Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Notes of a protein crystallographer: *quo vadis* structural biology?

Cele Abad-Zapatero‡

Department of Structural Biology, Abbott Laboratories, Abbott Park, Illinois 60064-6098, USA

‡ Present address: Center for Pharmaceutical Biotechnology, University of Illinois at Chicago, 900 So. Ashland St m/c 870, room 3152, Chicago, IL 60607-7173, USA. E-mail: caz@uic.edu.

Correspondence e-mail: xtalp1@aol.com

The question has been asked many times, ever since grade school and all through high school and college. It is a question that people in the streets accept as an important theme of the scientific discourse: what are the states of matter? The answer has always been deceptively simple: solid, liquid and gas. Everybody knows that! More scientifically sophisticated responders might add 'plasma' as a different configuration of material elements to yield a distinctly different state of matter. To this worn out answer one might reply: where do we include the living systems? Where do we include the myriad of living organisms that populate our air, soil, land and oceans? Explicitly none of these living forms is included in the three classic states of matter. They are neither totally solid nor liquid, nor are they completely gaseous. Traditionally, none of the well recognized states of matter includes the biosphere. By everyone's admission, living organisms are not solid, liquid or gas. They are a form of matter that includes all the other three accepted categories and then something else. They constitute the 'living state' of matter. Those two words connected are part of the title of an intriguing book by Albert Szent-Györgyi, the Hungarian biochemist who won the Nobel Prize in Physiology or Medicine in 1937 for his work on vitamin C. Late in his life, he searched for answers to the cancer puzzle in the intricacies of quantum mechanics (Szent-Györgyi, 1972).

Our body is given consistency by a series of solid structures that include bones, a well recognized solid entity that outlasts our own existence. Another key component of our bodies is a reddish solution of a large variety of complex chemicals that we refer to as blood. It carries nutrients to the nooks and crannies of our body and removes the unwanted waste products to maintain a healthy state. We inhale and exhale air, an accepted gaseous mixture of oxygen, nitrogen and other less important components. It is stored in our lungs and distributed to various cavities by a network of blood vessels and capillaries to maintain the metabolic processes required for life itself. Similar conclusions can be drawn from the analysis of other organisms. Physically and chemically these appear to be all the components of the living systems. What else do living organisms have that makes them living?

The existence of something else to explain the uniqueness of the living organisms has been part of the scientific discourse since the early ruminations of the human mind on the subject by the Greek school. External forces ranging from 'spirit' to 'soul' to 'élan vital' have been invoked to provide the additional component or element that will transform inanimate matter – admittedly composed of only solid, liquid and gaseous components – to living matter. So what is missing? The plain answer is that nothing is missing, absolutely nothing. But then, what is the difference between a dead cat or a dead cell and a jumping cat or a living cell?

Before I provide an answer to the question in terms of the current concepts of physical chemistry, I would like to take a brief historical digression. In my early days of intellectual curiosity regarding molecular biology, I remember reading the writings of Max Delbrück (1906–1981), the German-born biologist who together with Salvador Luria (1912–1991) and Alfred Hershey (1908–1997) were the founding popes of the 'phage group'. This group has been considered by some as one of the primary seeds for the origins of molecular biology as a field of study. Its members were trying to understand the nature of genes by studying the replication of phages. Delbrück has also been recognized as the unofficial mentor of James D. Watson, who had been introduced to the phage 'principles' earlier by Luria. In his early writings, Delbrück expressed the view that when a physicist looks at biology,

> This analysis should be done on the living cell's own terms and the *theories should be formulated without fear of contradicting molecular physics*. I believe that it is in this direction

> physicists will show the greatest zeal and will create a new intellectual approach to biology which would lend meaning to the ill-used term biophysics,

(Delbrück, 1949) (italics my emphasis). Within this framework, some historians have interpreted that the early studies on phage genetics and replication by the Delbrück school were driven by the possibility of 'finding new laws' that would apply only to biology and not to the then known chemical and physical sciences. The replication of the genetic material at the molecular level was considered to be a critical factor of the 'secret of life', if not the secret itself. The story is well known. Watson and Crick unravelled the structure of DNA in 1953 and Crick boasted and toasted at the Eagle pub in Cambridge, UK, that they have 'solved the secret of life' (Watson, 1980). The unraveling of the genetic code in the sixties by the insightful use of cell-free systems as a tool to analyze transcription and translation *in vitro* demonstrated once again that there appeared to be nothing revolutionary in the workings of the cell.

