PSI:Biology Evaluation Team Report

Introduction

The Protein Structure Initiative (PSI) was started in 2000 with the idea that protein structure determination could be made more efficient by developing pipelines for protein expression, crystallization, and structure determination by both X-ray and NMR techniques. Part of the goal of higher throughput structure determination was to increase the coverage of "fold" space with the goal of bringing every protein sequence within the realm of modeling based on sequence homology. The first phase from approximately 2000 to 2005 (PSI I) was devoted to developing pipelines for gene synthesis, protein expression, protein purification, and crystallization. A second phase running from 2005 to 2010 (PSI II) was devoted to using the pipelines to turn out as many structures as possible with four high-throughput centers contributing more than 4000 structures to the Protein Data Bank during that period. As part of PSI II modeling played a significant role in coordinating target selection and improving homology-based modeling and the use of new structures. In 2010, the focus of PSI was changed from determining as many structures as possible to focusing on more biologically or medically relevant proteins; for example human, or at least eukaryotic homologues of human proteins, and traditionally more difficult structures for example protein/protein complexes and membrane proteins. This third phase of PSI, titled PSI:Biology, was altered so that in addition to high-throughput centers (HTC), grants for biology partners were created and a separate set of grants for centers to specialize in membrane protein structures were created. The biology partners were to bring important biological problems that would benefit from structural insight to the HTCs and then use the structural information obtained as the basis for further work in understanding the functional and mechanistic aspects of the biological system. In addition, to broaden the access to the technology created by the HTCs, a system of Community Nominated Targets (CNTs) was created, where researchers could nominate targets to be picked up by the HTCs. In addition two research resources that were integral to PSI I and II were continued – the Structural Biology Knowledge Base (SBKB) and the PSI:Biology Materials Repository (MR).

The goals of this mid-point evaluation of PSI:Biology were:

- To assess how well the NIGMS-funded PSI:Biology program is performing based on its first three years by examining its progress toward the goals of the program and the impact of the program in leveraging structural information to enhance studies of protein function.
- To consider potential adjustments to the PSI:Biology program that would strengthen the impact of structure determination in the broader biological community.

Overview of the Evaluation Process

The evaluation team consisted of:

- Steven Sheriff, co-chair, Bristol-Myers Squibb Research & Development
- Judith Bond, co-chair, Emeritus Professor Pennsylvania State University School of Medicine, President, FASEB
- Jay Dunlap, Geisel Medical School at Dartmouth
- Jacquelyn Fetrow, Wake Forest University
- Millie Georgiadis, Indiana University School of Medicine, Indianapolis
- Lila Gierasch, University of Massachusetts, Amherst
- Klaus Schulten, University of Illinois, Urbana-Champaign
- Stephen Sprang, University of Montana
- Ken Taylor, Florida State University

- David Weber, University of Maryland School of Medicine

Principal Investigators of the High-Throughput Centers (HTCs), Membrane Protein Centers (MPC), Biology Partners, and Research Resources (SBKB and MR) were asked to submit short reports addressing the following areas:

- Highlight of outcomes (HTCs and MPCs to coordinate with Biology Partners to avoid duplication)
- Technology development (HTCs and MPCs only)
- Outcomes of collaborations with Community Nominated Targets (HTCs and MPCs only)
- Structure of the PSI:Biology program (Proposals to make it more effective)
- Publications and grants
- Training Efforts (HTCs and MPCs only)

Teams of ~5 members of the PSI:Biology Evaluation Team visited the four HTC centers to which the associated biology partners and relatively local MPCs were invited. They were asked to address the following questions in their oral presentations:

- What do you see as the main strengths of the PSI:Biology program as it is structured?
- What is working well?
- What is not working and/or where do you see the challenges?
- What strategies could improve achievements and accelerate the process of reaching goals?
- How has the PSI:Biology program provided greater scientific impact than an equivalent program of R01 awards?
- What has been the educational outreach of PSI:Biology awardees to the broader community?

