

REPORT  
of the  
STRUCTURAL BIOLOGY SUBCOMMITTEE  
of the  
BIOLOGICAL AND ENVIRONMENTAL RESEARCH ADVISORY  
COMMITTEE

In response to the charge letter of Dr. Martha Krebs, June 10, 1997

Executive Summary

Six years have elapsed since the previous report of the Structural Biology Subcommittee. Dramatic progress and advances in the field over this period led to the request to the Subcommittee to review the needs of structural biology and to advise the Office of Biological and Environmental Research (BER) on the portfolio of facilities and research grants currently being funded in this area. Two panels were convened: one to review the specific area of macromolecular crystallographic use of synchrotrons, which produced a report in July 1998, and a second to deal with a broader range of structural biology technologies and discoveries. The latter Subcommittee met on August 16 and 17, 1998, to discuss these areas. The major issues and recommendations of the Subcommittee are as follows:

**Review of recommendations from the 1992 Structural Biology Subcommittee report:**

The critical issues dealt with in the 1992 report, many of which continue to be central today, are (evaluation of the current situation regarding each issue is in italics):

- Improving synchrotron beamlines to improve efficiency and operations.  
*Much has been done to improve the current beamlines and construct new beamlines. Continued optimization of these beamlines will be needed in the future due to greatly increased user demand for beam time.*
- The need to provide neutrons for structural biology.  
*This continues to be a critical issue today. However, it is the sense of the Subcommittee that the need has shifted from neutron crystallography to lower-resolution neutron scattering.*
- Priorities were defined for development of the new APS and ALS synchrotron facilities. *The highest priority items (new beamlines for crystallography at APS, x-ray scattering at APS, and soft x-ray spectroscopy at ALS) are currently being funded by BER and other government agencies.*

- DOE, NSF, and NIH should play a role in the development of high field magnets for NMR. *Since 1992, 750 Mhz spectrometers have become common in NMR research groups. EMSL at PNNL is funded by BER to acquire a 900+ Mhz magnet for a user facility. Support should continue for development of Ghz class NMR spectrometers for the future.*
- DOE should fund computational and theoretical biology to help understand protein structure and function. *A research program in computational biology was initiated by BER and further funding focused on the area of structural genomics is recommended.*
- Coordination between DOE, NIH, and NSF in funding and support for structural biology. *Such coordination has increased over the past several years especially with the current OSTP working group. This coordination and cooperation should be expanded in the future.*

**Overview of the recommendations of the July 1998 Structural Biology Subcommittee report on use of synchrotron facilities for macromolecular crystallography:**

This Subcommittee reviewed the report of the Synchrotron Subcommittee produced in July 1998 on the use of synchrotron radiation for macromolecular crystallography and fully endorsed its recommendations. In several areas, this report will expand on important issues that were touched upon in the July 1998 report.

**Detector, other instrumentation, and computational development for x-ray crystallography<sup>1</sup>:**

*X-ray Detector Development*

The development of Charge Coupled Device (CCD) detector systems, which was largely funded by the BER Instrumentation Program, has probably been one of the most important developments over the last half decade enabling the exponential growth in synchrotron structural biology experiments. With the advent of third generation light sources, new detectors are needed to take full advantage of new capabilities.

*A two-pronged approach is recommended: providing the best currently available detectors for existing beamlines and encouraging longer-term research programs for better detectors. To accomplish the latter, detector research support needs to be increased, balancing the funding across the most promising new detector technologies, and eventually funding the integration of these new detectors into the beamlines.*

---

<sup>1</sup> In this section and those that follow, the subcommittee recommendations are given in italics.

### *Other Instrument and Computational Development*

The greatly increased rate of data acquisition at synchrotron light sources, the increased user demand, and the rise in number of less-expert users is placing dramatically greater demands on the data handling and reduction capacity of beamlines.

*It is recommended that sufficient human and hardware resources be provided to beamlines to implement software for data evaluation, processing, and analysis for the high rates of data acquisition and for increased numbers of non-specialist users. Furthermore, BER should foster the availability of high-speed network services for users and encourage development of a common, detector-independent format for diffraction images.*

### *Automation of Sample Handling*

With the greatly increased demand for synchrotron beamlines, especially at crystallographic stations, dramatic increases in efficiency can be realized through the implementation of automation for sample loading and handling.

*A BER-directed initiative is needed for the development of beamline automation of the critical, rate limiting steps. Funding and encouragement for the development of improved software for data collection and data quality analysis is essential.*

**Structural biology and the evolving need for high field nuclear magnetic resonance (NMR), mass spectrometry (MS), neutron diffraction and scattering, and x-ray scattering and spectroscopy (XAS) facilities**

### *Nuclear Magnetic Resonance (NMR)*

NMR is playing an increasingly important role in the determination of protein structure and the study of protein dynamics. Traditionally, DOE and BER have had only a small role in NMR development. The two most critical issues for NMR in the future are the development of higher field magnets to improve resolution and sensitivity and the relatively long, rate-limiting time required for data collection.

*These two issues should be addressed by a strong DOE role in the aggressive pursuit of the best possible initiatives to develop new high field instruments and by the creation of regional facilities, with cross-agency funding, for housing high field instruments to accommodate a large number of NMR researchers.*

### *Mass Spectrometry*

The past ten years have seen major advances in mass spectrometry that have propelled this method to the forefront in protein characterization. Nevertheless, little is being done on the national level to take advantage of this technique for exciting new areas such as the Human Genome Project and structural genomics.

*BER should seek to formulate a national consensus for integration of non-proprietary, large scale protein identification/characterization by mass spectrometry with the continuing structural*

*genomic/bioinformatic and computational biological efforts. To continue the advances in this area, BER should:*

- *participate with other agencies in formulating a national consensus for integration of large scale protein identification/characterization by mass spectrometry with the efforts in structural genomics/bioinformatics and computational biology*
- *continue support of Bio-AMS (accelerator mass spectrometry) and the development of a 'small' machine dedicated to <sup>14</sup>C biomedical research measurements*
- *leverage investments in detector technology development for new ion-optical strategies in ultra sensitive biological mass spectrometry*
- *foster development of mass spectrometric instrumentation for study of non-covalent protein complexes and encourage integration of attomole and sub-attomole separation technologies and chip-based strategies with mass spectrometric identification instrumentation*

### *Neutron Crystallography*

Advances in high resolution x-ray crystallography at synchrotron light sources and the greatly increased versatility of high resolution NMR for studying protein structure and dynamics together with difficulties inherent in the practice of neutron crystallography and the highly uncertain nature of neutron sources over the past decade have caused the niche for neutron crystallography to grow smaller.

*The recommendations of the Subcommittee are:*

- *not to build a protein crystallography station at ORNL at HFIR; the beamline can much better serve the community in a small angle neutron scattering (SANS) configuration. The protein neutron crystallography station at Los Alamos, already in the design/construction phase, together with the ILL instrument, LADI, should provide the ability to evaluate the future scientific case for neutron crystallography. Future operational support for the Los Alamos instrument by BER must be evaluated in terms of its performance and the size of the user community it attracts. Consideration should be given to what range of problems, including condensed matter science, could also be served by this instrument. Evaluation of neutron protein crystallography for future planning must engage a broad range of structural biologists in addition to neutron experts including leaders in competing technologies such as synchrotron crystallography, NMR, and microscopy.*

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

Small-angle scattering of x-rays or neutrons from biological macromolecules in solution is playing an increasingly important role in structural molecular biology. The techniques yield information on molecular associations and overall shapes of biomolecules in solution. Small-angle scattering can be particularly powerful when combined with data from high resolution techniques for studying the dynamic interactions and conformational flexibility inherent in the regulated functioning of molecular assemblies.

*The recommendations of the Subcommittee are that OBER should:*

- *partner with DOE/BES to ensure optimal utilization of the new ORNL cold neutron source, in particular by funding a second small-angle instrument optimized for and dedicated to biology*
- *support technologies to decrease the cost and increase the ease for production of deuterium labeled components*
- *proactively build the U.S. small-angle scattering community, with an emphasis on structural biology applications by:*
  - *facilitating access to high-intensity, well instrumented cold neutron beamlines and synchrotron sources*
  - *supporting operations of existing small-angle beamlines at SSRL and APS (designed for biological applications), and the proposed small-angle instruments at ORNL:*
    - *support for new users to get into the field through beamline staffing increases*
    - *support development of standardized user-friendly data handling and modeling software packages that can run on most common computer platforms*
- *support a multi-agency workshop to evaluate the future trends of small-angle scattering in structural biology*
- *partner with NIST to evaluate reflectometry (small-angle scattering from thin films and surfaces) potential for structural biology*

#### *X-ray Absorption Spectroscopy*

X-ray absorption spectroscopy (XAS) is a non-crystallographic solution method requiring synchrotron radiation that is used to probe the structure of specific atomic constituents in complex materials. XAS experiments provide information about the electronic structure of the absorbing atom as well as metrical details about its coordination or near neighbor environment.

