and dismount crystals have become available recently, but there is not yet a consensus on optimal design (or designs) for a variety of different beamline configurations. More than any other single improvement, automated sample handling is expected to increase beamline efficiency.
- *Automated data collection and processing*: Automatic recording of crystallographic data is needed to keep up with automated crystal mounting. This requires software that can detect satisfactory crystals automatically, interpret their diffraction patterns, develop an efficient experimental strategy, and execute the diffraction experiment. Afterwards, the

data need to be reduced to amplitudes automatically to be available in a timely fashion after the experiment. Hardware automation and sophisticated control systems that interface the processing software to the experiment are also required.

- *Capability for remote access and control of experiments*: The capability should exist for remote viewing of diffraction experiments under the overall supervision of beamline personnel or even for remote control by the investigators. Software to accomplish this needs to be made more robust and more generally available.
- *Common user interface*: A common user interface and output data format are essential to achieve most effective use of synchrotron X-radiation, by both on-site and remote-access users. This would allow users to take advantage of a number of diverse beamlines as time is available, with minimal training in environments that "appear" identical. The programs that actually operate the beamlines and robotics, process data, etc., would vary, but differences would be transparent to the user because of the common user interface and output format.

The importance of streamlined, robust operations is evident in the improved productivity of beamlines where staff resources have been devoted specifically to this task. In considering the above changes, the Subcommittee also considered whether efficiency of operation will dictate that specialized beamlines be developed to focus on one particular type of service and deliver that service most effectively. Thus, one might conceive of a beamline focusing on Multiwavelength Anomalous Dispersion (MAD) data collection, another for routine, robot-driven, single-wavelength data collection as well as crystal screening, another for specialized, difficult projects that require devoted attention. In an environment where all the beamlines are run centrally, this may be achievable. However, in a diverse Participating Research team/Collaborative Access Team (PRT/CAT)-based environment, it is not clear that such beamline specialization is feasible or desirable. A common solution is to schedule experiments by type so that equipment and optical elements are moved less frequently. Another possibility is to automate the optics and equipment necessary to switch between different types of experiments more quickly and straightforwardly. The subcommittee felt that improved automation of beamline optics is an excellent idea that will improve stability and reliability in general, but that other decisions in this area are best left to the individual facilities to pursue using some of the alternatives described above.

2. Evolution of user interaction with the beamline. The introduction of automation will significantly change the way the typical user will interact with the beamline. For standard crystal screening or data collection, the need for the user to come to the facility will be greatly decreased. The user will be able to send frozen samples to synchrotron facilities by express courier, and technical staff will load the cassettes that have been pre-loaded with multiple individual samples onto mounting robots for execution of the desired experiment: crystal screening, single-energy data collection, MAD data collection, etc. In many ways, this operation could run conceptually the same way DNA is sequenced today – as an outside service that accepts samples and delivers the desired products some days later. This approach in effect eliminates the need for an on-site user in cases where samples are frozen.

In an alternative operating mode, the capability for a remote user to follow and even to control experiments would be highly desirable in many cases, for example, when screening for good

crystals, or working with new crystals of proteins or protein complexes where the experiments are not expected to be routine.

These considerations will result in an evolution in the types of personnel necessary to staff the beamlines in the coming years. New areas of staff specialization will be:

- Engineering: staff familiar with the physics and optics necessary to design, trouble-shoot and operate a beamline, and also familiar with the range of robotics and automation that will characterize beamlines in the future.
- Technical: personnel trained to operate beamline equipment and robotics, but not as expert in crystallography or beamline physics and optics as the engineering staff. These people would run the day to day operations and maintain high throughput.
- Crystallography: personnel who will solve structures for those users who have no experience with crystallography, and who will provide expert advice on experimental design for other experiments. These individuals will also be an invaluable resource to help the more experienced users on their cutting-edge, very difficult crystallographic problems.

The emphasis will shift away from the existing situation of beamline personnel who must service a broad range of users on site with the attendant safety issues, the risks to the complex equipment that is run by less experienced users, and the "diplomatic" attitude and interpersonal skills necessary to service a diverse group of scientists. Instead, the above categories of personnel will be less focused on personal service, but are more closely devoted to achieving the users' and the beamlines' needs most effectively and efficiently.