For various reasons, the committee was unable to obtain information from investigators who had proposed targets through the community nomination process and from modelers on the following sorts of questions, which the committee felt would be important in assessing impact of PSI:Biology in the broader sense:

- Are the intellectual and material products of the current PSI:Biology program resources being well used by the community?
- How involved is the broad scientific community in the different aspects of the PSI:Biology program?
- How well does the community-nominated structure determination program meet the needs of the community?
- Are the policies for sharing data and materials with the community effective?

As this information was not available, the committee relied where necessary on derived but imperfect estimates of impact, such as publications, citations, functional annotation analysis, on anecdotal information from site visit participants, and on the collective experience of the Evaluation Team.

Overview of the Perceptions of the PSI:Biology Evaluation Team

The PSI:Biology Evaluation Team was unanimous in the following views:

- An impressive number of high quality protein structures have been determined by PSI-funded investigators (both through HTCs and membrane centers) over the last 13 years. Structure

determination during the PSI:Biology phase continues to be strong, especially from the membrane centers.

- Some of the accomplishments (methodological as well as determined structures) could not have been readily achieved through R01-type investigator-initiated grants. The establishment of high throughput and membrane centers enabled a concentration of resources for efficiency and risk-taking.
- Some reluctance/inability was observed for some of the HTCs and MPCs to think and design experiments in terms of the new PSI:Biology goals of enhancing the impact of their work on protein function, rather than the original goals of solving as many structures as possible.
- Evaluation of the success of PSI:Biology (in terms of leveraging structural information to enhance studies of protein function, and having impact on the broader biological community) was limited by the relatively short time (3 years) this phase of the program has been operating.
- The outreach of PSI:Biology has been inadequate in that it has not reached the broader biological community. Failure of outreach encompasses many areas including making the community aware of the resources available to it through PSI:Biology, providing training in technology and techniques that can be used on a smaller scale than an HTC or MPC, and engaging the broader community in making use of the structures that have been determined.
- In PSI I and II, the Program was not perceived to be an appropriate venue to train students and post-doctoral fellows. Training has not been a high priority of the PSI:Biology Program, and more attention needs to be devoted to this.
- No program should be allowed to continue indefinitely, and this applies to PSI:Biology
- Before PSI:Biology is retired, it should be renewed for a period not exceeding one more term (of 3 to 5 years) with some reduction in outlay and considerable restructuring to help it reach its promise.
- NIH must start planning now for maintaining the parts of PSI:Biology that provide unique capabilities and resources that are not available elsewhere via other funding mechanisms. This principally consists of the pipelines in the HTCs for expression construct design, gene synthesis, protein expression, and purification pipelines.
- Mechanisms for achieving sophisticated technological developments in the area of structural biology will also need to be continued in some capacity from the NIH, so that state-of-the-art research does not stagnate on the development front, particularly for studies of mammalian complexes and/or other medically relevant targets.

An Overview of Strength and Weaknesses of PSI:Biology as Currently Structured

The PSI:Biology program has been very successful in some respects. Much of the research that is being pursued could not be done in an environment of regular individual investigator R01-type grants, because grant evaluation mechanisms generally discourage high-risk research (for example, “no crystal, no grant”). However, NIH has instituted some mechanisms for “high risk, high impact” research, and these are possible vehicles for the work currently done by the Centers. Importantly, PSI centers have had the resources and the mission to develop methods, reagents and technologies to express and crystallize proteins and determine their structures with high efficiency.

The achievements of some membrane protein centers have been impressive. However, the success of these centers has varied considerably. The number of structures of membrane proteins determined by the general community has increased considerably during the time of PSI:Biology, but PSI:Biology has had a particularly significant impact in the number of structures of human membrane proteins. The MPCs have also established good inter-center collaborations and interactions.

The HTCs developed excellent capabilities for high throughput screening and expression of proteins over the course of the last 13 years, and are very valuable resources for structure determination. However many of the projects being developed are technology driven, chosen because they can capitalize on the existing high throughput structure pipelines, rather than being driven by biological interest or impact. Thus some of the centers have had a tendency to continue to focus on the goals of PSI I and II rather than the new goals of PSI:Biology. Tension was observed between some of the HTCs and some of the partners. In many instances, a large number of structures were determined but no evidence was presented that functions or other integrated biological questions are being pursued. While a few efforts have been well received by the scientific community at large, much work remains invisible with the majority of publications reporting structures arising from PSI-supported efforts never having been cited. Moreover, the leaders of the HTCs noted that coordination and interaction between HTCs has diminished since the end of PSI II and with that loss of interaction technology development has slowed down in large part.