*The Subcommittee recommends:*

- *The current investment in synchrotron beamlines for XAS is adequate but may need to be expanded over the next 3–5 years if the user base increases*
- *Insufficient staffing of XAS beamlines remains a critical issue as with crystallography beamlines*
- *Implementation of high performance, high energy resolution solid state detectors in association with MAD crystallographic beamlines would enable combined XAS/crystallographic studies to be done routinely*
- *Further investments in the development of improved and new detector technologies will enhance capabilities for new experiments including whole cell and time-resolved XAS. Improvements in data processing and analysis theory and algorithms are also important*

#### **Review of the portfolio balance of the BER structural biology research grants**

The BER program provides critical support for the construction and ongoing operations of structural biology beamlines at the synchrotrons, at neutron sources, as well as for instrumentation development, computational biology, and biological research.

*The Subcommittee concluded that BER should:*

- *maintain at least its current critical level of investment in facility development and operations for both x-rays and neutrons*
- *focus research investments in areas that couple to the unique strengths that it has developed in association with the national user facilities and laboratories*
- *encourage research involving the development of novel tools that will enhance access and throughput of the shared, multi-user facilities*
- *expand the level of investment in advanced instrumentation research, especially for detectors*

## **BER program in computational biology and the relationship to structural genomics**

### *Structural Genomics*

The goal of structural genomics is to determine the three dimensional structures of a large number of proteins achieving as full a coverage of various genomes and as large a diversity of protein folds as possible using x-ray crystallography and NMR. These experimentally determined structures will be extended and supplemented with homology models of related proteins. This effort constitutes a massive data production effort that, in its most extensive form, contains parallels to the Human Genome Project.

*Because of the newness of this field and the current absence of a “national” effort to undertake a structural genomics effort, the Subcommittee recommends that a broadly based panel of scientists should be assembled to evaluate the importance, feasibility, and cost of a cross-agency federally-funded national structural genomics program.*

### *Computational Structural Biology*

Computational structural biology is a loosely defined field that, in its broadest definition, encompassing molecular dynamics, quantum mechanical studies of enzyme mechanisms, protein folding and reverse folding studies, and sequence analysis tools. The Subcommittee focused on computational areas that are most directly related to structural genomics:

*The emphasis of the current BER research grant portfolio in the general area is good and should be continued. Increases in funding, and the mechanisms by which this should occur, should come as part of larger scale projects that would be an integral component of any broadly-based initiative in structural genomics emerging from the workshop proposed above.*

## **Coordination of funding efforts between BER, BES, NIH, NSF, and NIST**

The need for interagency cooperation in structural biology arises because various research portfolios or initiatives have significant components or user bases funded by different government sources. This is especially the case for large, shared multiuse user facilities such as those providing synchrotron or neutron beams. Mechanisms for

interagency coordination include, among others, joint solicitation/funding of proposals, coordinated focused programs, and working groups at the interagency level.

*We strongly endorse the interaction within different divisions of the same agency and among different government agencies in areas where optimization of resource allocation and joint planning and development are appropriate. Such interactions should be fostered and encouraged.*

*The Subcommittee specifically recommends continuation/initiation of cooperative activities among the federal agencies in the following important areas:*

- *synchrotron beamline operations and construction*
- *x-ray and neutron detector development*
- *high-speed network links connecting synchrotron beamlines and research institutions*
- *high field magnet development for NMR and NMR user facilities*
- *mass spectrometer instrument development*
- *neutron scattering instrumentation and beamlines*
- *structural genomics including its computational structural biology component*

*Five specific areas where joint agency planning is important are:*

- *the new beamlines on the cold neutron source at HFIR*
- *the upgrade of SSRL (SPEAR3)*
- *adequate operating support for CHESS*
- *the requested upgrades for NSLS, and*
- *operation of general user crystallography beamlines at the six synchrotron x-ray facilities*

## I. Introduction and Charge of the Subcommittee

The six years that have elapsed since the previous Structural Biology Subcommittee Report of 1992 have seen a dramatic increase in the use of Department of Energy synchrotron light sources by the life sciences and especially by structural biology. In addition, exciting advances have been made in a variety of structural technologies. Three-dimensional structure determination by nuclear magnetic resonance (NMR) has evolved to the stage where the solution of structures of proteins up to 25-30 kilodaltons has become routine and proteins in the 40-45 kilodalton range are being tackled. In the realm of protein characterization, mass spectrometry has developed into a critically important analytical method that can measure the mass of proteins up to 250 kilodaltons and evaluate the primary structure with remarkable accuracy using only femto- to picomoles of material. The significant progress achieved in the human genome project has created the need to determine the three-dimensional structures of a large number (hundreds to thousands) of novel proteins from a range of medically relevant species producing a new field called "structural genomics". Finally, after a period of relative turmoil at DOE laboratories for neutron scattering in the U.S. marked by interruptions of reactor operating schedules (at Oak Ridge and Brookhaven) and the cancellation of the Advanced Neutron Source project, DOE/BES has settled upon a strategy for neutron source investment that encompasses upgrades to existing facilities as well as the planned next generation Spallation Neutron Source (SNS). These developments offer the structural biology community new opportunities to consider.

In light of these developments, the Structural Biology Subcommittee of the Biological and Environmental Research Advisory Committee (BERAC) was requested to review the needs of structural biology in these respective areas and to advise the office of Biological and Environmental Research (BER) on the portfolio of research grants currently being funded in this area. This request appeared in the form of two charges.

One, dated May 28, 1998, was focused on the use of synchrotron radiation for macromolecular crystallography. A committee of expert protein crystallographers was convened (called the 'Synchrotron Subcommittee' in this report), met on July 13, 1998, and produced a report (called herein the 'Structural Biology Subcommittee Report of July 1998'). The other charge to the Structural Biology Subcommittee, dated June 10, 1997, was broader and dealt with the range of structural technologies and discoveries described above. A second panel of scientists was selected to address this charge. This Subcommittee met on August 16 and 17, 1998, to discuss these areas. This report summarizes the conclusions and recommendations of this latter Structural Biology Subcommittee.

This report is not intended to cover all methods in structural biology. Imaging techniques such as electron microscopy and soft x-ray could be included in a future review.



## II. Review of Recommendations from the 1992 Structural Biology Subcommittee Report

In reviewing the Structural Biology report of 1992, the Subcommittee was struck by the fact that many of the issues that were considered most critical six years ago remain of prime concern today. The recommendations of the 1992 report will be listed (*italics*) followed by the comments of the current Subcommittee on the progress made in the implementation of these recommendations (plain text) and then by the relationship to the recommendations of the current review (**bold text**):

*1. Improve facility and beamline function to maximize efficient operation. This includes funding the basic operation of SSRL full time and enhancing the structural biology beamlines at NSLS with respect to hardware, personnel, and laboratory space to permit effective use of these facilities. It also included improving management and peer-review for these user facilities.*

Virtually the same recommendations were the primary outcome of the discussion of Synchrotron Subcommittee report of July 1998, and are endorsed by this Subcommittee, i.e., to increase staffing for beamlines, to improve hardware, and to optimize peer-review for more rapid and effective allotment of beam time to the users. However, it should be emphasized that considerable progress has been made since 1992 in this area. SSRL is now running at maximal operational capacity. The existing structural biology beamlines at SSRL and NSLS are functioning very effectively as service facilities with hundreds of satisfied users passing through these facilities annually. Laboratory buildings were constructed at both NSLS and SSRL to permit structural biology experiments to be performed more effectively on site. User support personnel were added to these beamlines to help support the user services. Most interestingly, the 1992 recommendations were incorporated into the design for the new synchrotrons, ALS and APS, that have come into operation in the past 2 to 4 years. For example, the basic facility designs included laboratory modules.

