A bittersweet celebration of crystallography

The Protein Structure Initiative will end next year; the aftershocks of this ending should be minimized for the benefit of the broader biology research community.

The United Nations has named 2014 the International Year of Crystallography, in honor of the centennial of the birth of X-ray crystallography. There is certainly much to celebrate about the impact that crystallography has had, especially in providing us with a basic understanding of protein chemistry and helping to decode complex biological processes and to design lifesaving drugs.

The celebration is bittersweet for many in the structural biology field, though. In September 2013, the US National Institute of General Medical Sciences (NIGMS) announced that its Protein Structure Initiative (PSI), a large-scale structural genomics project that has received about $70 million USD each year since 2000, will not be renewed after the current phase comes to an end in 2015.

The PSI was created to solve novel three-dimensional protein structures in a high-throughput manner. The first phase of the project (2000–2005) involved much methodological development to set up efficient structure determination pipelines in specialized centers. The second phase, PSI-2 (2005–2010), focused on high-throughput structure determination to broadly cover protein domain families. In the third phase, PSI:Biology (2010–2015), the emphasis shifted to solving challenging targets such as membrane proteins.

About 90% of PSI targets since 2000 were solved using crystallography, the rest mainly by nuclear magnetic resonance spectroscopy. The PSI has solved more than 6,500 protein structures to date—including notoriously difficult targets such as G protein–coupled receptors—representing about 6.5% of the total number of protein structures in the Protein Data Bank. The PSI also accepts target suggestions from the community; disseminates news, tools, protocols and other useful information through its Structural Biology Knowledgebase (in partnership with Nature Publishing Group); and makes protein expression plasmids available through its Materials Repository.

But despite the concrete successes of the PSI, the biologically exciting protein structures emerging from PSI:Biology efforts and the recommendation of a scientific advisory panel to NIGMS to extend the current PSI:Biology program for an additional 3–5 years, budget cuts and spending reassessments have led NIGMS to announce the PSI's end. An internal US National Institutes of Health committee and an external scientific committee have been created to help NIGMS determine whether and how any of the substantial PSI resources and infrastructure should be preserved.

NIGMS and these transition-planning committees now have the opportunity to set an example for other funding agencies on how to strategically wind down a $1 billion project. The 15 years' worth of infrastructure, high-efficiency pipelines, human expertise and huge quantities of data generated by the PSI should be preserved as much as possible to minimize the negative impact on structural biology research in the United States.

The PSI has been first and foremost about creating resources, tools and methods that benefit the entire structural biology community and beyond. The committees should consider recommending that some of the most productive PSI centers be maintained as community facilities to support individual labs in solving structures of interest. The wealth of information and tools aggregated by the Knowledgebase, and the raw data and metadata amassed by the PSI centers—valuable for data mining, algorithm training and pipeline efficiency, but not available in any protein database—should not be lost.

An underappreciated benefit of the PSI has been its emphasis on solving unique protein folds, which is a key step in determining unknown functions and also enables homology modeling from related structures. Individual investigator grants, however, are usually given to projects that study proteins of known high biomedical interest, with the unfortunate result that a vast number of proteins with currently unknown structures and functions are ignored. The NIGMS committees should ponder whether a systematic effort to broadly sample protein folds could be continued on a smaller scale.

Although critics of the PSI have continually asserted that innovative research is best done in individual labs, 'big science' can be very effective for confronting challenging research problems. In order to solve the structures of large protein complexes, hybrid methods—which may include crystallography and nuclear magnetic resonance, as well as electron microscopy, mass spectrometry and computational modeling—are needed. Individual labs typically do not have the resources or expertise to use multiple techniques, let alone develop new integrated methods. The committees should recommend that NIGMS continue to strongly facilitate team research to tackle particularly challenging structural biology problems.

As part of the Year of Crystallography, we should celebrate the many contributions of the PSI in advancing structural biology. We hope that NIGMS will ensure that its termination is implemented as painlessly as possible.