3. Common user application process and rapid access to beamlines. The nature of biological research usually requires rapid access to beam time. Whether for new structures or for modifications to known structures, in most cases leading-edge science is very time sensitive. Waiting times of months to a year are not acceptable. With the availability of a larger number of macromolecular crystallography beamlines, more uniformity in the user interface and data output, and the predicted trend towards "remote" users, it becomes more important that experiments are done in a timely fashion than that they are performed at a specific beamline. Consequently, we recommend the development of a simple, uniform process for requesting crystallography beam time across the DOE complex of synchrotrons. Such a scheme should allow applications to be acted on quickly and time assigned to the user in a matter of days (especially for crystal screening) to a couple of weeks at most. Such a scheme can succeed only with a focused effort in developing new web-based tools, only with associated database and management staff, and only if the demand for beam time does not exceed supply. It would also require additional administrative support in user operations and scheduling at each of the synchrotrons due the need for a much greater degree of coordination and a much faster turnaround time.

X-ray Crystallography and the DOE-BER Genomes to Life Program

The Genomes to Life "facilities roadmap" proposes a facility for characterization and imaging of molecular machines (Facility #3). This facility is to focus on structural and biophysical analysis of macromolecular complexes ("the machines of life"). Several major synchrotron-based

techniques should be made key, and very likely would become, central elements of this facility. X-ray crystallography, which is responsible for many of the dramatic breakthroughs in understanding molecular machines in the past several years, must be a core component of any such facility. The atomic or near-atomic resolution structures of the RNA polymerase-II complex, of the ribosome and its subunits, of the F1 ATPase and of the proteasome all show the remarkable power of macromolecular crystallography in providing a structural foundation for understanding the biophysics of very large molecular machines. A full understanding of the properties and function of molecular machines would be greatly enhanced with high-resolution structural information on more of these complexes. In particular, the molecular modeling and simulation of the behavior of these machines will be put on a substantially improved basis with a known time-averaged three-dimensional structure. For example, high-resolution modeling of 70S ribosomal structure and activity is made much more practical with high-resolution crystal structures of 30S and 50S subunits.

Role of DOE-BER in Structural Genomics Initiatives

The DOE should continue to play a crucial role in structural genomics efforts both in the overall facilities and in beamline technology development and implementation. The synchrotron facilities themselves require continued development and support of capabilities for improved operational efficiency, flux and brightness, and positional stability. The further development of technologies for beamline instrumentation, including automation and detectors, and optimized data collection is critical. Structural genomics efforts over the next few years also will require standardized approaches to evaluate crystal quality, to decide on data collection strategies, to collect data, to evaluate data quality, and to carry out initial structure determinations. These steps require further advances in both data collection capabilities, and in decision-making and analysis of crystal and data quality. Further, the DOE can actively encourage cross-beamline compatibility of both sample-handling devices and output data formats. Meeting these needs will benefit all of biological crystallography, not only the structural genomics projects. DOE should also cooperate with NIH to seek to leverage the NIH investment in structural genomics in ways which directly benefit its own mission, including the Genomes-to-Life initiative.

Collaborative Implementation of Standardized Technologies

As many users of macromolecular X-ray synchrotron beamlines simultaneously use multiple beamlines, inter-operability and compatibility are highly important. This is even more important for structural genomics efforts. Support for the dissemination of software and hardware that promote compatibility among beamlines would be extremely beneficial. This could include robotic systems for sample mounting, user interfaces, and data processing. For example, BLU-ICE is a very useful and inter-operable beamline operation system and user control interface, but installation at new sites requires highly specialized knowledge that is not universally available and is expensive. If support for installation and maintenance of such software were available in a centralized manner, then many sites might use it, improving the compatibility of these sites both within a given site and among sites themselves.