Having put those versions of the secret of life behind, Delbrück and other disciples of the phage group looked into the mechanisms of phototaxis and other signal-transduction mechanisms in the fungus *Phycomyces* as the possible hiding place of those putative new laws of nature. As we know now, the search has failed. Living organisms operate by the same logic and rules as the inanimate world. In addition, the continued study of organisms during the second half of the 20th century using the tools of chemistry, physical chemistry, X-ray diffraction and others has established beyond reasonable doubt that

> rather than searching for radically different ways of studying organisms or for new laws of nature that will be manifest in living beings, what biology needs to do to fulfill its program of understanding and manipulation is to take seriously what we already know to be true. It is not new principles that we need but a willingness to accept the consequences of the fact that biological systems occupy a different region of space of physical relations than do simpler physico-chemical systems, a region in which the objects are characterized, first, by a very great internal physical and chemical heterogeneity and, second, by a dynamic exchange between processes internal to the objects and the world outside them. That is, organisms are internally heterogeneous open systems,

(Lewontin, 2000). Assuming this overarching paradigm, 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 thorough, well reasoned and compelling. In brief, he suggests that the fusion of structural molecular biology and structural cell biology will provide an extended framework for understanding of biological systems in the next decade and 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 understanding certain biological processes better. However, I am skeptical that the simple 'structural' extension from molecules to cells will provide the full answers to the complexities of the biological systems expressed above. What else do we need? I think that what we need is to 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 Professor Prigogine (1917–2003) and his coworkers at the Free University of Brussels and the University of Texas at Austin. His efforts were recognized by the Nobel Foundation that awarded him the Nobel Prize in Chemistry in 1977 for his contributions to non-equilibrium thermodynamics and particularly the theory of dissipative structures.

Our stable macromolecules (*i.e.* proteins, nucleic acids and crystals) are conservative structures (or equilibrium structures) in so far as they are stable after their generation and no further energy or matter are required to maintain them. They are results of the process of classical or equilibrium thermodynamics in isolated systems. In contrast, dissipative structures are structures formed and maintained far from equilibrium by a constant flux of matter and energy from outside the system. Dissipative structures can be formed with relatively simple chemical entities (as in the well known Belousov–Zhabotinskii reaction) but their possibilities and richness will undoubtedly be expanded by the presence of the complex and heterogeneous molecules that make the biological milieu.

The work of Prigogine and colleagues has shown that dissipative structures display two types of behavior. In the linear regime and close to the equilibrium conditions, the order that they create tends to be destroyed by the environmental fluctuations and disturbances. However, under far from equilibrium conditions, the order can be maintained and new forms and novel equilibrium states can be explored and stabilized. Critical to maintaining dissipative structures far from thermodynamic equilibrium are the flows of matter and energy from the outside environment. These flows could include not only the well accepted flow of nutrients and water but also the flow of other effector molecules such as ions, signaling molecules and even the inter- and intramolecular flow of phosphate groups. What are the implications of this expanded framework for the future of structural biology? Where should the new generations look for the developments of field in the 21st century?

The current status of the methodology of macromolecular crystallography has been extensively reviewed recently by Dauter (2006). He concludes that the automatic (*i.e.* structural genomics) and investigative (*i.e.* special crystallographic methods) approaches for structure solving will continue to

coexist in parallel in the near future. I have previously discussed what I think would be important developments in the field (Abad-Zapatero, 2002) and the possible impact of robots (Abad-Zapatero, 2005). On the computational side, we should strive for improved refinement methods and a superior description of the structures in relation to the structure factors. Direct methods of protein structure determination in reciprocal space, using the information gathered from the distribution of atoms in space for the known structures, are still not available. Molecular replacement will of course continue to be the 'semi-direct' method of choice for structures for which a structural homolog exist but, are direct methods still an impossible dream? On the experimental side, neutron diffraction (Bennett *et al.*, 2006), time-resolved Laue crystallography (Ren *et al.*, 1999) and diffraction experiments based on fibrous (or partially ordered) samples will continue to provide invaluable data as these methods are applied to a wider range of biological problems (Dickinson *et al.*, 2005).

