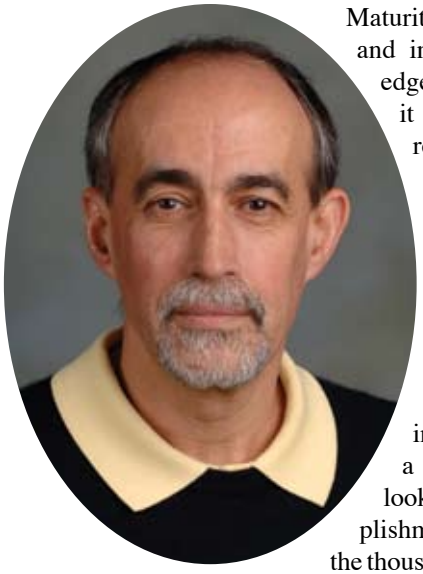
Notes of a Protein Crystallographer: Quo vadis Structural Biology?



Maturity in our personal lives and in science is a double-edged sword. On one side, it is quite satisfactory to reach the middle state in our lives with a sense of accomplishment and pride and look ahead to next stage. Similarly, it is comforting to see the young and revolutionary science that structural biology was in the early sixties reach a point of maturity, and look around at its accomplishments as represented by the thousands of macromolecular structures deposited at the PDB. More important is to examine the critical insights that these structures have provided into in all branches of biology, chemistry, medicine and drug discovery. However, the question is inevitable: What lies ahead? Is it a calm and subdued middle age going to be followed by death, or will there be a rejuvenation and rebirth? Will the future of structural biology lay dormant within the many branches of science that it has helped to advance (biochemistry, cell biology, medicine and others), or will it experience a rebirth by developing new methods to explore the complexity of the living organisms?

This issue has been explored in the last few years by several members of the community. How far and deep we have been able to penetrate into the molecular machinery of biological systems at the beginning of the 21st century, from Vesalius to Palade and Perutz has been insightfully reviewed (Harrison, 2004). After the anatomical discoveries of the renaissance, the structural cell biology tradition of Palade in the first part of the 20th century extended naturally into the structural molecular biology represented by Perutz that we practice today. Harrison's analysis is thoroughly well reasoned and compelling suggesting that the fusion of 'structural molecular biology' and 'structural cell biology' will provide an extended framework for the understanding of biological systems in the next decade. He discusses the roles of structural genomics and computational modeling in that context (Harrison, 2004). This suggested fusion of the two structural traditions represented by Perutz (molecular) and Palade (cellular) will undoubtedly aid in a better understanding of certain biological processes.

The impact of more traditional, well-focused, and slower (i.e. systems-oriented) approaches to discovery in relation to the high-throughput, more expedient (i.e. discovery-oriented), structural genomics strategy was discussed in more detail by Stevens soon thereafter (Stevens, 2004). More recently, Dauter has superbly reviewed the current state and prospects of macromolecular crystallography with a detailed review of the methods and techniques currently in use and the ones that will be appearing in the near future. Both Stevens and Dauter seem confident that the two approaches (high-throughput and specific focus) will continue to provide a constant stream of macromolecular structures that will continue to add to our databases of biological structures and will expand our understanding of living systems (Stevens, 2004; Dauter, 2006). Will this be enough?

I am skeptical that the simple 'structural' extension from molecules to cells will provide the full answers to the complexities of biological systems. A recent essay has been published (Abad-Zapatero, 2007) that provides a historical and scientific context to support this viewpoint. What else do we need? I think that what we need is to the put the living systems within the proper set of physico-chemical principles under which they operate. What is the conceptual framework that encompasses these open, highly heterogeneous and complex systems? The technical term is *dissipative structures*. The term was coined by R. Landauer in 1961 but has been studied, analyzed, and disseminated in the scientific literature by the work of the late Prof. Prigogine (1917-2003) and his coworkers at the Free University of Brussels and the University of Texas at Austin.

In the end, it is the interplay among the conservative molecular entities that we study by single-crystal diffraction methods and the dissipative structures that these molecules make possible that results in the magic of life. This broader conceptual framework suggested above will help us put all this information in the context of systems biology. The concepts of non-equilibrium thermodynamics and dissipative structures have to enter into the domain of modern structural biology if it is to proceed to the next level of understanding. These are concepts that go beyond the commonly accepted notions of intermolecular interactions (be it protein-protein, or protein-nucleic acids) because they include the ideas and notions of flows (fluxes) of matter, energy and information and the sharing of metabolites and chemical intermediates as effectors or facilitators of those interactions. New generations of structural biologists should be introduced to these concepts so that little by little they percolate into the fabric of structural biology and form a part of its intellectual framework. This extension should bring the methods, techniques and *modus operandi* of biochemistry back to the forefront in a novel and more comprehensive way.

Biochemistry is important and I do share the view expressed recently by Arthur Kornberg and others that biochemistry matters "because it does something that genomics, proteomics and other 'omics' cannot yet do" (Kornberg, 2004). As he argues, in the past we have used *in vitro* cell-free systems to gain insights into fermentation, transcription, translation and so many other biological processes. What are those 'cell-free systems' but stable dissipative structures that we can control, manipulate and study their inputs and outputs to infer their complex behavior? We need many more of those self-sustaining systems to gain a deeper understanding of the subtleties of biological systems. This has also been suggested by Harrison (Harrison, 2004) to understand processes ranging from clathrin coating to the motions of the mitotic spindle and beyond. Using the sophistication and experience of traditional biochemists, we need cell-containing or cell-free systems to assay processes such as various biological oscillators,

biological clocks, kinase cascades, cell replication and robust, reproducible and self-sustained signal-transduction systems as well as many other critical biological processes that we do not understand yet at the molecular or cellular level. We may understand the 'parts' but the 'whole' still eludes us.