Detectors

Two-dimensional X-ray detectors are a mission-critical part of any crystallography beamline, and when they fail, the beamline may have little recourse but to shut down until the detector is repaired, which can easily require a month's time. These detectors are too costly to stock spares for each beamline ($700K - $1.3M), and due to the range of detectors in use at each synchrotron facility, it is costly to maintain a detector pool suitable for all beamlines unless there is commonality across the set of beamlines. This issue has been raised in earlier BERAC reports, but a uniform solution is not yet in hand. Given the aggregate purchasing power of the DOE-BER and its interagency partners (NSF, NIH and others) and the large number of detectors in use on crystallography beamlines at synchrotron sources, the granting agencies are in a strong position to negotiate new contracts with detector suppliers that shift the burden for maintaining a detector pool from the beamline or synchrotron facility to the detector supplier. One mechanism might be a repair-exchange program, where a new/reconditioned detector that meets or exceeds the specifications of the failed detector is express shipped within 24 hours to the site, after which the failed detector is returned to the supplier *via* ground transportation. The cost of underwriting this service would be added to the purchase price, but when amortized over the useful life of the detector (approximately 5 years), the overall cost to the granting agencies should be significantly less than the cost of maintaining detector pools at multiple sites. Even if the direct cost savings were minimal, this plan would minimize down time from detector failures, whose projected costs are far greater. There is also a need to continue to develop even more advanced detectors for crystallography (and other) synchrotron experiments. Given the importance of this endeavor, and the relatively small amount of resources currently available for such programs, BER should consider working with BES and other interested agencies like NIH to foster new efforts in this direction.

Status of Synchrotron Beams for Crystallography

Improvements in beam reliability and in angular and positional stability and the trend toward lower emittance at all synchrotron sources have been highly beneficial to macromolecular crystallographic experiments. The increased brightness of 3rd generation sources has been a boon to macromolecular crystallography. Special improvements at some facilities, such as the "top-up" operations mode at the Advanced Photon Source (APS), provide especially stable operations. The Subcommittee agrees with the conclusion of the 2002 BioSync Report that higher average brightness than currently available from existing 3rd generation storage rings is not likely to result in "proportional increases in biological throughput" and that on balance a relatively small number of today's problems would benefit. Indeed, many of the experiments done today take advantage of the brightness offered by insertion devices on 3rd generation sources in ways other than simply delivering more photons to the sample; for example, by exploiting the low divergence inherent to bright beams for small samples, large unit cells or high energy resolution. As structural biologists tackle ever more challenging biological problems, it is important that efforts to provide greater excellence in synchrotron beams continue apace.

X-ray Absorption Spectroscopy

Over the last decade, the number of beam lines available for biological X-ray absorption spectroscopy (XAS) has decreased, mostly as beam lines have been converted to use for crystallography due to the overwhelming demand for new biological crystal structures. Although this has been offset somewhat by increased capabilities (higher flux and brightness, better detectors), the increase in sample throughput has not kept pace with the increase in demand. Consequently, it is more difficult to obtain synchrotron access for biological XAS today than it was 5 years ago. This is hampering the efforts of young investigators to enter the field and beginning to limit opportunities for technological and scientific innovation.

The applications of XAS have increased significantly, driven by the discovery of new metalloproteins, by the extension of XAS to ligand edges, and especially by the ability to study ever more dilute samples. However, the growth in XAS has not kept pace with the explosive growth of X-ray diffraction, and as a consequence, XAS represents a significantly smaller fraction of structural biology than in the past. Within this perspective, however, XAS provides direct information on oxidation state and electronic structural properties and uniquely accurate bond-lengths in comparison with those typically obtained by macromolecular crystallography. Without the information provided by XAS, it would be impossible to interpret some of the features of protein crystal structures or to understand the details of the chemistry of metalloprotein active sites. It is therefore critical that future synchrotron planning by DOE-BER, in coordination with NIH, prevent further erosion of the number of beam lines, and support of those beamlines, for XAS.

Spatially and Temporally Resolved XAS

Looking to the future of XAS vis á vis new synchrotron sources, the major impact of the high brightness of 3rd generation sources is to allow measurements to be made on much smaller samples. This is an enabling technology for spatially and temporally resolved measurements. X-ray microprobe spectroscopy can be used to obtain spatially-resolved metal speciation, extremely valuable information both for biology and for environmental studies. We encourage DOE-BER, again in appropriate partnership with other interested agencies, to leverage this common interest of the environmental and biological communities to build a critical mass of scientists that can facilitate the construction, instrumentation, and effective operation of state-of-the-art X-ray microprobe beam lines. Temporally resolved XAS is likewise facilitated by high-brightness beam lines because stopped-flow instrumentation uses small sample volumes to give rapid mixing and to reduce sample consumption. Time-resolved measurements are having an important impact on studies of catalysis and inorganic chemistry, but are just beginning to impact biological studies. The addition of spatial and temporal resolution will increase the demand for XAS beam time. It is important that this demand be met by new beamline construction, and not by replacing existing static XAS beam lines, which remain in high demand, with microprobe beam lines.