The coordination between structural and functional studies needs more attention. Biology partners were all assigned to a center (HTC or MPC) by NIGMS. Nevertheless more success and enthusiasm was engendered when partners and centers communicated before a formal assignment was made. Some dissatisfaction was expressed about projects that resulted from "arranged marriages". Moreover, the role of the biology partners in the implementation of the collaborative research is not clear. Is the role of the partners solely as consultants to direct the selection of targets, or to truly integrate structure-function studies? More than 60% of the biology partners are headed by structural biologists. The reason for this was discussed at several of the site visits. Among the suggestions raised to explain this situation was that it took a structural biologist to act as a broker between the HTCs and functional biologists. However, more input from the biologists is required to achieve the aims of the PSI:Biology.

One of the goals of the mid-term evaluation was to assess the impact of the program in leveraging structural information to enhance studies of protein function. However, convincing evidence was lacking in the written material and at the site visits that the goals of HTC projects were to use structural information to address particular problems or hypotheses in biology. The evaluation team perceived a lack of focus on how the collaboration between HTCs and biology partners should be leveraged to impact the biological problems of interest. This is a most serious weakness of the current program.

The Structural Biology Knowledgebase and Materials Repository were considered very valuable resources. However, their value is mainly to the members of PSI:Biology rather than to the broader biologic community.

The broader biological community is largely unaware of PSI and individuals outside of PSI, who are aware, do not feel invited or empowered to enter into the structure. A later section in this report will focus on possible ways to increase the success of outreach activities.

The question of whether the PSI:Biology program provides greater scientific impact than an equivalent program of R01 awards is difficult to answer at this stage. The HTCs, partners, and the evaluation team all agree that the HT centers provide invaluable resources and some of these capabilities could not be accomplished by an R01 mechanism. However, in terms of innovation it is not clear that centers perform better than individuals funded through R01 mechanisms. Some advances made by PSI centers rested on previous work done in R01-funded laboratories.

Some Specific Recommendations for the Continuation of PSI:Biology

The consensus of the Evaluation Team was that PSI:Biology should be continued for another term after 2015. Opinions varied within the committee about whether that term should be limited to 3 years or to another 5 year term. Also opinions varied within the committee about whether an extension of PSI:Biology should be funded as fully as possible given the likely reductions in the NIH budget due to sequestration and other budgetary pressures or whether reductions beyond that level should be built into this term.

The large scope of the demands on PSI:Biology requires strong leadership. Significant creative effort above and beyond the considerable achievements in structural biology is needed. This includes engaging the scientific community, as a whole, to illustrate: (1) the value of coordinated team science; and (2) to create methods and strategies for incorporating individual laboratories throughout the country in a meaningful way, but without diluting PSI:Biology productivity. Some specific issues that need to be addressed include:

1. Prudently identify scientific priorities that answer biological questions while simultaneously pursuing technological development.
2. Decide what technologies to develop as well as which ones to de-emphasize.
3. Organize how "big science" can effectively partner with individual laboratories throughout the country.
4. Engage the scientific community for assistance in new technological developments, when necessary; this should be done by communicating the technologies that are deemed "high priority".
5. Take an aggressive role to "empower" laboratories throughout the country by disseminating expertise and reagents to the scientific community. Some groups from within the PSI, which are already achieving such success need to be organized into an executive role within the PSI to assure the PSI as a whole become national leaders for expanding the field of structural biology. Their executive goal must be to empower a wider scope of scientists, including biologists, with expertise in structural biology techniques that are appropriate for rapid dissemination.

