greatly increase PCI/PSI coverage of the protein complexes and make it more adequate to the task. Naturally, with the advances in the experimental and computational methodologies the structural resolution of the models will be improving, furthering their utility for the scientific community.

The effort offers virtually unlimited perspectives for further development. Other types of complexes (e.g., protein-DNA, protein-ligand, etc.) and types of data (e.g., functional classification and characterization) can be included. The enhancement of the structural resolution and advancement of the experimental/modeling methodology will make possible the description of the dynamic changes in protein structure and the kinetics of protein association, providing a more detailed description of these interactions for deeper insights into the basic principles of life processes at the molecular level.

Conclusion

The large-scale, systematic, community-wide determination and structural characterization of protein complexes will happen regardless of the current decision on the continuation of funding for PSI. It is already happening in other countries and will happen in the United States, simply because it is the direction where the science is going. Arguing against it is like arguing against automobiles in 1890s (saying that horses are a better way of transportation, which I am sure was true at the time) or space exploration in 1960s (with the logic of how many lunches can be provided for the cost of a single flight to the moon). The only issue is whether it will happen now (within a few years) or later down the road with time and resources wasted, progress slowed down, and the quality of biomedical research in the United States and other countries damaged. Therefore, PSI has to live and thrive by significantly increasing its focus on protein complexes.

REFERENCES

Aloy, P., Bottcher, B., Ceulemans, H., Leutwein, C., Mellwig, C., Fischer, S., Gavin, A.-C., Bork, P., Superti-Furga, G., Serrano, L., et al. (2004). Structure-based assembly of protein complexes in yeast. Science *303*, 2026–2029.

Binkowski, T.A., Joachimiak, A., and Liang, J. (2005). Protein surface analysis for function annotation in high-throughput structural genomics pipeline. Protein Sci. *14*, 2972–2981.

Janin, J. (2007). Structural Genomics: Winning the second half of the game. Structure *15*, 1347–1349.

Lijnzaad, P., and Argos, P. (1997). Hydrophobic patches on protein subunit interfaces: Characteristics and prediction. Proteins *28*, 333–343.

Lu, L., Lu, H., and Skolnick, J. (2002). MULTIPROSPECTOR: An algorithm for the prediction of protein-protein interactions by multimeric threading. Proteins *49*, 350–364.

Russell, R.B., Alber, F., Aloy, P., Davis, F.P., Korkin, D., Pichaud, M., Topf, M., and Sali, A. (2004). A structural perspective on protein–protein interactions. Curr. Opin. Struct. Biol. *14*, 313–324.

Topf, M., Baker, M.L., Marti-Renom, M.A., Chiu, W., and Sali, A. (2006). Refinement of protein structures by iterative comparative modeling and CryoEM density fitting. J. Mol. Biol. *357*, 1655–1668.

Tovchigrechko, A., Wells, C.A., and Vakser, I.A. (2002). Docking of protein models. Protein Sci. *11*, 1888–1896.

Vakser, I.A., Matar, O.G., and Lam, C.F. (1999). A systematic study of low-resolution recognition in protein-protein complexes. Proc. Natl. Acad. Sci. USA *96*, 8477–8482.

Wojciechowski, M., and Skolnick, J. (2002). Docking of small ligands to low-resolution and theoretically predicted receptor structures. J. Comput. Chem. *23*, 189–197.

The Soul of a New Structure-Function Machine

Sylvie Doublié[1,*] and Mark A. Rould[2,*]
[1]Department of Microbiology and Molecular Genetics
[2]Department of Molecular Physiology and Biophysics
University of Vermont, Burlington, VT 05405, USA
*Correspondence: Sylvie.Doublie@uvm.edu (S.D.), Mark.Rould@uvm.edu (M.A.R.)
DOI 10.1016/j.str.2007.12.004

It was at the 2001 American Crystallographic Association meeting that we witnessed the first reports from structural genomics (SG) centers and companies. As newly independent crystallographers we had set up our laboratories a mere three years prior. We looked over to the legion of similarly junior colleagues seated around us. We did not utter a word but it was clear we were all thinking the same thing: If these centers and companies can churn out structures that fast, are our small biologically oriented crystallographic labs destined to go the way of the dodo? Would these speed-demons eventually tackle the structural science that we deliberately pursued? We were convinced our nascent research programs were doomed to extinction as less-efficient generators of structural results. We felt like Indiana Jones, running for our lives from the formidable SG rolling boulder.