The introduction of robots to perform some of the repetitive and tedious operations at synchrotron beamlines will expedite the structure solution of the many macromolecular structures still unknown. In addition, it will facilitate the screening for the favorable crystallization conditions necessary to crystallize and solve additional membrane proteins, and ever more challenging protein–protein and protein–DNA/RNA complexes. This will extend the scope of the molecular biology that we practice today and will expand the catalog of known structures available (Jiang & Sweet, 2004). However, it was also suggested that we should develop novel experimental designs to expand our domains of structural enquiry. Experiments that would permit us to gather structural information complementary to the atomic details that we can obtain once our molecules of interest are frozen in a crystal lattice (Abad-Zapatero, 2005). Super-resolution microscopes, more powerful imaging software and even more powerful and effective fluorescent labels have already been suggested by Harrison to bridge the gap between the behavior of the molecular assemblies *in vitro* and within the milieu of the living cell (Harrison, 2004). What else lies ahead? *Quo vadis* structural biology?

Structural biologists should continue to navigate beyond the comfortable confines of single-crystal diffraction methods or even cryo-electron microscopy to see what lies beyond this well known meridian. I would suggest that what extends east and west of this longitude is very attractive indeed. They will find that although apparently sailing in opposite directions, as they travel far enough, they will meet. By combining their complementary structural results our understanding of the 'living state' will be much more comprehensive and detailed.

On the east if you wish, they should extend the diffraction methodology to the possibility of reliable and robust structural information obtained from single-molecule diffraction methods. The imminent access to fourth-generation X-ray sources, referred to as XFEL (X-ray Free Electron Lasers), could make this dream a reality. Reconstruction of the molecular structure based on the diffraction data from one (or a few) differently oriented molecules would have a tremendous impact in structural biology (Miao *et al.*, 2001). Many important biological molecules or complexes that might never be crystallized could unveil their structural secrets using this approach. Very encouraging results are being obtained by the pioneers in the field using 25 fs pulses from the first XFEL sources (FLASH at DESY, Germany) producing soft X-rays (Chapman *et al.*, 2006). The most recent results will be reported at the upcoming 9th International Conference on Biology and Synchrotron Radiation (BSR2007, Programme Overview, http:////www.srs.ac.uk/bsr2007/pages/programme_overview.html).

Looking west, the panorama is broader but no less attractive. The younger generations of structural biologists should consider also what is possible now using other methods of obtaining structural information. Diffraction Enhanced Imaging (DEI) methods are coming of age and what lies ahead looks 'bright' and 'sharp'. These methods allow us to see in superb detail what one could call 'tissular' biology of an increasing variety of biological samples (Chapman *et al.*, 1997). The methodology is being optimized rapidly for softer (living tissues, cells) samples to be able to bridge the gap between molecular and macroscopic biology and the future looks very promising indeed (Majumdar *et al.*, 2004; Yoon *et al.*, 2007).

Moreover, the combination of these imaging techniques with the spectroscopic techniques adapted to micrometre resolution sampling would permit the molecular mapping and distribution of the chemical species present in the biological samples. It is worth noting that large synchrotron sources are not really necessary to obtain this combination of imaging and microspectroscopy structural data. Excellent results using DEI methods to image femoral heads and cartilage–bone interactions have appeared from ELETTRA (Majumdar *et al.*, 2004). Most recently, the Pohang Light Source (PLS) in Korea, has been used to image biological tissues with low absorption contrast such as renal cell carcinoma and prostate cancer (Yoon *et al.*, 2007). Currently, beamlines for mid-IR, far-IR (with and without Fourier transform), spectro-microscopy (SM beamline with 40 nm resolution) are available to users at the Canadian Light Source (CLS). In the near future, a unique biomedical imaging beamline (BMIT) will add to the wide range of imaging techniques that will be available at this 2.9 GeV source with one of the smallest footprints among the existing synchrotron radiation sources (Chapman, 2006). Naturally, these developments are the result of the experience gained at the more mature sources (ALS, APS, ESRF, NSLS, SLS, SSRL, SPring-8).

My personal dream would be to be able to perform microdiffraction and spectro-microscopy experiments with biologically relevant samples with micrometre or better spatial resolution and with time-resolved capabilities. In this way, structural biologists could map in space and time the dissipative structures that are the essence of all biological systems. The diffraction methods will unveil the details of any existing order among the macromolecular components, and the combination of imaging and spectromicroscopy would identify

and map the chemical entities in space and time, under controlled conditions. These comprehensive data would be critical for the theoretical and computational biologists attempting to unravel the complexities of systems biology.