Fourth generation synchrotron sources are unlikely to have a large initial impact on XAS studies of biological samples. Higher average brightness, coupled with more idealized beam shapes,

could enable higher-brightness microfocussed beams. In the time domain, the same capabilities that permit femtosecond time-resolved X-ray diffraction can also be used to make femtosecond time-resolved XAS measurements. However, it is likely that, at least initially, femtosecond time-resolved XAS will be limited to inorganic samples rather than more dilute biological samples. These experiments, which can be carried out with an X-FEL, will be especially exciting if pulses in the few fsec regime are obtained. They can be expected to yield fundamental insights into the structural dynamics of chemical reactions, some of which, e.g., light-induced processes, are of central importance to biology.

Small Angle X-ray Scattering/Diffraction

For materials without periodic order, e.g., biomolecules in solution or fibers such as found in muscle, small angle X-ray scattering/diffraction (SAXS/D) is a method that provides valuable structural information including average size (radius of gyration), molecular weight and volume, and a vector distribution function (analogous to a "Patterson" map for a crystalline sample). Such information is obtained by careful measurements made at low scattering angles, typically much smaller than recorded for diffraction patterns from most crystalline biomolecules. As a consequence, the resolution available is not at the atomic level as obtained from crystallography. However, SAXS/D measurements can be carried out in a time-resolved manner (currently with time resolution in the few millisecond or slightly shorter time regime). SAXS/D provides very useful information on large and complex biomolecules, for example, the shapes of molecular complexes, and how shapes change during reactions under physiological conditions, and the nature of the "folded" or "unfolded" states in solution. Consequently, SAXS/D studies are finding wide applications in areas as diverse as the fundamental biophysics of protein folding to investigating the molecular basis of the "protein misfolding" diseases, such as mad cow disease and Alzheimer's.

Over the past few years, synchrotron radiation has also made it possible to measure scattering curves that extend well beyond the traditional small angle or Guinier region. Such curves can provide much more information about the specimen than does traditional small-angle analysis, including, at least in principle, the low-resolution structures of particles like icosahedral viruses.

The importance of low-order Bragg reflections in producing accurate electron density maps has been under-appreciated until recently. Such reflections are difficult to measure, especially from large unit cells, with conventional crystal data collection apparatus, but can be readily measured using a slightly modified small angle instrument on a synchrotron. Demand for these sorts of crystal measurements at small angle beam lines could create additional demand and make important contributions to the structural biology of large macromolecular assemblies.

Due to the very low scattering signal relative to background and the sensitivity to parasitic scatter, synchrotron X-radiation indeed is the primary enabling ingredient for biological small-angle scattering studies. Synchrotron X-rays have proven to be essential for biological time-resolved studies. The SAXS/D scientific community is relatively small, but we agree with the conclusions in the 2002 BioSync Report that there is significant potential for growth and a much broader application of this technique in a manner complementary to crystallographic studies. In addition, we agree that the temptation to convert additional beam lines, such as those used for SAXS/D, into crystallography stations should be strongly resisted. SAXS/D experiments are challenging and the beamlines available to carry them out are not "turn key". Hence providing adequate skilled scientific and technical staff at the beamlines to further their development, while supporting outside user groups, is very important. We urge DOE-BER to continue to recognize the importance of this area and seek, in partnership with NIH, to effectively staff and maintain/upgrade the existing beam lines. As the field evolves, additional capacity may be needed in the future, but this should be well established and documented before building additional beam lines. We also observe that it may be possible to generate synergy with other

fields where SAXS/D plays an important role, most notably in materials and polymers. Joint efforts to develop new instrumentation and software could benefit both communities.

SAXS/D experiments can in principle access even faster time scales (10s to 100s of microseconds) given suitable approaches to sample handling (e.g., faster mixing systems, more rapid means of fast initiation of the event to be studied), and progress is already being made in this area. Radiation damage can be managed in some cases by approaches like rapid flowing of sample solutions. Increased average brightness that could be provided by ERLs or X-FELs could play a role in pushing this time domain. However, it should be noted that other bottlenecks such as adequate detectors and sample handling technologies are currently limiting and need additional R&D. With the extreme brightness of an X-FEL, it is possible to imagine recording a two-dimensional scattering pattern on the ultrafast time scale provided by the pulses of such an accelerator (in an experiment analogous to that of single-molecule imaging discussed elsewhere in this document). If significant challenges with sample manipulation and background scatter can be solved, it may be feasible to study structural changes at low resolution on a much faster time scale than the current limit in the microsecond to millisecond regime. Such experiments could well be tried in the early phases of operation of an X-FEL.