The committee also recommends that a major reorganization should occur in the next term. Some ideas for reorganization are enumerated below:

1. Restructure the High-Throughput Crystallography centers to focus exclusively on what is unique to them in the academic community – the ability to produce many constructs and attempt to express them in many different systems, i.e. become Multi-construct Protein Production (MCPP) centers.
    a. Technology development would be focused on expression systems – improving them, adding additional known systems to a center's pipeline, or developing novel systems. This is an important and underappreciated aspect of protein biochemistry.
    b. A secondary emphasis on technological development of other complementary high-throughput techniques should be encouraged.
    c. A scaled-down structure-determination service focused on community-nominated targets or limited collaborations with NIH-funded scientists (see point 6 below), could be maintained.
2. Priority setting of projects undertaken by the centers must occur. Certain topics should be emphasized and others de-emphasized. For example,
    a. The search for additional protein folds should be totally de-emphasized.

   b. On the other hand, the study of intact proteins rather than domains should receive more attention and examining mammalian protein-protein interactions, other mammalian complexes, and/or complexes relevant to human health should be emphasized.
3. The themes for centers should be designed around specific biological systems of interacting macromolecules, i.e., "networks" rather than on multiple biomedical themes as they have chosen. Membrane centers should focus on proteins (and their complexes) with common biochemical functions, or that participate in a common pathway, rather than being generalists. Membrane centers could also focus on specific areas of technological development with high potential impact.
   a. The number of membrane centers should be reduced.
   b. Any future membrane centers should be focused on human (or at least eukaryotic) proteins, with exceptions justified on an individual basis relating to biological impact. For example, structures of specific bacterial membrane proteins might be justified on the basis of their importance to antibiotic resistance.
4. The biology partnerships should be restructured:
   a. One possibility would be to limit the number of biology partnerships per center to two each.
   b. Alternatively, or in addition, create a project manager position to coordinate between the biology partner and the center. One of the biggest problems appeared to be a lack of coordination between structural studies done by centers and associated functional studies performed by partners. The project manager would be a person dedicated solely to coordinating the projects. This person would most likely reside in the center and should have the sole responsibility of keeping track of all partner projects. Ideally, a person with both scientific and managerial experience would fill this position.
   c. Biology partnerships should be written in conjunction with an HTC or MPC and should be timed to commence from the beginning of the program rather than phasing in.
5. In lieu of or in addition to "Community-nominated targets", provision should be made to support collaborations between centers and NIH-funded scientists to pursue specific targets of high biomedical significance. These focused collaborations could be developed at any time within the term of center funding. They could be funded by R01 grants or supplements (or other appropriate mechanisms) developed in collaboration with the centers. These grants will be more effective if nominators were permitted (or required) to put "skin in the game" to advance their targets priority. Research proposals that attack barriers to progress, which by definition are risky endeavors, might become more palatable to study sections if the proposal involved the concentration of expertise in the centers. Research grant-funded collaborations between centers and members of the scientific community would also ensure that centers focus on problems of high relevance to the NIH mission, provide a significant means of outreach to the scientific community and, importantly, serve as one mechanism (among others) to sustain PSI centers after the PSI program itself is ended.
6. More emphasis needs to be put on outreach to the broader biological (or at least NIH) community for input in terms of community nominated targets, and perhaps more importantly training to disseminate the knowledge obtained by the PSI as broadly as possible with thoughts towards appropriate scaling of technology at the university or state/region level. In addition, NIH staff should find a mechanism to alert potential grantees about the resources available to individual investigators as a result of activities of the PSI.
7. All funded projects, whether partner grants or centers, should be asked to describe and defend their commitments to training.

Outreach: Comments and Recommendations

The Evaluation Team did not find convincing evidence that PSI:Biology is effectively reaching the broader scientific community. More emphasis needs to be placed on outreach for the remainder of the current phase of PSI:Biology and any future phase. The members/leaders of the HTCs realize that they are unique and have specialized capabilities, but they do not appear to appreciate both their responsibility in sharing and disseminating their knowledge and expertise broadly and the value of effective outreach in making the biology community feel "vested" in PSI:Biology. The members of the HTCs are encouraged to make a commitment to educational outreach, which involves both understanding available tools for outreach as well as personally engaging the broader scientific community.