The use of the concepts and methods of non-equilibrium thermodynamics will aid in understanding the stability, dynamics and control of these open thermodynamic systems and in the design and implementation of new ones. This will open doors to a better understanding of the results obtained by genomics, proteomics and any other 'omics' that we might invent, and will extend to true 'systems biology'. Systems biology modeling should be more than the catalog, description and computer modeling of interactions, no matter how intricate (Giot *et al*., 2003). It should include the detailed spatial and temporal mapping of all components, interacting forces and corresponding fluxes acting on the system. Steven Strogatz, a well known mathematical biophysicist has expressed this idea very concisely: "Our models of complex systems will never advance beyond caricatures until we can find a way to infer local dynamics from data" (Strogatz, 2002).

The insights and understanding gained within this expanded framework will take us from the detailed study of the individual parts at the molecular and pathway level into the true meaning of systems biology, well beyond the simple notion of protein-protein interactions or even protein-nucleic acid interactions (Giot *et al*., 2003). It is conceivable that by expanding our vision of structural biology to include stable, fully integrated dissipative structures, we could open the door to understanding the deregulation existing in the multitude of pathologies associated with cancer, immune disorders, depression and others complex diseases for which our knowledge is still rather limited.

References: Abad-Zapatero, C. (2007). *Acta Cryst*. **D63**, 660-664, Dauter, Z. (2006). *Acta Cryst*. **D62**, 1-11., Giot, L., Bader, J., Brouwer, C., Chaudhuri, A., Kuang, B., Li, Y. et al. (2003). *Science* ***302***, 1727-1736., Harrison, S. (2004). *Nat. Struct. Biol.* **11**, 12-15., Kornberg, A. (2004). *Nat. Struct. Biol*. 11, 493-497, Stevens, R. (2004). *Nat. Struct. Biol*. 11, 293-295, Strogatz, S. Fermi's "Little Discovery" and the Future of Chaos and Complexity Theory. In *The Next Fifty Years. Science in the First Half of the Twenty-First Century.* p. 121. Edited by John Brockman. Vintage Books. A Division of Random House Inc. New York. 2002.

Cele Abad-Zapatero

"Molecular anatomy will be the foundation of medicine in the 21st century , as was human anatomy five centuries earlier […]. In as much as synchrotron radiation is the primary means by which large scale biological structure information will be obtained in the future, continued support is of utmost importance"

Arthur Kornberg

John W. Backus- Father of Fortran - (1915-2007)

In March a note posted to the CCP4 news group stated that John W. Backus, the 'father of Fortran', had passed away.Backus assembled and led the IBM team that created Fortran (short for Formula Translator) which was first released in 1957. The NY Times obituary (March19, 2007) stated that Mr. Backus and his youthful team, then all in their 20s and 30s, devised a programming language that resembled a combination of English shorthand and algebra. It was very similar to the algebraic formulas that scientists and engineers used in their daily work. With Fortran they were no longer dependent on a programming priesthood to translate their science and engineering problems into a language a computer would understand. His team was an eclectic bunch that included a *crystallographer*, a cryptographer, a chess wizard, an employee on loan from United Aircraft, a researcher from the Massachusetts Institute of Technology and a young woman who joined the project straight out of Vassar.

* *

A second note posted to the news group by Bob Sweet included the following: I'm pretty sure that the crystallographer was David Sayre, known in crystallography for the Sayre's Equation (*Acta Cryst*. 5, 60-65 (1952) a fundamental relationship in direct methods. Also, maybe not so well known, he has been a major driving force behind the method of visualizing single molecules or cells from diffraction patterns: (J. Miao, H. N. Chapman, J. Kirz, D. Sayre and K. O. Hodgson, Taking X-ray Diffraction to the Limit: Macromolecular Structures from Femtosecond X-ray Pulses and Diffraction Microscopy of Cells with Synchrotron Radiation, *Ann. Rev. Biophys. Biomol. Struct*. **33**, 157-176 (2004).)

He and I used to use adjacent darkrooms at the NSLS for developing x-ray films (the '80's). I'd meet him on the long walk, ask what he was doing, and smile sympathetically when he said he was going to image single yeast cells. Well, they're essentially doing it now. One never wants to underestimate David Sayre's ability to find phases.

(Editors note: David spoke on this project at the IUCR Congress in Florence where he described the technique used in the 2D imaging of the yeast cell (Acta. Cryst **A62***, 248-261 (2006) and he is now working on extending this to 3D imaging).*

* *

To test Bob Sweet's memory the editor followed up with a note to David asking if he really was that crystallographer and was extremely pleased to receive the following:

In 1954 Peter Friedlander and I were at the Johnson Foundation of the University of Pennsylvania, working on the structures of benzanthracene and 7,12-dimethylbenzanthracene (7,12-DMBA), hoping to cast some light on 7,12-DMBA being a much stronger carcinogen than benzanthracene itself. Peter, working on benzanthracene, was finding a planar polycyclic structure, but 7,12-DMBA, which at that time was generally thought of by chemists as also being planar, was showing signs of crowding of the methyl groups and non-planarity. Wishing to check further, we decided to see whether least-squares 3D refinement would confirm this difference in the structures. At that time the only 3D least-squares program was Durward Cruickshank's program for the MADAM computer in Manchester England, but there was a closer machine, an IBM