SAXS/D and the DOE-BER Genomes to Life Initiative.

One of the primary goals of the Genomes to Life initiative is to characterize and understand the nature of the complex machinery of life ("molecular machines"). SAXS/D can be used to study low resolution structures of such large and complex assemblies in conditions that are at or near those normally found *in vivo*. In the simplest sense, such low-resolution information can be used to estimate phases for crystallographic data from complexes that can be crystallized, which with iteration can lead to high-resolution structures. Such approaches are already proving valuable for low-resolution diffraction data from large molecules like viruses. More importantly, low-resolution information can be used to orient subunit structures obtained from high resolution methods such as crystallography and NMR, and can be used to study changes in structure as a function of time. Hence, we believe that SAXS/D has an important and complementary role to play in the emerging Genomes to Life science portfolio, and this should be recognized in the programmatic planning for this new initiative and the facilities that will support it.

SAXS/D and the Structural Genomics Initiatives.

The NIH-funded efforts in structural genomics depend primarily upon high-resolution structural information available from synchrotron-based X-ray crystallography. SAXS/D could play a role in providing complementary information that is useful in optimizing the structure determination pipelines. For example, SAXS could be used to screen crystallization conditions, providing information on aggregation states of the protein. Rapidly obtained lower-resolution information could, together with sequence and other information, lead to more reliable computational predictions of structure.

Structural Biology and Fourth Generation Synchrotron X-radiation Sources

Comparison of Source Properties of Third and Fourth Generation Sources

Virtually all synchrotron radiation used today is provided by storage rings. Linear accelerators (linacs) can also be used to produce synchrotron radiation by passing electron beams through insertion devices but, to date, designs capable of high brightness and economical operation have not been realized. Recent developments have given rise to a series of proposals around the world for two new classes of accelerators called Energy Recovery Linacs (ERLs) and X-ray free electron lasers (X-FELs). These devices will represent the next, fourth generation of synchrotron X-radiation sources. Third generation sources are compared to ERL and X-FEL sources in the Table below. All numbers are approximate.

Properties of Current 3rd-Generation and Proposed Next-Generation X-ray Sources

	3rd generation	ERL	X-FEL
Pulse length (sec)	~10^{-10}	~10^{-13}	~10^{-13}
Peak brightness (ph/(s•mrad^2•mm^2•0.1% bandwidth))	10^{23}-10^{25}	10^{25}-10^{26}	10^{32}-10^{33}
Average brightness (ph/(s•mrad^2•mm^2•0.1% bandwidth))	10^{19}-10^{21}	10^{22}	10^{22}-10^{25}
Beam cross-section	~10:1 elliptical	~circular	~circular
Coherence fraction	<0.1%	0.3-15%	~100%
Linewidth (ΔE/E)	1-2%	0.02-1%	0.1%
Maximum X-ray photon energy (keV)	>50	>50	~15

ERL sources. Compared to current 3rd generation synchrotron radiation sources, ERLs have the potential to deliver higher average brightness in the X-ray region of the spectrum. In addition, ERLs can produce very short pulses because they do not have the constraints imposed by an equilibrium system like a storage ring. The unique uses of ERLs are likely to be found by exploiting these short pulses. The emittance of an ERL is determined mainly by the performance of the injection system and by the accelerating linac, and should be nearly identical in the horizontal and vertical directions. Thus, the photon beam source will be almost spherical, compared to the elliptical source points of storage rings. Such source properties are advantageous for micro-focus applications and will result in a higher fraction of coherence than current 3rd generation storage ring sources. ERL undulator sources will be able to provide a narrower linewidth than storage ring undulator sources because the electron energy spread of the driving linac is in general smaller than that of a storage ring. ERLs are currently in a pre-design phase, and several proposals exist for R&D and prototypes. Building an ERL to operate with a high degree of reliability and efficiency in the X-ray region would require overcoming significant technical challenges in areas such as gun design and power dissipation in the cavities. Given suitably funded and focused R&D, these challenges should not be insurmountable.