**Why then are these same issues at the forefront of the recommendations for synchrotron use today? The answer is that the demand for these facilities has increased and is anticipated to continue to increase dramatically in the foreseeable future. To respond to this demand, a number of new beamlines are currently in commissioning stage at each of the synchrotrons servicing this community. All these facilities (i.e., both existing and new ones) need to be properly funded and implemented—hence the continued recommendation for improved personnel, hardware, and user beam time review systems.**

*2. Recommendation to provide neutron sources for structural biology. This included bringing the High Flux Beam Reactor (HFBR) to full power and expanding the biological neutron program and instrumentation at the HFBR specifically by addition of a second neutron crystallography station. It also was proposed to appoint a review committee to consider the relative merits of pulsed versus reactor-based neutron sources for structural biology applications.*

The DOE neutron sources in the U.S. have been subject to repeated reviews, and the situation for neutrons continues to be critical today. The problems at Brookhaven have prevented the HFBR from running and its future continues to be unknown. At the same time, advances in NMR and ultra-high resolution crystallography at synchrotron sources have unexpectedly overtaken the scientific niche previously occupied by neutron protein crystallography. In contrast, there is increasing interest and potential for scattering from non-crystalline systems using both neutrons and X-rays. The potential for the High Flux Isotope Reactor (HFIR) at Oak Ridge for scattering applications in structural biology will dramatically increase upon completion of the cold neutron source upgrade project within the next two years.

The BER investment strategy must reflect this shift in the 'niche' for neutrons in structural biology from high-resolution crystallography to low-resolution scattering from multi-component systems. It is recommended that BER consider investing in a small-angle scattering instrument at Oak Ridge dedicated to structural biology.

3. *Priorities were defined for the development of the new synchrotron facilities: APS and ALS. These placed macromolecular crystallography at APS as first priority, followed by small angle scattering and diffraction at APS, and then soft x-ray microscopy and spectroscopy at ALS.*

Certainly, the first priority has been largely accomplished with the current construction of eight new beamlines at APS for macromolecular crystallography. BER is funding two user beamlines at APS (SBC CAT) for macromolecular crystallography. Two beamlines at APS are devoted to x-ray scattering or x-ray absorption (BIO CAT) under funding from NIH. BER is also funding the macromolecular crystallography and soft x-ray spectroscopy programs at the ALS.

The rising interest in solution x-ray scattering and spectroscopy in the U.S. resonates with the trends in neutron scattering. Now there are a number of synchrotron beamlines dedicated to these techniques, and investment in operations and user support becomes critical.

4. *DOE, together with NSF and NIH, should play a role in the development of very high field magnets for NMR.*

Since 1992, 750 MHz spectrometers have become common in NMR research groups across the U.S. The Environmental Molecular Science Laboratory (EMSL) at PNNL is funded by BER to acquire a 900+ MHz magnet for NMR that would be part of a user facility. BER has also contributed to the 900 MHz solid state NMR currently being constructed at the University of Pennsylvania.

The Subcommittee continues to support the recommendations that DOE work with other federal agencies to advance the development of GHz class, and beyond, magnets for NMR applications as well as user facilities to make this technology broadly available.

5. *DOE should build a computational and theoretical structural biology capability at the national laboratories to help understand protein structure and function.*

A research program in computational biology has been funded by BER although largely at universities rather than at the national laboratories.

**This area is addressed in the recommendations on computational biology below.**

6. *DOE, NIH, NIST, AND NSF should work together to coordinate funding and support for structural biology.*

Coordination between the major structural biology funding agencies has increased significantly over the past five years. This effort has culminated recently in the formation of an OSTP working group consisting of representatives of DOE-BES, DOE-BER, NIH-GM, NIH-NCRR, NSF, and NIST working together formally to oversee and optimize funding for structural biology over the next several years.

**Further recommendations on this issue are presented below in Section VIII on interagency cooperation and include most of the areas that encompass the major recommendations of this report.**

### III. Overview of the Recommendations of the July 1998 Structural Biology Subcommittee Report on Macromolecular Crystallographic Use of Synchrotron Facilities

This Subcommittee reviewed the report of the Synchrotron Subcommittee produced in July 1998 on the use of synchrotron radiation for macromolecular crystallography. The conclusions and recommendations as condensed in the executive summary of that report are recapitulated here:

#### Improvements recommended for current beamlines<sup>2</sup>

The number of beamlines currently in use or in construction for macromolecular crystallography has increased significantly in the past six to seven years. These beamlines serve a broad geographical distribution of users at six synchrotrons meeting structural biology x-ray needs (ALS, APS, CAMD, CHESS, NSLS, and SSRL). It was the sense of the Synchrotron Subcommittee that, from an efficiency, productivity, and timeliness perspective, the first priority should be to upgrade these existing beamlines to maximize utilization. This involves the following investments:

1. *Staffing*: increases are recommended in staffing at the rate of one FTE/beamline, on average, to improve round-the-clock user support, increase efficiency of operation, and better provide for less-experienced and non-specialist users. The Synchrotron Subcommittee regards this as a minimal increase relative to what is really needed for a fully efficient and functional operation. *Estimated cost: annual incremental expenditure of \$3M/yr (~\$125K/FTE).*
2. *Hardware upgrades*: Improved hardware such as newer CCD detectors, backup detectors to avoid wasting time after catastrophic failures, and improved optics and instrumentation to increase beam intensity or broaden versatility of a beamline. *Estimated cost: initial expenditure of ~\$10M. Estimated annual capital requirement of \$250K/yr/beamline to maintain effective state of hardware (some of which is already accounted for in existing budgets).*

#### Improved access process for synchrotron beam time and user education

Increased use and modern demands of research require changing the way beam time is allocated:

1. Rapid access to beam time within a month or two using rolling reviews
2. Regional or facility-wide access system for time slot applications
3. Increase number of longer term 'program' allocations
4. Provide for 'instant access' for hour-long time blocks for sample evaluation

---

<sup>2</sup> For the purposes of this report, a beamline is defined throughout as an independently operating station, i.e., an experimental station/hutch.

5. Development of common cross-facility user-friendly graphics interfaces across beamlines
6. Institute programs, such as workshops, to educate less-experienced and non-specialist users

### **General Facility Operations and Upgrades**

1. The Synchrotron Subcommittee supports the recommendations of the Birgeneau/Shen BESAC report urging continued funding for effective sustained basic operations for all four DOE synchrotrons. The Synchrotron Subcommittee also supports the beamline upgrades for APS and NSLS (priority 2) and the facility upgrades for NSLS and SSRL (priority 2.5), recommended by Birgeneau/Shen, as necessary for continued support of the macromolecular crystallographic community.
2. Funding of the operating shortfall of CHESS of ~\$0.5M/yr by the government agencies cooperating under the OSTP auspices.

### **Future research efforts: Detector, automation, and methodology development**

To foster continued expansion of facility access, efficiency, and capabilities, the Synchrotron Subcommittee recommends support of research in:

1. Advanced detectors
2. Develop beamline automation
3. Methodology development, such as optimizing data collection, processing, and refinement strategies to enhance synchrotron use. Development of new web-based tools to permit active access to beamline experimentation by remote users

### **New beamlines**

A number of crystallography beamlines are currently in the commissioning, development of proposal stage at six of the synchrotron facilities. The Synchrotron Subcommittee urges a case-by-case review of new beamline proposals by funding agencies with weight given to innovative applications.

### ***Comments and Recommendations***

*Considerable time was spent reviewing and discussing the recommendations of the Synchrotron Subcommittee. This Subcommittee endorses those recommendations. In several areas, this report will expand on important issues that were touched upon in the July, 1998 report. Such areas include an expanded discussion of detector development and other instrumentation (Section IV).*

## IV Detector, other instrumentation, and computational development for x-ray crystallography

### X-ray Detector Development

#### Background

The 1992 BioSync Report identified fast, efficient x-ray detectors as the most critical technological need for effective utilization of synchrotron radiation resources. The development of Charge Coupled Device (CCD) detector systems, which was largely funded by the BER Instrumentation Program, has probably been one of the most important developments over the last half decade enabling the exponential growth in synchrotron structural biology experiments.