I can refer to a classic example of how to accomplish this in the domain of the well established field of electrophysiology. Alan Hodgkin and Andrew Huxley were able to describe mathematically the neuron dynamics from their measurements in the squid axon (Hodgkin & Huxley, 1952). They kept the voltage across the nerve membrane at various fixed values ('patch clamp' experiments) and then they measured the flow of sodium, potassium and other currents through the membrane as a function of the voltage. These controlled measurements allowed them to derive the precise nonlinear mathematical dynamics of a single neuron. For this pioneer insight into 'systems biology' they were awarded the Nobel Prize in Medicine or Physiology in 1963.

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. The broader conceptual framework suggested above will help us in putting all this information in the context of systems biology. The concepts of nonequilibrium thermodynamics and of 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 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 back the methods, techniques and *modus operandi* of biochemistry to the forefront again 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 the 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 yet understand 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 nonequilibrium 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 the doors to a better understanding of the results obtained by genomics, proteomics and any other 'omics' that we might invent, and extending 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 the components, the interacting forces and the corresponding fluxes acting on the system. Steven Strogatz, a well known mathematical biophysicist has expressed this idea very concisely (Strogatz, 2002),

> our models of complex systems will never advance beyond caricatures until we can find a way to infer local dynamics from data.

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 and 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 would open the door to understanding the deregulation exiting in the multitude of pathologies associated with cancer, immune disorders, depression and others complex diseases for which our knowledge is still rather limited.

I wish to thank Drs J. Greer, K. Longenecker, C. Park, G. Stamper and V. Stoll for their comments on earlier versions of this essay. The constructive suggestions from the anonymous referees and the editor to improve this essay are greatly appreciated.

References

Abad-Zapatero, C. (2002). *Z. Kristallogr.* **217**, 1–3.
Abad-Zapatero, C. (2005). *Acta Cryst.* D**61**, 1432–1435.
Bennett, B., Langan, P., Coates, L., Mustyakimov, M., Schoenborn, B., Howel, E. E. & Dealwis, C. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 18493–18498.
Chapman, D. (2006). Canadian Light Source. Bioimaging group (BMIT). http://www.lightsource.ca/bioimaging.
Chapman, D., Thomlinson, W., Johnston, R. E., Washburn, D., Pisano, E., Gmür, N., Zhong, Z., Menk, R., Arfelli, F. & Sayers, D. (1997). *Phys. Med. Biol.* **42**, 2015–2025.
Chapman, H. N. *et al.* (2006). *Nature Phys.* **2**, 839–843.
Dauter, Z. (2006). *Acta Cryst.* D**62**, 1–11.
Delbrück, M. (1949). *Trans. Connecticut Acad. Sci.* **38**, 175–191.

Dickinson, M., Farman, G., Frye, M., Bekyarova, T., Gore, D., Maughan, D. & Irving, T. (2005). *Nature (London)*, **433**, 330–334.
Giot, L. *et al.* (2003). *Science*, **302**, 1727–1736.
Harrison, S. (2004). *Nature Struct. Biol.* **11**, 12–15.
Hodgkin, A. L. & Huxley, A. F. (1952). *J. Physiol.* **117**, 500–544.
Jiang, J. & Sweet, R. (2004). *J. Synchrotron Rad.* **11**, 319–327.
Kornberg, A. (2004). *Nature Struct. Biol.* **11**, 493–497.
Lewontin. (2000). *The Triple Helix. Gene, Organism and Environment.* Cambridge, Massachusetts: Harvard University Press.
Majumdar, S., Issever, A., Burghardt, A., Lotz, J., Arfelli, F., Rigon, L., Heitner, G. & Menk, R.-H. (2004). *Eur. Radiology*, **14**, 1440–1448.
Miao, J., Sayre, D. & Hodgson, K. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 6641–6645.
Ren, Z., Bourgeois, D., Helliwell, J., Moffatt, K., Srajer, V. & Stoddard, B. (1999). *J. Synchrotron Rad.* **6**, 891–917.
Strogatz, S. (2002). *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*, edited by J. Brockman, p. 121. New York: Vintage Books. A Division of Random House Inc.
Szent-Györgyi, A. (1972). *The Living State: With Observations on Cancer.* New York: Academic Press.
Watson, J. (1980). *The Double Helix. A Personal Account of the Discovery of the Structure of DNA (Norton Critical Edition)*, edited by G. S. Stent. New York: W. W. Norton & Co. Inc.
Yoon, Y., Sung, D. J., Lee, J. H., Kim, A. R., Oh, C. W., Je, J. H., Weon, B. M., Seol, S. K., Pyun, A., Hwu, Y., Margaritondo, G., Joo, K. J. & Yoon, D. K. (2007). *Int. J. Urol.* **14**, 996–1003.