X-FEL sources. Keeping in mind that 3rd generation sources provide sufficient flux and average brightness for most present applications in structural biology, the advantages of ERLs will be in the better time resolution and probably in the application of the larger coherent fraction. However, in these two areas, ERLs will be far less powerful than X-FELs due to the limited number of photons per ERL pulse. In an X-FEL, extremely short electron bunches of low emittance will be created by a laser-driven injection system, accelerated in a linear accelerator, further compressed by suitable magnetic chicanes to a length of ~50 µm, and fed into a very long undulator. At the beginning of the undulator, normal spontaneous undulator radiation will be generated. Due to their perfect alignment, the generated photon beam and the electron beam will interact, causing each electron bunch to be divided into microbunches of size similar to the wavelength of the photon beam. The tremendous increase in brilliance of X-FELs will derive from the coherent scattering of all electrons within each microbunch. Due to the process of "self-amplified spontaneous emission" (SASE), each bunch of an X-FEL will provide high brightness in a narrow bandpass. The maximum repetition rates available for single-pulse experiments will be 120 Hz and 10 Hz, respectively, for the Linac Coherent Light Source (LCLS) and TESLA X-FELs. In principle it is possible to distribute the electron pulses or part of a pulse train onto several undulators at the expense of the overall repetition rate. It is also possible to use the electron beam after it has passed through one SASE undulator for a second one with a lower resonance energy. However, even under this conditions only a limited number of SASE undulators can be driven by a single linac. In addition to extremely high peak brightness, X-FELs will provide a beam that is almost totally coherent in the transverse direction. This property will require special care for optics and beamline design but provides tremendous scientific potential. The specific benefits of coherence to biological applications have not been determined.

Ultrafast Structural Processes

Use by the biological community of the X-FEL source brightness and spatial coherence is discussed later in this report. The further key property of X-FELs, their ultrashort pulse duration, would have immediate applications to ultrafast time-resolved macromolecular crystallography. The shortest time resolution achieved in this nascent field is now 150 psec using synchrotron X-rays from 3rd generation sources, limited by the X-ray pulse duration, which in turn is limited by the electron bunch length. In a recent study (F. Schotte, *et al*. (2003) *Science*, in press), small conformational changes were found to propagate through a protein structure in less than 100 psec, and therefore remain temporally unresolved. Ultrafast spectroscopic data suggest that structural changes occur on this time scale in a number of important, light-sensitive biological systems such as the photosynthetic reaction center, light-harvesting complexes, and photoreceptors such as rhodopsin, bacteriorhodopsin, photoactive yellow protein and LOV domains (Z. Ren *et al*. (2001) *Biochem*. **40**, 13788-13801; S. Crosson, S. Rajagopal & K. Moffat (2003) *Biochem*. **42**, 2-10). Indeed, the genome of the plant *Arabidopsis thaliana* encodes at least 58 proteins that respond to light, including nine that are homologous to known signal transduction proteins (The Arabidopsis Genome Initiative (2000) *Nature* **408**, 796-815).

X-FELs would extend the time resolution of macromolecular crystallography more than three orders of magnitude into the chemical time scale (~100 fsec and in later developments, to ~10 fsec). With this time resolution, protein structural changes associated with a photochemical event could be witnessed as forces develop from the epicenter and propagate throughout the protein, making it possible to "watch" the fastest events in protein functions. Such studies would provide essential data for understanding the dynamics that underlie chemical processes in proteins. That is, the fundamental structural principles that govern such processes as bond breaking, bond making, isomerization, and rapid conformational change do not depend on whether the reactions being studied are driven by a physical process such as absorption of light, or by a chemical process such as binding of ligand.

Imaging of non translation periodic objects with XFEL radiation

The transverse coherence of X-FELs, coupled with their very high peak brightness and short pulse length, permits, at least in principle, the direct imaging of non-crystalline (and hence non-periodic) biological molecules. The physical basis for such measurements lies in the fact that a molecule, when illuminated with X-rays having a coherence length comparable to the distances between scattering atoms, will give rise to a diffraction pattern. Such patterns are characterized by a continuous variation of intensity and not by the intense spikes seen from crystals (Bragg peaks) that result from the amplification due to the crystal lattice. In fact, Bragg diffraction actually "samples" the continuous molecular transform (sometimes also called a speckle pattern) at a discrete frequency. In principle, the full, high-resolution, 3-D structure can be obtained from a series of such two-dimensional molecular transforms, if they can be recorded to high enough resolution and with sufficient signal to noise. This approach is somewhat analogous to the methodology that has revolutionized the practice of electron cryo-microscopy but for which there are still practical limits on resolution (4-5 Å at present, although experts predict that 3-4 Å will be attainable). The molecular transform also offers another advantage in that it can be sampled at a higher frequency than can Bragg diffraction. It has been shown that an iterative numerical approach together with "oversampling" can lead directly to solving the classic phase problem in crystallography, i.e., provide a means to obtain the phases which, together with the intensities, are needed to produce an image of the structure (J. Miao, J. Kirz & D. Sayre (2000) *Acta Crystallogr. D* **56**, 1312-1315; J. Miao, K. O. Hodgson & D. Sayre (2001) *Proc. Nat. Acad. Sci. USA* **98**, 6641–6645). Various aspects and limitations of these concepts are elaborated in the sections below.