Nevertheless, the current state-of-the-art beamlines are very powerful and will not be utilized to their full potential even with the best available CCD detectors. In some cases, this is because the detectors are poorly integrated into the beamlines, in need of better software, or hampered by nonstandard data formats and inadequate analysis protocols. In other cases, the limitations arise from the detectors themselves. For example, at insertion device beamlines, even the newest CCD detectors are not capable of reading out data as fast as it can be collected nor do they have the resolution to capture all the data from very large macromolecular assemblies. Other experiments are still completely detector limited, including those relying on time-resolved Laue diffraction, rapid measurement of very thin-sliced data images, energy discrimination, or complex local manipulation of the diffraction signal. Most fundamentally, the inflexible read-out property of CCDs limits the kinds of experiments that can be performed.

#### Critical Issues

Much remains to be done on developing detector systems and ancillary instrumentation. Fortunately, newer technologies now in early stages of development have the clear potential to meet many of the present and future needs of synchrotron radiation researchers. These include Pixel Array Detectors (PADs) and devices based on amorphous silicon and on superconductors. The challenge will be to apportion resources wisely to provide for existing needs with the best commercially available detectors, while simultaneously nurturing the research necessary to develop future detector technologies. The complexity of the devices that are being designed, the expensive custom-fabrication costs during development, and a stable infrastructure of design personnel all require consistent long-term funding for these crucial programs to succeed. Detector research has become increasingly complex and expensive, and the time required to fully develop a new detector technology has increased. This has led to a decrease in the number of detector research groups, even as the needs for better detectors have grown. Care must be exercised to prevent the dissolution of the remaining productive groups.

Once the next generation of detectors is developed, additional resources will be required to properly integrate them into the beamlines.

### *Recommendations*

1. *We recommend support of a two pronged approach:*
  - a. *The best commercial detectors currently available should be installed and integrated on existing beamlines.<sup>3</sup>*
  - b. *Long-term detector research programs must be encouraged*

The urgency of the current needs must not come at the expense of long-term research. A wise sense of balance, with consistent funding over many years, will be needed.

2. *We recommend increasing the support of detector research.*

The increased cost of developing the new generation of detectors will require a sustained increase in funding. Although detector instrumentation research has been a small part of the BER program, it has proven essential to the success of other parts of the program; this is certain to continue to be true over the next decade. Detector research clearly benefits biological synchrotron workers supported by all the funding agencies. In the past, there has been wasteful duplication of effort in the detector research programs of the various agencies. We recommend a coordinated, cross-agency selection of detector research projects, with one agency taking the lead in each case to avoid a multiplication of administrative overhead and reporting requirements. In this regard, projects which address needs of both neutron and x-ray detection should be considered.

3. *Detector research projects will have to be carefully selected to cover the field of possibilities while still maintaining enough competition to keep the research active and healthy.*

Although some of the recent detector developments have come out of the national laboratories, significantly more have come from the university and industrial sectors. BER should support the best, proven expertise available, irrespective of the sector in which it is found. The most effective detector research projects have been those in which the development teams included scientists who needed the detectors to perform their biological research, and in which the resultant devices were judged not on their intrinsic merits, but upon the quality of the biological research that resulted. This approach is to be encouraged.

---

<sup>3</sup> This recommendation is consistent with the Structural Biology Subcommittee Report of July, 1998.

4. *Finally, as detector research nears application, increased resources will have to be dedicated to integrating the new generation of detectors into the beamlines.*

Integration will involve the development of better software, more accurate calibration procedures, networked analysis and data acquisition methods, and remote beamline control for the users via web-based technology.

## **Other Instrument and Computational Development**

### **Background**

Crystallographic and scattering experiments at new, brilliant synchrotron sources produce data at a much greater rate than ever before. The advent of the CCD detector as the standard for synchrotron beamlines also vastly increases the data acquisition rate by eliminating the very slow readout time for older image plate or film detectors. This places much greater demands on the computing environment at the beamline. Users must be able to access their data easily to evaluate the progress of an experiment, to reduce the diffraction images to indexed, integrated intensities, to evaluate the quality of data at the level of integrated intensities, and to transfer data to their home laboratories for further analysis.

Beamline computing environments have in general suffered from the shortage of staff to develop and implement software that would help to streamline experiments. As a result, the computing environment often is poorer at a synchrotron beamline than in the user's home laboratory, where data acquisition occurs orders of magnitude more slowly. Software and hardware generally lag in several areas.

Conversion of two-dimensional diffraction images to indexed intensities for the thousands of Bragg reflections in a typical experiment is an active area of research in macromolecular crystallography. Thus, several software packages, each with advantages and disadvantages, are presently in use. It is imperative that data processing software be readily accessible to users during the experiment. Each experienced user has a preferred data processing system. In general, beamline staff try to accommodate the wishes of all users, but this is not always possible because the programs tend to be detector specific. Lack of a common format for diffraction images hinders implementation of detector-general software.

The growing number of non-specialist and less-specialist users places great demands on beamline staff and software to support user experiments.

### **Critical Issues**

With respect to the data rate problem, beamlines must be equipped with sufficient data storage and networking resources to accommodate the needs of users. At today's prices, purchase of sufficient data storage capacity should not be a problem for any beamline. Beamline computing environments should be able to accommodate users who bring



their own workstations or disk drives for data processing and archiving. For the long term, synchrotron facilities should be connected to the high speed national research network so that data can be transmitted to the home laboratory electronically.

A common image-file format is critically needed for macromolecular crystallography. This will foster the development of cross-detector software and greatly simplify the job of beamline staff in implementing new data processing software. It will also prevent beamline operators and users from being at the mercy of software vendors whenever a new detector is used. This is especially important in view of the large number of detector types currently in use and the anticipated development of new detector technologies.

There is a need for development of software in several areas.

- As data are acquired rapidly, the user typically cannot keep a record of the experiment in a laboratory notebook as in the days of more leisurely experiments. Software is needed to keep an electronic log of the experiment, including information important to the user, *e.g.*, beam intensity, X-ray energy, time of day, camera settings, exposure time, image name and number.
- For many of today's challenging structural problems, an important application of synchrotron radiation is in screening crystals or solution samples to determine the feasibility of synchrotron data collection. Software is needed for rapid and convenient analysis of diffraction patterns and small-angle scattering patterns to assess sample quality.
- Improvements to data processing software are needed. Special needs of the challenging projects that users bring to synchrotron sources are highly mosaic crystals, weakly diffracting crystals, very large unit cells, and polychromatic data.
- A longer term goal is to allow users to interact with, or at least to view, experiments at the synchrotron or neutron facility from their home laboratory. This will require development of specialized software and access to a high-speed research network.

Because the need for software improvements is quite broad, BER should not expect to address all problems, but should concentrate its efforts on the special needs created by the research possibilities of synchrotron and, if appropriate, neutron facilities.

There is a growing number of non-specialist users who lack the expertise to solve the structural problems they bring to the synchrotron source. Non-structuralist users are generally viewed as an additional responsibility by beamline operators, and indeed require more staff and software support. However, an excellent synergy could be developed between such users and beamline scientific staff seeking collaborations in structural biology research projects. Such interdisciplinary collaborations could be of great benefit to the advancement of biology.

### *Recommendations*

Greater priority should be given to development of software that is closer to the experiment, among the many software development projects BER may be asked to support.

1. *We recommend that sufficient human and hardware resources be provided to synchrotron beamlines to implement software for data evaluation, processing and analysis that is suited to the rates of data acquisition.*  
This includes all data processing and analysis software packages in wide use in the user community. Users should also be given remote access to evaluative software from their home laboratories.
2. *We recommend that BER foster the availability of high-speed network services to users of structural biology beamlines.*  
The goal is transfer of image data from the synchrotron to the home laboratory over a high-speed (100 Mbits/sec or faster) research network that is isolated from the commodity Internet. Such a network could also be used for remote interaction with the experiment by users in their home laboratories.
3. *We recommend that BER encourage efforts to develop a common, detector-independent format for two-dimensional diffraction images.*
4. *We recommend that BER look kindly upon requests from beamline operators to augment their facilities, especially in scientific staff, to accommodate the greater support needed by demonstrated numbers of non-structuralist users.*

## Automation of Sample Handling

### Background

At present, tremendous effort is being expended in the direction of high-throughput protein purification and crystallization by a number of different research laboratories. In addition, more intense beamlines have resulted in more rapid data collection and faster detector readout in reduced data collection periods. In many cases, more beam time is now spent on mounting and screening individual samples than on a full data collection run of a single sample; this results in the increasingly inefficient use of valuable beam time. For example, micro-crystals of membrane proteins have been observed to diffract, but deteriorate quickly even under low temperature conditions. For a successful and complete data collection, multiple sample crystals are often required leading to significant time spent in mounting and alignment.

