nature structural biology
molecular form & function
november 2001
vol. 8 no. 11

Whither crystallography

The start of the 21st century has seen two great years for protein crystallography. It is difficult to remember when so many long-awaited structures have been solved in so short an amount of time. Indeed one crystallographer (who will remain anonymous) commented following the publication of the structure of the ribosome[1–3]: "it's finished things for us; we might as well all pack up and go home". Not a serious comment perhaps, but is there a nugget of truth within it? What big goals are left for protein crystallography?

Historical milestones

Once it was demonstrated that X-rays diffracting from a protein crystal contain sufficient information to determine its atomic structure, the more complex puzzles were at first relatively obvious: to obtain the structure of a larger protein, of an enzyme–substrate complex, of an intact virus particle, of a protein bound to nucleic acid, of a membrane protein, *etc*. Each new structure provided a specific step forward, each a harder puzzle due to its size, the resolution needed or any number of other complicating factors.

In more recent years, the proliferation of protein structures has made the milestones seem closer together; nonetheless certain results standout — GroEL–GroES, F_1-ATPase, the nucleosome, a bacterial potassium ion channel, photosystem I and the ribosome, to list but a few. The unveiling of such structures provides a moment of revelation, but with only a limited number of unique folds available, such revelations may become less frequent.

Many of the milestone structures owe their status to their size — photosystems and ribosomes are large enough that they can even be seen under light microscopes. Beyond searching for the next 'big' target, perhaps other frontiers are worth exploring.

Space and time

One of the challenges could be to determine structures at ultrahigh resolution. In the Protein Data Bank, only a handful of proteins are solved to below 1 Å resolution — the highest resolution structure of a protein is of crambin at 0.54 Å (ref. 4) — and many structures are reported in the 2–2.5 Å range. While the positions of amino acid side chains are clearly defined at such resolution, for many enzymes, they provide barely enough information to determine the exact mechanisms of catalysis. Certainly no hydrogen atoms can be placed — for that a resolution better than 1.2 Å is needed — but to really get a grip on the chemistry of most reactions, the shapes of the electron clouds, not just the positions of the atoms, need to be accurately mapped.

Even in an ultrahigh resolution structure, the dynamics of a protein, including the gross and subtle conformational changes involved in its function, are usually absent from the picture. However, considerable progress has now been made in catching the short-lived structures that form in various reaction intermediates. These studies typically involve careful manipulation of crystals to synchronize and slow down the reactions so that as much of the reaction intermediate as possible is present during data collection. For example, this approach has been used to study the light-induced change of bacteriorhodopsin (for example ref. 5) and the reaction cycle of isopenicillin N synthase[6].

In the search for ever-increasing resolution, whether spatial or temporal, technological advances are opening up new possibilities. Third generation synchrotron sources bring faster collection of more extensive data sets. What might be achieved with an even more intense source such as a free electron X-ray laser seems almost in the realm of science fiction (see for example refs 7,8).

Integrated approaches

Protein crystallography is also diversifying to form partnerships with other structural techniques. The fitting of crystallographic structures into cryoelectron microscopy-derived densities is now almost common place (see for example refs 9,10). Atomic force microscopy has been coupled to crystallography with great success (see for example ref. 11). Two papers in this issue of *Nature Structural Biology* demonstrate how NMR studies complement structures determined by crystallography[12,13]. Indeed the whole field of structural genomics is an integration of structural methods on a grand scale.

It has never been true that the justification for solving a protein structure was, as for conquering Mt. Everest, because it was there. The form illuminates the function. Whatever the technical achievements or the complexities of the puzzles to be solved, we believe that the reason for being interested in the structure of a protein, or indeed any biological molecule, is to understand how it performs the tasks for which evolution has shaped it. The crystallographers' credo that you cannot understand a protein's function without knowing its structure now pervades all of molecular and cellular biology.

Looking forward from what has certainly been an exceptionally exciting couple of years, it can be emphatically said that this is not the time for packing up and going home. There are huge goals for crystallographers to chase in all directions.

1. Ban, N., Nissen, P., Hansen, J., Moore, P.B. & Steitz T.A. *Science* **289**, 905–920 (2000).
2. Wimberly, B.T. *et al. Nature* **407**, 327–339 (2000).
3. Yusupov, M.M. *et al. Science* **292**, 883–896 (2001).
4. Jelsch, C. *et al. Proc. Nat. Acad. Sci. USA* **97**, 3171–3176 (2000).
5. Edman, K. *et al. Nature* **401**, 822–826 (1999).
6. Burzlaff N.I. *et al. Nature* **401**, 721–724 (1999).
7. Hajdu, J. *Curr. Opin. Struct. Biol.* **10**, 569–573 (2000).
8. Miao. J., Hodgson, K.O. & Sayre, D. *Proc. Nat. Acad. Sci. USA* **98**, 6641–6645 (2001).
9. van den Ent, F., Amos, L.A. & Lowe, J. *Nature* **413**, 39–44 (2001).
10. Simpson, A.A. *et al. Nature* **408**, 745–750 (2000).
11. Smith, D.E. *et al. Nature* **413**, 748–752 (2001).
12. Chou, J.J., Li, S., Klee C.B. & Bax, A. *Nature Struct. Biol.* **8**, 990–997 (2001).
13. Mulder, F.A.A., Mittermaier, A., Hon, B., Dahlquist, F.W. & Kay, L.E. *Nature Struct. Biol.* **8**, 932–935 (2001).