For objects that can be made in reproducible conformations and in large quantities, X-FEL radiation might provide a highly novel approach. In this case, the sample must remain unperturbed only for the duration of one X-ray pulse. This allows exposure of a sample to a much higher dose than would be possible in a normal diffraction experiment, provided the pulse is short enough. It may be possible with currently available optics to focus a very short X-FEL pulse (~100 fsec) sufficiently for the scattering signal from a microscopic object to be recorded on an appropriate two-dimensional detector, provided the detector is free of any noise.

Since the samples are assumed to be identical, the scattering pattern of a large number of randomly oriented, individual particles could be used to sample all of reciprocal space. However, the scattering image of a single sample must contain sufficient information to be distinguished from images of other samples recorded at slightly different orientations. The problem of determining the orientation of each particle from its scattering images is conceptually similar to the one in single-particle reconstruction by electron cryo-microscopy. Once the scattering contribution in reciprocal space is obtained, the reconstruction of the object should be straightforward due to the oversampled nature of the scattering information.

The maximum flux density a sample can withstand depends on the photon energy and the duration of the X-ray pulse (R. Neutze *et al.* (2000) *Nature*, **408**, 752–757). The main inelastic damage effect is photo-absorption, leading to atoms with an electron hole mainly in the K-shells (hollow atoms) and a photo-electron. For very small samples, the photo-electron will leave the sample without further interaction. The primary decay process for K-shell holes of light atoms generates Auger electrons, which must be assumed to interact inelastically with their environment, generating additional ions. Since the overall photo-absorption cross section drops with increasing photon energy, higher X-ray energies are preferrable to minimize this effect. Shorter X-ray pulses also reduce the effects of inelastically produced ions. The shorter the pulse the higher is the sustainable pulse flux density since the inertial moments of the atoms slow down the Coulomb explosion which follows the generation of ions due to inelastic processes.

The scattering experiment must be carried out in vacuum in order to keep the background as low as possible. State of the art electrospray techniques could be used, as in mass spectrometry, to generate a jet of biomolecules that will intersect the photon beam. More elaborate techniques to position the molecules such as optical tweezers could also be investigated. It is obvious that a number of other technical developments will be necessary for this experimental technique to be successful in practice.

The smallest sample that can be investigated using this approach is still under theoretical investigation. In general, the larger the object, the greater the number of photons that will be scattered, containing more orientation information. From the present, preliminary knowledge and the parameters of the first X-FELs (LCLS and TESLA-XFEL with ~100-fsec pulses of more than 10^{12} photons per pulse), the smallest particle size that seems feasible to investigate at high resolution is approximately the size of a large virus (~100 nm). Significantly smaller objects could be accessible if the pulses from future X-FELs were as short as 10 fsec, or even 1 fsec. As mentioned above, higher photon energies compared to the envisaged 8-12 keV could also be beneficial. A larger number of photons per pulse would be of practical importance by allowing less stringent focusing conditions for these experiments. Such improvements could conceivably bring this experiment into the size regime where it is possible to study large macromolecular assemblies. As also noted earlier, it is in principle possible to directly solve the phase problem for such oversampled diffraction patterns, and hence to directly reconstruct the real electron density (Miao *et al*. (2001) *Proc. Nat. Acad. Sci. USA* **98**, 6641–6645).

If this experiment can be realized, the achievable resolution is determined mainly by two factors: (i) the degree of conformational reproducibility among individual particles and (ii) the accuracy of determining the orientation of an individual particle from its diffraction image. If such

methods can be used to study biological molecules, and if the formidable sample orientation problem can be solved (the scattering patterns of each of several thousand molecules must be oriented and aligned in order to reconstruct the electron density), coherent single-molecule x-ray scattering could have a major impact in biology through its ability to image large, complex and non-crystalline biomolecules at or near atomic resolution.