The logical follow-up to automated systems of protein preparation is automation of crystal mounting, alignment, and data collection. Automation of microcrystal manipulation, mounting, alignment, and data collection has been successful for 'small molecules' where structures have been solved using microcrystals smaller than  $1 \mu\text{m}^3$ . For these inorganic samples, a micromanipulator has been used to handle crystals that are  $0.1 \mu\text{m}$  in size using technology similar to that developed for manipulation of single cells under a microscope. Extension of this approach using microcrystals has been explored with several proteins (e.g., bacteriorhodopsin, 5 by 30 by  $40 \mu\text{m}$  on edge).

One automation method system currently under development exploits technology developed for the Human Genome Project. The goal is to reduce the amount of time necessary for crystal mounting and optical alignment using automated robotics workstations. These steps often consume a significant amount of synchrotron beam time as one has to go in and out of the beam 'hutch' to mount and align the samples. Ideas for future development include optical methods for rapid crystal centering and new hutch designs to allow for alignment/mounting in parallel with data collection. In addition, small synchrotron beams (1  $\mu\text{m}$  in size using Kirkpatrick Baez grazing incidence focusing mirrors) may minimize the effects of crystal decay by allowing for selection of different parts of the crystal for optimizing diffraction and for data collection. This capability would benefit from automation.

Automation technology will be invaluable for high throughput structure analysis that will be required in structural genomics and high-throughput structure-based drug design.

### Critical Issues

A non-trivial fraction of synchrotron beam time assigned to protein crystallography is devoted to crystal mounting, alignment and screening. This fraction increases with more intense beams and faster data collection. Some of this 'lost' beam time can be recaptured using automation. Since the majority of synchrotron sources are DOE operated, DOE should take the forefront in initiating research into beamline automation.

Some efficiencies have been realized by researchers pre-screening their samples in the home laboratory. Nevertheless, many poor data sets are still collected due to the long time required to adequately evaluate data quality relative to data collection time. Development of software for the more efficient evaluation of data quality is critical for the optimal use of our current resources.

### Recommendations

We therefore make the following recommendations:

1. *BER support for the development of beamline automation.*  
This effort includes automating crystal mounting and alignment.
2. *BER encouragement for the development of improved software for data collection and data quality analysis is recommended.*  
Improved software is needed to analyze the data immediately as they are collected.

## V. Structural Biology and the Evolving Need for High Field Nuclear Magnetic Resonance (NMR), Mass Spectrometry (MS), Neutron Diffraction and Scattering, and X-ray Scattering and Spectroscopy (XAS) Facilities

### Nuclear Magnetic Resonance (NMR)

#### Background

The increasingly important role of NMR in structural biology is powerfully illustrated by the number of novel structures of biological macromolecules determined by this technique over the last few years. For example, in 1996 approximately 100 new NMR structures were deposited in the PDB, compared to ca. 400 crystal structures, a remarkable number given the fact that the structural NMR field is still in the developing phase compared to the mature area of crystallography. In addition, a growing number of NMR groups have been established at universities, research institutions, and industry, reflecting the vigor of the field. But NMR is not only a technique for structure determination, it is also able to provide a wealth of information which is complementary to the structural data. NMR data on dynamics and solvation of proteins and nucleic acids is becoming available as well as NMR characterization of partially folded or unfolded protein and polypeptide states.

#### Critical Issues

Compared to x-ray crystallography, where data collection at synchrotron sources can be completed in a short period of time, NMR data collection still represents a major investment in real time and in equipment time and may require several months for one structure. One way of increasing the number of NMR structures may lie in the creation of high-throughput service facilities as started in Japan, where as many as a hundred NMR instruments will be provided for data collection by outside users.

A second research goal is to increase the size of the protein or macromolecular complex that can be analyzed by NMR. This requires the development of even higher-field spectrometers, which will yield higher sensitivity and better resolution, thereby improving the quality of structures of all sizes. In particular, instruments at 1 GHz or higher will need to be developed for this purpose, challenging current magnet technology and pushing magnet developers to new materials. The availability of such higher field instruments will be crucial for the further advancement of structural NMR.

#### *Recommendations*

1. *A strong DOE role in the aggressive pursuit of the best possible initiatives to develop high field instruments is recommended.*

This can best be accomplished by BER identifying the most promising high-field NMR magnet research effort and partnering with it.

2. *The creation of regional facilities for housing new high field instruments is highly desirable in order to allow access to this improved technology for a large number of NMR researchers. Coordination across funding agencies will be essential for these facilities since the NMR users are likely to be funded from a variety of sources.*

## Mass Spectrometry

### Background

Revolutionary new methodologies for the sequencing and structural characterization of bio-macromolecules at the picomole level and below have emerged during the past decade. These developments were triggered by the virtually simultaneous discovery of two powerful new ionization/desorption techniques that provided mass spectrometric-based technologies with the inherent capability to detect and analyze high molecular weight polar, labile biopolymers and their digests for the first time. These are matrix-assisted laser desorption ionization (MALDI) and electrospray ionization (ESI). Together with the commercialization of a variety of new strategies in mass spectrometric instrumentation, MALDI and ESI have quickly provided the biomedical research community with a suite of tools of unprecedented power for the detection and comprehensive identification of proteins at the picomole level and below and with the structural characterization of their post-translational and xenobiotic modifications. They have become the methods of choice for protein characterization in the field of protein biology, and are currently showing promise for large-scale rapid sequencing of human gene polymorphisms of importance in assessing individual susceptibility to various diseases.

During the same period, a major international effort was initiated to sequence the human genome, as well as a growing list of bacterial and other model genomes. Since 1990, these achievements are the result of a well focused, brute-force technological effort supported by DOE and NIH. The expected accomplishment of this goal within the next 3 to 4 years sets the stage for the challenge to understand the function of the large number of proteins that will be discovered in the massive DNA sequencing effort. Meanwhile, mass spectrometry has emerged quickly and somewhat unexpectedly as the key technology with the inherent power and speed required for comprehensive characterization of the machinery of cells. Confronting this task is the next logical major challenge in deciphering the precise molecular basis of homeostasis and cell dysfunction. Pursuit of this challenge would establish a direct link between the functioning machinery of cells and the field of functional genomics with its exploding genomic, protein, and expressed sequence tag (EST) databases.

However, while the technology is ripe for such timely mobilization, mass spectrometry in the U.S. remains an individual investigator-initiated effort that includes five NCRR research resources. These research resources are focused, through investigator-initiated projects, on application of mass spectrometric methods in a somewhat *ad hoc* manner. They tackle a wide variety of biomedical research problems involving the detection, sequencing and structure elucidation of cellular constituents. Although some

groups already have much of the expertise required, none are presently geared to address the genome challenge with resources on a scale comparable to those that are beginning to emerge for structural genomics.

A second significant development is isotope ratio mass spectrometry using accelerator technology and atom counting techniques. This development has been pioneered for biological tracer studies, especially for  $^{14}\text{C}$ , by Lawrence Livermore National Laboratory and provides absolute sensitivity for  $^{14}\text{C}$  detection of some 5–6 orders of magnitude over scintillation counting or fluorography.

### Critical Issues

There is a need for DOE leadership, coordination, and support for a major effort for the rapid identification of all proteins expressed in the estimated 250 human cell types and their correlation with structural and functional genomics. Based on development of mass spectrometric technologies involving sample handling robotics for multidimensional chromatography and computer-based data analysis, these goals could be achieved rather quickly with a comparatively modest level of dedicated support. Investment in computer-based intelligent application of, data management for, and archiving is required for the rapidly emerging mass spectrometry technology.

There are critical needs to foster design and development of new ultimate-sensitivity ion-optical strategies for detection and identification of the protein composition of a single cell. This goal must be carried out in the context of development of new strategies for handling, separating, and ionizing sub-attomole quantities of proteins and other important biomacromolecules.