Currently PSI:Biology's educational outreach is largely limited to workshops and symposia. However, these specialized programs miss large segments of the broader scientific community. It is recommended that workshops include community participants at the planning stage as well as speakers. Featuring biological problems or technologies that are underrepresented in PSI:Biology will attract a wider spectrum of community scientists as speakers and as attendees. Although some HTC staff members have identified biological areas that are underserved within their specific targets or structure families others have not done this so well. Publicizing workshops with prominent speakers from outside PSI:Biology would have the dual effect of bringing in an audience and informing them about PSI:Biology as a research resource or partner. If the workshop is technology centric, then they need a fraction of speakers working in this area that are not in the PSI:Biology. Contact lists developed by tracking speakers and registered attendees would be invaluable in guiding selection of the topics for future conferences, and also be an indicator of broader community interest.

PSI:Biology centers are training their graduate students and post-docs well, but they are not reaching out to serve the education needs of the broader community. Many of the centers do engage in training, but this is on a more one-on-one kind of training, which probably came about through individual contact, i.e. "in reach" rather than outreach. While this is good, the HTCs are encouraged to keep track of whether the trainees were PSI:Biology or from the broader community, and use the information to enhance the contact lists.

A majority of the members of the evaluation team think that the leaders of HTCs might consider formal courses with extensive practical experience as an efficient outreach activity. Training of graduate students and postdocs is the most effective near-term device for engaging the broadest community of investigators not directly involved in PSI:Biology, but at the same time has broad long-range effects through its impact on the next generation of scientists. These courses would have to recruit outsiders to reach a larger number of students, and contact lists from the workshops would be a starting point. For example, Biomedical Technology Research Centers (P41 centers) are required to have courses to train scientists in the application of their new technologies. The courses should be designed to teach techniques that the centers have developed and so should bring in graduate students, research technicians, and post-docs as teaching assistants. Collectively, the HTCs might consider holding two of these courses each year and membrane centers one course per year. The nine membrane centers and the four HTCs collectively have the capacity to cover essentially all of the technologies that are being developed. Whether biology partners should hold a course would depend on the extent and depth they have been involved in technology development. The training that would be of greatest benefit to the structural biology community would be in areas such as protein expression, cell free expression, new crystallization strategies such as lipidic cubic phases and bicelles, and use of software for modeling or structure determination. Moreover, protein expression cuts across all fields and also benefits

biochemists that are not immediately engaged in structure determination. Courses also have the effect of building contacts between the participants better than participation in a workshop or symposium. Centers could expand their contact lists by tracking applicants/attendees from PSI:Biology and the broader community.

Some Mechanisms to Consider for Future Funding of HTC Cores

At some point, whether it is in two years, five years or seven years from now the impact of PSI will be diminishing to an extent that no longer justifies a set-aside funding mechanism. Thus, NIH must start planning now for the successful transition of the PSI from its current set-aside funding structure to some other form of funding for its essential core, which is the Multi-construct Protein Production Centers. One model would be to maintain the best of the current centers under a different funding mechanism. Alternatively, a broader vision might include the creation of additional regionally distributed, but smaller centers.

While the evaluation team welcomes creative new forms of funding for centers that focus mainly on service, the team did want to point to two different currently existing funding mechanisms that might be appropriate for funding the PSI cores.

- Biomedical Technology Research Centers.
    - Biomedical Technology Research Centers (BTRC) (P41 funded facilities) have five basic components: Technological Research and Development (TR&D); Driving Biological Projects (DBP); Collaboration and Service; Training; and Dissemination. One or more of the HTCs and some MPCs might be converted to this sort of funding in a fairly straight-forward way, where, for example, protein expression or membrane-protein crystallization could form the technological basis for the TR&D, Biology Partners could be the Driving Biological Projects, and Community nominated targets could be collaboration and service. As dealt with elsewhere in this report outreach in the form of training and, especially, dissemination have not been the strong points of PSI whereas a P41-like mechanism would foster this commitment.
- A mechanism similar to the high-throughput screening and chemical probes development facility of the National Center for Advancing Translational Sciences (NCATS).
    - The NCATS program solicits projects that could make use of the chemical libraries and expensive infrastructure within NCATS with the idea that these technologies should not be “rebuilt” at every University throughout the country. In this model, after preliminary work with respect to feasibility and some review a formal agreement is developed between NCATS and the Office of Research and Development at the PI’s university, but no money is exchanged.