Microcrystal Diffraction

Even if the bold imaging proposal fails, the sample-handling techniques developed should be amenable to structural studies using extremely small crystals, too small to be examined by today's 3rd generation sources. The Subcommittee considered how diffraction images might be obtained from micro- or nanocrystals, although a single X-FEL pulse would destroy each microcrystalline sample. Consider, for example, a spray of microscopic water droplets, each arranged to contain a single microcrystal but minimum water, traversing the tightly-collimated focal spot from an X-FEL. Each microcrystal would be in a random orientation with respect to the beam (although hydrodynamic, electrostatic or magnetic forces might be used to at least partially orient the crystal prior to reaching the focal spot). As the microcrystal traverses the focal spot it would give rise to a monochromatic "still" diffraction image. The volume of reciprocal space sampled in one "still" image would depend on the angular convergence and energy dispersion of the X-ray beam at the crystal. Identifying that volume in the reciprocal lattice would depend on the ability to index the diffraction pattern. This does not require the highest resolution data; it may be sufficient to index a small number of the stronger, low-order reflections in a prominent diffraction zone. Successful indexing would allow assignment of each pixel in the "still" image, even those at high resolution, to a location in reciprocal space. A large series of such "still" images ultimately would span all of reciprocal space, and may allow the necessary signal-to-noise in the high-resolution diffraction pattern to be built up.

Indeed, the intensity from a single, focussed X-FEL shot on a micrometer (μm)-size lysozyme crystal is estimated to provide significant intensity to about 2-Å spacings. This estimate assumes a flux density for which 1-10% of all atoms will be ionized during the X-ray pulse, effectively destroying the sample. Further assumptions were (i) ideal detector efficiency, (ii) background contributions that derive only from the sample itself, and (iii) microcrystals of similar size. Although the necessary experimental techniques and data evaluation procedures for these experiments have not been developed, the potential of microcrystal diffraction using X-FEL sources is worth serious consideration since it is frequently easier to obtain microcrystals than ones suitable for more conventional crystallographic approaches.

A similar approach may be feasible with two-dimensional microcrystals. Many membrane proteins assemble into two-dimensional patches but are recalcitrant to growth of three-dimensional crystals. Provided that many two-dimensional crystals of approximately the same size can be obtained, a data set could in principle be collected from a larger number of exposures from different samples. The minimal sample size required for obtaining significant intensity from a two-dimensional micro crystal is (10 μm)^2 for lysozyme-size molecules and diffraction data to about 2.5-Å spacings.

Holography

Another example is Fourier-type holographic imaging of nanometer-scale biological objects using a quasi-spherical reference wave. In this case, the achievable resolution is the resolution to which the reference scatterer is known, and the X-FEL beam would need to be sufficiently monochromatic that the longitudinal coherence matched the size of the object. Such a technique could be useful for whole cells or organelles. More than one X-FEL pulse would be needed to record the entire hologram and, since none of these samples can be produced in identical copies, the achievable resolution will be limited by radiation damage. If such a hologram can be recorded, it should be straightforward to reconstruct the object.

Summary

Because the X-ray synchrotron sources currently serving the biological community are in practice spatially incoherent for the kinds of experiments being done today, the unique opportunities afforded by an intense, spatially coherent X-FEL source are not yet appreciated, nor well understood. Nevertheless, novel ideas that take advantage of spatial coherence will undoubtedly emerge and may have an impact in the biological sciences. The possibility that high-resolution structural information could be acquired from non-crystalline nanoscale particles with a single fsec X-ray exposure is quite intriguing, and must be pursued. Many obstacles remain, and it is not yet clear if coherent single-molecule X-ray scattering is realistic, particularly for proteins. However, the potential breakthrough importance of measurements that produce nm-scale structures for non-crystalline biological molecules is so great that this experimental approach needs to be supported and explored. While such biological experiments do not on their own provide a compelling reason to build an XFEL, they do add to the body of compelling science that justifies construction of such a machine. Fortunately, XFELs appear well on their way to being built in the US and in Europe. In this context, it is important that resources be provided for R&D, at comparably small marginal cost, to explore the extent to which coherent single-molecule X-ray scattering can be used to image biological molecules.