### *Recommendations*

1. *Participate with other agencies in formulating a national consensus for integration of non-proprietary, large scale protein identification/characterization by mass spectrometry with the continuing efforts in structural genomics/bioinformatics and computational biology.*
2. *Continue support of Bio-AMS (accelerator mass spectrometry) and the development of a 'small' machine dedicated to biomedical research measurements.*
3. *Leverage investments in detector technology development (i.e., IC pixel arrays, see Section IV above) for new ion-optical strategies in ultra sensitive biological mass spectrometry. This should be a cooperative effort with the NIH and NSF.*
4. *Foster, in cooperation with NIH and NSF, opportunities in instrument development such as study of non-covalent protein complexes and integration of attomole and sub-attomole separation technologies and chip-based strategies with mass spectrometric identification instrumentation.*

## Neutron Crystallography

### Background

X-ray and neutron diffraction techniques are similar in both their experimental methodologies and in the resulting informational content. The impetus to carry out high resolution neutron crystallography was based on its ability to locate experimentally hydrogen (or deuterium) atoms in large molecules. This is particularly significant in the study of protein behavior because typically half the total number of atoms are hydrogen and these dominate much of the chemistry and physical structure of proteins. The application of high resolution neutron crystallography to assign hydrogen atom positions in proteins and to differentiate between hydrogen and deuterium atoms has been focused mainly on structural issues in three research areas: 1) protein reaction mechanisms, 2) protein dynamics (using H/D exchange), and 3) protein-water interactions. Protein hydration is a principal player in the physical chemistry of the molecule. Because deuterium has a scattering length of the same magnitude as oxygen or carbon, water molecules scatter neutrons with about 3 times the magnitude as they do in the x-ray case.

Consequently, in the past, high resolution neutron crystallography filled a small niche focusing on carefully designed experiments for specific systems where the inherent technical problems could be overcome. These problems are related to the low flux of the available neutron sources and the large backgrounds created by incoherent scattering from hydrogen atoms in the sample. The only available solution to the problems was to grow extremely large crystals ( $> 1 \text{ mm}^3$ ), and if possible use deuterated proteins. Unfortunately, very few protein crystals can be grown anywhere near this size. Superimposed upon these significant technical issues has been the serious impediment of lack of availability of neutron sources over the past decade. Thus, over the last 15 years, less than 10 different high resolution ( $< 3 \text{ \AA}$ ) 3D protein structures have been determined by neutron crystallography. Lower resolution neutron diffraction experiments have also found their niches in structural biology. Medium-resolution (3–8  $\text{\AA}$ ) neutron diffraction has played a role in fiber and membrane structural studies locating water-filled pores in membrane proteins structures, or the position on water molecules in different forms of DNA. Low-resolution ( $> 8 \text{ \AA}$ ) neutron diffraction with contrast variation has been useful in locating disordered nucleic acids in viruses or detergent molecules in membrane protein crystals. These low-resolution methods do not require  $1 \text{ mm}^3$  crystals, and can handle cells up to 1000  $\text{\AA}$ .

### Critical Issues

In the past, even considering the heroic efforts that were necessary to complete a high resolution neutron protein structure, resources were requested for neutron crystallography based on the ability to extract unique structural information. There was also the promise of new instrumentation and methodology that would potentially improve the technical situation. However, with the advent of high resolution NMR and

very brilliant synchrotron x-radiation sources, the need for neutron crystallography may no longer be as persuasive as it was. Hydrogen exchange (H/D) as a probe for protein dynamics is done as a by-product of many NMR structure determinations. These NMR data are collected quite readily and have a dynamic range at least an order of magnitude greater than can be achieved by neutrons. Furthermore, the NMR experiments are more versatile and superior because they are performed in solution phase rather than in the solid state of the crystal. The H/D exchange information available from NMR structures provides the biophysicist a rich database. As NMR methods are being developed to handle larger and larger structures, there is less need for complementary, lower resolution data, from neutrons.

Another challenge to the importance of high resolution neutron crystallography for structural biology is ultra-high resolution protein crystallography—an increasingly common by-product of synchrotron analyses. There are a growing number of structures being analyzed at resolutions better than 1 Å using high intensity synchrotron radiation and cryo-crystallography. At the ACA meeting in Washington (July 1998) there were six posters presenting ultra-high resolution crystal structures on a variety of enzyme structures. An estimated 15–20% of crystals measured at the synchrotron sources diffract to better than 1 Å resolution for structures up to about 50 kDa. For larger systems there are fewer statistics to make this estimate reliably at this time, but examples are appearing in the 100s of kDa size range. At these resolutions, individual hydrogen atom positions can be determined with a similar confidence level to that available from neutron crystallography. The difference is that very much smaller, much more readily produced crystals can be used and data collection times are orders of magnitude faster opening up the technique to a much broader range of systems. One arena where neutrons still provide an advantage is in cases where it is important to obtain data on the locations of hydrogens at ambient temperatures, e.g., for studying solvent structure in crystals.

In addition to these challenges to high resolution neutron crystallography, there are also challenges for medium- and low-resolution neutron diffraction studies. X-ray studies of viruses now routinely locate the disordered nucleic acids in these structures, and there are indications that analysis of low-angle x-ray diffraction data at synchrotron sources may hold promise for further advances in the utilization of x-ray data for locating less ordered structures in crystals. There are two operating neutron protein crystallography stations in the world today; a quasi-Laue neutron station at the JRR-3M research reactor in Japan and a quasi-Laue neutron diffraction instrument at the Institut Laue Langevin (ILL) that has been demonstrated to be able to collect a 2 Å resolution data set from 2 x 2 x 1.5 mm<sup>3</sup> crystals of deuterated lysozyme in 10 days. This instrument represents the current state-of-the-art in neutron crystallography.

### *Recommendations*

In consideration of the following facts:

1. Based on the advances in alternative technologies, the niche for neutron crystallographic studies appears to be growing smaller. The impressive successes of NMR



solution methods and ultra-high resolution synchrotron-based x-ray structures are relatively recent and could not have been predicted several years ago. However, this development is a fact that needs to be weighed in planning future strategies for neutron crystallography.

2. The design and construction of a neutron crystallography station at Los Alamos has begun and there has been a proposal to build a neutron crystallography station at ORNL.

The recommendations of the Subcommittee are:

1. *Not to build a protein crystallography station at ORNL on the new cold source being installed at HFIR; the beamline can much better serve the community in a small-angle neutron scattering (SANS) configuration.*

This resource will be especially important given the historic dearth of cold neutrons in the U.S.

2. The protein crystallography station at Los Alamos is already in the design/construction phase. The station uses a "quasi-Laue" configuration that has been proposed to solve a portion of the neutron flux problem, but is as yet unproven. The ILL instrument (LADI) is in operation. The ILL and Los Alamos projects provide logical foci for continuing to evaluate the future scientific case for neutron crystallography. *Future operational support for the Los Alamos instrument by BER must be evaluated in terms of its performance and the size of the user community it attracts. Because of the rapid development of competing technologies, consideration might be given to what range of problems, including condensed matter science, could be served by the instrument.*
3. *Evaluation of neutron protein crystallography for future planning must engage a broad range of structural biologists. In addition to a group of neutron experts it should include leaders in competing technologies such as synchrotron crystallography, NMR, and microscopy.*

## Neutron Scattering and X-ray Scattering

### Background

Small-angle scattering of x-rays or neutrons from biological macromolecules in solution is playing an increasingly important role in structural molecular biology. The techniques yield information on molecular associations and overall shapes of biomolecules in solution. In the case of neutron scattering, deuterium labeling with contrast variation allows one to extract information on the shapes and dispositions of individual components of complex assemblies of biomolecules. Small-angle scattering therefore can be extremely powerful when combined with data from high resolution techniques for studying the dynamic interactions and conformational flexibility inherent in the regulated functioning of molecular networks that, for example, transmit and amplify signals, or are involved in energy transduction, transport, mechanical movement, etc. Recent years have seen a number of advances in sources and instrumentation for small-angle scattering that have

yielded gains in the flux of x-rays and/or neutrons on samples. These advances have facilitated more rapid experiments on smaller samples using lower concentrations. Small-angle instrumentation at synchrotron sources has made possible time-resolved studies of protein conformational changes, for example, during protein folding. At the same time, advances in biotechnology have made sample production and deuterium labeling easier and cheaper. In addition, the availability of faster and cheaper computers has allowed more investigators to do more sophisticated modeling to interpret scattering data. These advances are having their impact on the field, as we see that in the last two years, publications of small-angle x-ray and neutron scattering studies of biomolecules have doubled as the technology becomes more accessible and more sophisticated.