Fast forward six years: we survived. Sure we occasionally got scooped like everybody else, but mostly by competitors in individual laboratories, not by SG centers. We didn't get overrun because we and the SG centers were by-and-large running in different directions. That said, in what ways have the SG centers had an impact on hypothesis-driven structural research?

In aggregate, SG centers across the globe have been productive, having already deposited over 6000 structures in the Protein Data Bank (Janin, 2007). At these about 2800 are from the NIH-sponsored Protein Structure Initiative (PSI), which aims to provide representative folds for most of protein fold-space (http://www.nigms.nih.gov/Initiatives/PSI/).

It thus appears that the SG centers are accomplishing the task originally put before them. On a per-structure basis, the SG centers seem more cost-effective than individual labs, especially in the more recent ''production'' years of PSI-2. The question is: Are the SG centers generating ''science'' more efficiently than hypothesis-driven structural research? That is, if the NIH dollar for structural work is distributed according to the science generated, is the PSI worth the equivalent 200–250 individual R01 grants that it costs? With renewal of the PSI just around the corner (2010), and in light of the statistic that currently only 8% of R01s are funded on their first try (down from 21% in 1998) (Couzin and Miller, 2007), this is the question of the day.

To be sure, all of us in the crystallographic community have benefited from the advances in protein expression, automation of crystallization screens, robotic handling of crystals at synchrotron beamlines, and development of crystallographic software, all of which were generated by the Specialized Centers of the PSI. In particular, the new generation of software has brought macromolecular structure determination almost to the automated level of small molecule crystallography for some, but not nearly all, macromolecular specimens. This brave new era in which the means to obtain high resolution structures is available to an ever-increasing body of apprentice structural biologists is evidenced by the half dozen or so new structures solved by the 50 novice crystallographers each year at the week-long RapiData Workshops run by Bob Sweet at the National Synchrotron Light Source (http://www.px.nsls.bnl.gov/rapidata2007/). Via the PSI, the structural genomics centers have not only made high-resolution structure determination methods faster and easier, but in doing so have also made the field of structural biology more attractive to many more new and established investigators.

If the PSI is successful in its mission, it will provide the means to approximately model the individual domains of the majority of proteins for which there exists no crystallographic or NMR structure, based entirely on sequence similarity. This might seem an attractive proposition to an investigator whose research is focused on such a protein, but what does it get her? Perhaps most importantly she gains some idea of approximately where individual residues are located within each domain, which ones are likely to be inside the core, which ones are likely to be on the surface, which ones are perhaps close enough to be interacting with each other within the domain. She might even be able to model how the individual domains of the protein could be oriented relative to each other, and thus build an overall model for the protein. It would surely suggest a myriad of experiments, many of which would reasonably be aimed at validating the derived protein model itself. Would anyone, or the study section of any granting agency, be satisfied with such a model, with all of its ''maybes'' and ''likelies,'' in an age when high-resolution structure has never been more accessible by X-ray and NMR methods? It is somewhat ironic that from the PSI were forged the powerful tools that obviate its own existence, by significantly lowering the hurdles involved in pursuing bona fide structural information.

Biological structures are sought in order to assist in the understanding of a protein's function. Given the rapidly-growing list of proteins (identified for example by microarray techniques) whose functions are of intrinsic biological and medical interest, it seems wasteful of limited financial, instrumental, and personnel resources to solve structures simply for the sake of structural information. By itself, knowledge of a protein's structure brings us only marginally closer to understanding its function. Only in combination with the results of other biophysical, biochemical, and genetic experiments is function elucidated. Altogether this suggests that the investigator-initiated collaborative grants which bring together biochemists and molecular and structural biologists are the more efficient mechanism for generating biologically and biomedically relevant results than the PSI. Given the increasing demand for experimentally-derived structural data, the influx of young scientists to the field of structural biology, and the availability of hardware and software to enable the latter to produce the former, all that is lacking is adequate funding to fuel exponential growth not only of the number of protein structures solved, but also of the number of proteins whose function has been revealed. Collectively and inclusively, we are the soul of the new structure-function machine.

REFERENCES

Couzin, J., and Miller, G. (2007). NIH budget: boom and bust. Science *316*, 356–361.

Janin, J. (2007). Structural genomics: winning the second half of the game. Structure *15*, 1347–1349.