### Critical Issues

Small-angle neutron scattering in biology requires the highest intensity cold neutron sources available. The U.S. has had a critical shortage of cold neutrons for the past two decades. During that period, the ILL in Grenoble, France, has been operating a highly successful scattering program with applications in materials, polymers, and biological systems. In general, European scientists have broad training and experience in scattering techniques, while in the U.S. DOE labs the focus has been strongly directed to high resolution diffraction techniques, except for NIST where the focus has been on SANS and reflectometry. The current move toward structural and functional genomics requires broader application of scattering and diffraction techniques in order to understand how networks of biomolecules interact to achieve coordinated function. The four critical issues that need to be addressed for U.S. scientists to be able to make this step are:

- Availability of competitive scattering instrumentation on cold neutron beam lines
- Availability of deuterium labeled samples for neutron contrast variation studies
- Support for effective operation for existing small-angle x-ray scattering beam lines at SSRL and APS
- The means to support new users to use more diverse scattering techniques.

### *Recommendations*

*Partner strongly with DOE/BES to ensure optimal utilization of the new ORNL cold neutron source (to be operational in the year 2000). BES is funding one small-angle scattering instrument on that cold source that will be used primarily by the materials/polymer science community. BER could support a second small-angle instrument optimized for and dedicated to biology.*

1. *Support for technologies to decrease the cost and increase the ease for production of deuterium labeled components will be important for the long term success of these approaches.*  
Neutron contrast variation techniques depend critically upon the ability to label protein components, or selected protein domains with deuterium.

2. *Proactively build the U.S. small-angle scattering community, with an emphasis on structural biology applications.*

Planning for resource allocations should allow for a significant growth in the use of small-angle scattering techniques using both neutrons and x-rays. The technique is now poised for expansion into a much wider range of laboratories. Specifically:

- *Facilitate access to high-intensity, well instrumented cold neutron beamlines and synchrotron sources. Support operations of existing small-angle beamlines at SSRL and APS (designed for biological applications), and the proposed small-angle instruments at ORNL.*

Investment in user support should include software development, sample environment, hardware for time-resolved studies, and biological sample preparation support facilities where needed.

- *Support for new users to get into the field.*  
This will require adequate staffing at the major facilities to provide technical support for data acquisition and interpretation. Partnering with the very successful small-angle neutron scattering program at NIST would be beneficial.
- *Support development of standardized user-friendly data handling and modeling software packages that can run on most common computer platforms.*

4. *Support a workshop to evaluate the trends of small-angle scattering in structural biology.*

This workshop should be specifically charged with evaluating pulsed and reactor sources for small-angle scattering applications, with particular emphasis on evaluating if the SNS will be suitable for small-angle scattering applications. In the case of pulsed sources, both short and long pulsed sources should be considered. Since NIST is currently stewarding the most competitive small-angle scattering capabilities for biology in the U.S., they should be prominently involved in the workshop.

5. *Partner with NIST to evaluate reflectometry (small-angle scattering from thin films and surfaces) potential in structural biology.*

NIST is currently sponsoring an NSF planning project for a dedicated reflectometer for biological applications. The outcome of this project should be evaluated in terms of user participation.

## X-ray Absorption Spectroscopy

### Background

X-ray absorption spectroscopy (XAS) is a non-crystallographic method that is used to probe the structure of specific atomic constituents in complex materials. XAS studies of biological systems all require the use of synchrotron radiation since no conventional source provides the required high intensity over the range of wavelengths needed for the experiments. Hence, such experiments depend upon the availability of specialized beamlines and instrumentation at the synchrotron facilities. From these XAS experiments, one obtains information that includes electronic structural information about the absorbing atom as well as metrical details on its coordination or near neighbor

environment. For example, XAS studies have been especially important in determining the structure and function of metal sites in proteins where they play crucial roles in a range of important biological reactivity ranging from electron transfer to catalysis or metabolism involving small molecules like CO, O<sub>2</sub> and NO.

XAS measurements are complementary to crystallographic studies in that they can be performed on solutions (or frozen solutions), hence are applicable to biological materials that have not or cannot be crystallized. The method can be used to characterize reaction intermediates that can be produced either by stop-flow or freeze trapping. XAS can provide highly accurate metrical details about metal coordination. A specific limitation, however, is that it gives only local structural information about the absorbing atomic site. In this respect, it is especially complementary to x-ray crystallography, which does not produce as accurate metrical data for metalloproteins.

### Issues and Opportunities

The growing number of proteins identified, especially those deriving from the human and microbial genome projects, will lead to a wide range of new and interesting problems, including many where metal ions are essential constituents of biological structure and function. XAS studies of soluble and membrane bound proteins will provide valuable information about structure and function of the active sites in these systems, some of which will be quite complex involving multi-component, multisubunit assemblies.

One new area of opportunity is the combination of XAS and crystallography. By doing XAS studies on the same samples as used for crystallographic analysis, very accurate metal coordination information and direct information on metal oxidation state and electronic structure can be obtained. The XAS information is also very important in both understanding the chemical role of the metal and in determining and monitoring redox states. Application to reaction intermediates generated and freeze trapped in the crystalline state provides an approach to the study of active site metal ion function.

Another emerging area is the combination of good spatial resolution with XAS studies. Using specialized focusing optics, x-ray beams with dimensions of tens of microns to sub micron can be used to probe the element spatial distribution in a chemically sensitive way. For example, it is possible to monitor the uptake of toxic metal ions by plants or microbes, following both their spatial distribution and changes in their chemical speciation.

Advances in both x-ray optics and detectors are essential to continue to push the state-of-the-art in XAS applications to biological materials. Further developments in theory and analysis programs are also needed to broaden the range of applicability and to make the techniques more widely available to the non-specialist users.

## *Recommendations*

1. *The Subcommittee believes that, at an absolute minimum, the current investment in synchrotron beamlines used for XAS studies needs to be maintained and may well have to be expanded over the next 3–5 years. The Subcommittee recognized the specialized nature of XAS and the fact that it does not (and likely will not in the future) have the extensive user base of crystallography. Hence, it was felt that resources should be coordinated and focused in a way so as to optimize user service and science and that wider proliferation of such beamlines was probably not currently warranted.*

However, the potential need for increased capacity should be monitored and appropriate plans made if demand and new scientific opportunity warrant. This is an issue for coordination at least with NIH/NCRR since their program also supports XAS synchrotron facilities.

2. *Adequate staffing remains a critical issue as with structural biology synchrotron beamlines for both crystallography and scattering.*

XAS beamlines are on average not even staffed as well as crystallography lines and this continues to create heavy demands on available support staff. It also limits the effective training and dissemination of the technique to new user groups. In considering improvements to staffing levels, XAS should also be considered and again this is an interagency issue.

3. *The implementation of high performance, high energy resolution solid state detectors in association with MAD crystallographic beamlines would enable combined XAS/crystallographic studies to be done routinely.*

It would be very desirable to create such capabilities in selected locations.

4. *Further investments in the development of improved and new detector technologies will enable more effective use of the newer very high intensity beamlines, will push viable concentration limits lower, and will enhance capabilities for special classes of experiments including whole cell XAS and time-resolved XAS. Improvements in data processing and analysis theory and algorithms are also important in improving ease of use and efficiency as well as providing new capabilities.*

Such efforts are important and should be fostered.

## VI. Review of the Portfolio Balance of the BER Structural Biology Research Grants

### Background

The BER program provides critical support for the construction and ongoing operations of structural biology beamlines at the synchrotrons as well as facilities for research using neutrons. Also within the BER portfolio is support for instrumentation development, computational biology and biological research. The Subcommittee discussed the balance of funding and reached consensus on several points regarding future investments.

### *Recommendations*

1. *We recommend that BER maintain at least its current level of investment in facility development and operations.*

Such an investment is critical for enabling the structural biology community to make effective use of the synchrotron and neutron facilities (note, however, the comments about redirection of the neutron program in Section v of this report). Inflationary growth should be provided for these programs.

2. *The Subcommittee felt strongly that BER should focus research investments in areas that couple to the unique strengths that it has developed in association with the national user facilities and laboratories.*

Such a research focus would couple to the important and significant investments already being made and give rise to synergism that would enable both new research and continued instrumentation developments at the state-of-the-art.

3. *We encourage research involving the development of novel tools that will enhance access and throughput of the shared, multi-user facilities.*

Investments in areas like collaboratories and robotics could yield significant dividends. As other agencies are planning investments in some of these areas, coordination between agencies is very appropriate.

4. *We believe that on balance the level of investment in advanced instrumentation research is too modest.*

This is especially the case for detectors (see Section IV above). Research investments which couple to the national user facilities and laboratories should continue to go beyond exclusive support of research at the national laboratories. For example, many of the important instruments, that have facilitated the use of the national synchrotron sources, were developed with BER funding to universities and industry.

## VII. BER Program in Computational Biology and the Relationship to Structural Genomics

### Structural Genomics

#### Background

The goal of structural genomics is to determine the three dimensional structures of a large number of proteins, achieving as full a coverage of various genomes as possible. Protein structures will be obtained experimentally from x-ray crystallography and from NMR, and these experimentally determined structures will be extended and supplemented with homology models of related proteins. The criteria used in choosing proteins for crystallization will include evidence that they represent folds that have not yet been seen and/or that they are members of sequence families for which there are no representative three-dimensional structures. These criteria, as well as the exploitation of the structures that are solved, will require continuing developments in computational structural biology that are closely linked to the experimental effort.

#### Critical Issues

Most current structural work is motivated by the biological or medical significance of a particular protein. For this reason, many important structures require years of effort to complete. In contrast, inherent to structural genomics is the need to determine a large number of structures fairly quickly. This requires a major effort in high throughput cloning, expression, purification, and crystallization, much of which would have to be carried out in parallel so as to maximize throughput. More generally, the entire project requires coordination between groups with different areas of expertise. This is not easily accomplished within the framework of individual investigator-initiated research proposals. It is possible that the project might be accomplished primarily in the national laboratories and, if so, serious consideration must be given to the entire cost; not, for example, just of a beamline but rather of the underlying experimental infrastructure that will be required to 'feed' the beamline with adequate quantities of purified protein for crystallization and structure determination.

A related issue is that structural genomics is not by its nature 'hypothesis driven' in the sense that it does not focus on the determination of the structure and function of a particular protein target of biological interest. Rather it constitutes a massive data production effort that, in its most extensive form, contains some parallels to the Human Genome Project. It is not clear how to assign relative priorities to current research in structural biology and to the resources that will be required for an effective effort in structural genomics. There is no precedent for this type of effort in structural biology and the reaction and interest of the community needs to be assessed. Moreover, given the possible extent of the overall project, it should have a broad interagency focus.

A further issue is the possibility of one or more structural genomics efforts in the private sector. A number of initiatives of this type have been discussed and the interaction and competition of public and private efforts needs to be considered.

### *Recommendation*

*A broadly based panel of scientists should be assembled to evaluate the importance, feasibility, and cost of a cross-agency federally-funded national structural genomics effort conceptually like the Human Genome Program.*

Partnership with the NIH is essential to the success of such a workshop.

## Computational Structural Biology

### Background

Computational structural biology is a loosely defined field that, in its broadest definition, encompasses diverse research areas ranging from molecular dynamics and quantum mechanical studies of enzyme mechanisms to protein folding and reverse folding studies to sequence analysis tools based on mathematical methods such as dynamic programming or Hidden Markov models. The areas that are most directly related to an experimental effort in structural genomics include the development of new tools in sequence analysis including multiple sequence alignment, structure-based analysis of amino acid sequences, protein structure prediction based on fold recognition and homology model building methods, and structure-based analysis of protein function. The application of these tools to entire genomes and to sequence databases and the clustering of proteins into sequence and structural families is an essential step in choosing proteins for experimental study and for exploiting new sequence and structural data as they become available.

### Critical Issues

Many areas of computational structural biology are represented by large and active research communities that have been funded under standard grant mechanisms for some time. Computational research related to structural genomics is at a much earlier stage of development and involves a community with diverse scientific backgrounds. Funding, though available, has come from a variety of sources. It is notable in this regard that many of the developments in structural genomics originated in Europe; for example at the EMBL in Heidelberg, the MRC in Cambridge, and at a number of closely linked labs in London. In many cases, progress was due in part to the ability to assemble a centralized, stable, and interdisciplinary group, a model not easily duplicated in the U.S. While it is not clear that this model is necessarily desirable in the long term (indeed a number of outstanding European investigators have moved to the U.S.) it is important to determine the best way to assure stable funding sources and to provide the type of interdisciplinary training required in structural genomics.



Questions that arise include: a) what areas of computational structural biology should be included in a structural genomics effort, b) how to relate the funding efforts of different government agencies, c) whether the national laboratories should establish computational biology groups that are closely linked to experimental efforts in structural genomics, d) the desirability and possible mechanisms for funding academic centers (involving one or more universities) in computational structural genomics, e) the need to link computational structural biology to planned investments in high performance computing (ASCI), and f) what mechanisms are best suited to provide interdisciplinary training in computational structural genomics.

### *Recommendations*

Computational structural genomics is an important emerging research area, both in its impact on current genome sequencing projects and in its crucial contribution to possible experimental initiatives in structural genomics.

1. *The current BER research grant portfolio is strongly focused in this general area; the decision to move in this direction was a good one. This emphasis should be continued and should be used as a foundation for larger scale projects that would be an integral component of any broadly-based initiative in structural genomics.*
2. *Questions as to possible new funding mechanisms and the establishment of research centers should be considered in the context of the workshop proposed above in the general area of structural genomics.*

## VIII. Coordination of Funding Efforts Between BER, BES, NIH, and NSF

### Background

The need for interagency cooperation arises when a given research portfolio or initiative has significant components or user bases funded by different sources. This is especially the case for large, shared multiuse user facilities such as those providing synchrotron or neutron beams. Consider the synchrotron situation where the operation, maintenance and upkeep of the basic facilities is the purview of either DOE (for ALS, APS, NSLS, or SSRL) or NSF (CHESS), but the facilities serve an increasingly large fraction of research funded by another agency (generally NIH in the case of structural biology research).

The issue of operations support for the basic facility has been considered and discussed by a number of earlier subcommittees, including the Birgeneau-Shen BESAC Subcommittee and the BERAC Synchrotron Subcommittee. There is a strong consensus and feeling that operations support for a given facility should be the primary responsibility of a single division or group within a single agency and this Subcommittee strongly endorses this point. However, it is highly appropriate for other agencies to provide funds for construction and operations of specialized resources for research relevant to their mission and/or for use by their grantees of such facilities. It is also appropriate, when desirable and necessary to share the funding of major upgrades or improvements to the basic facilities. This latter point was also made by the Birgeneau-Shen BESAC Subcommittee in the context of the synchrotron upgrades.

Mechanisms for interagency coordination include, among others, joint solicitation/funding of proposals, coordinated focused programs, and working groups at the interagency level. The latter mechanism is currently being used to address supporting the growing needs for access to the synchrotron facilities by the macromolecular crystallography community. This group has been formed under the auspices of the OSTP, includes representatives of DOE-BER, DOE-BES, NIH-GM, NIH-NCRR, NSF, and NIST, and serves as an excellent model for how such processes can be structured to function effectively.

### *Recommendations*

1. *We strongly endorse the interaction within different divisions of the same agency and among different government agencies in areas where optimization of resource allocation and joint planning and development are appropriate. Such interactions should be fostered and encouraged.*
2. *The Subcommittee specifically recommends continuation/initiation of cooperative activities among the federal agencies in the following areas:*
  - *synchrotron beamline operations and constructions*
  - *x-ray and neutron detectors and other specialized instrumentation for structural biology applications*
  - *high-speed network links between synchrotron beamlines and research institutions*
  - *high field magnet development for NMR and NMR user facilities*

- *mass spectrometer instrument development*
  - *neutron scattering instrumentation and beamlines*
  - *structural genomics including its computational structural biology component*
3. *Five specific areas where joint agency planning is important are:*
- *the new beamlines on the cold neutron source at HFIR*
  - *the upgrade of SSRL (SPEAR3)*
  - *adequate operating support for CHESS*
  - *the requested upgrades for NSLS*
  - *the development and operation of general user crystallography beamlines at the six synchrotron x-ray facilities*

These facilities either currently have, or will develop, the potential to support an increasing presence of structural biology research. We encourage BER to work together with BES and NIH and NSF to see that the significant potential of these resources is achieved. This recommendation is particularly important in light of the growing structural biology user community at the synchrotrons, the diminishing importance of neutron crystallography discussed in other sections of this report, and the rising interest in x-ray and neutron scattering from non-crystalline systems.