JENNIFER L. MARTIN and NATHAN P. COWIESON

Protein structure and function determination using synchrotron radiation

The opening of the Australian Synchrotron heralds an exciting new era for Australian structural biology research.

Key molecules of life

Proteins carry out vital biological functions as diverse as energy production, defence against deadly bacteria, biosynthesis of complex chemicals and nanosecond neurotransmission. But proteins can also represent formidable opponents in the ongoing battle with disease. Proteins are, at least in part, responsible for the morbidity and mortality linked to the uncontrolled growth of cancerous tissues, the worldwide epidemics of viral infection, the venoms of the globe's nastiest critters, and the devastating metabolic muddle associated with diabetes.

The human genome encodes approximately 30 000 proteins. Each must be expressed at the right time, in response to the right signal and in the right amounts to ensure that the balance in an individual tips towards health rather than disease. What's more, proteins can be modified in a number of ways (phosphorylation, glycosylation, sulfation) and at multiple positions, giving rise to a mind-bogglingly complex array of biomolecules. How do we even begin to go about understanding the function of individual proteins in such a complex soup? One vital approach is to investigate protein structure, allowing us to pinpoint functional 'hot spots', show structural changes required for activity and classify protein function based on fold.

Jennifer L. Martin and **Nathan P. Cowieson** are at the Institute for Molecular Bioscience, University of Queensland, Brisbane, Queensland.

Protein structure

Proteins exhibit an astonishing range of structural complexity. At the most fundamental level, the primary structure is characterised by the amino acid sequence of the polypeptide chain. Secondary structure describes local arrangements of the polypeptide backbone (α-helices and β-sheet) that are stabilised by backbone hydrogen bonds. Generally, 80–90% of a protein chain adopts a conformation consistent with one or the other of these two local types of structure. Secondary structure elements 'fold' together and the arrangement of secondary structure elements in a protein is the tertiary structure. The SCOP database (scop.mrc-lmb.cam.ac.uk/scop/index.html), one of several that describes experimentally observed protein folds, currently numbers at about 1000 (June 2007). The total number of folds is estimated to be between 2000 and 10 000, so there is still some way to go before the protein fold 'universe' is mapped. Finally, protein molecules often interact with other protein molecules to form oligomers and large, multi-subunit complexes. This last level of organisation is referred to as quaternary structure.

Synchrotrons

Synchrotrons are powerful tools for studying protein structure, by providing intense radiation across a broad wavelength spectrum. Synchrotrons were first applied to the study of proteins in the 1970s, and through subsequent years the quantity and range of techniques for protein analysis have increased dramatically. Today, synchrotron radiation is used, for example, for near- and far-edge x-ray absorption experiments to probe the coordination of metals in proteins; protein x-ray crystallography (PX) to resolve the 3D coordinates of protein atoms; synchrotron radiation circular dichroism (SRCD) to yield information about protein secondary structure; wide, small and ultra-small angle x-ray scattering to inform on the size and shape of proteins in solution; and powder, polycrystalline and fibre diffraction to identify low-resolution features in protein structures. Three of the most widely used techniques, SRCD, PX and small angle x-ray scattering (SAXS), are discussed here. Between them, these three approaches cover secondary, tertiary and quaternary protein structure.

SRCD

Circular dichroism (CD) spectroscopy uses polarised light in the UV range, which is absorbed by proteins. Using radiation from UV lamps, the technique is commonly applied to assign percentage of secondary structure, to demonstrate that a protein is folded and to investigate protein structural changes in response to environmental changes such as temperature. Replacing the traditional UV lamp with the intense, broad-spectrum light of a synchrotron (Fig. 1) leads to dramatic improvements in the signal-to-noise ratio, allowing analysis of low-volume or low-concentration samples and measurement of very subtle changes in protein structure. For example, when the protease carboxypeptidase A is complexed with the protease inhibitor latexin, there is

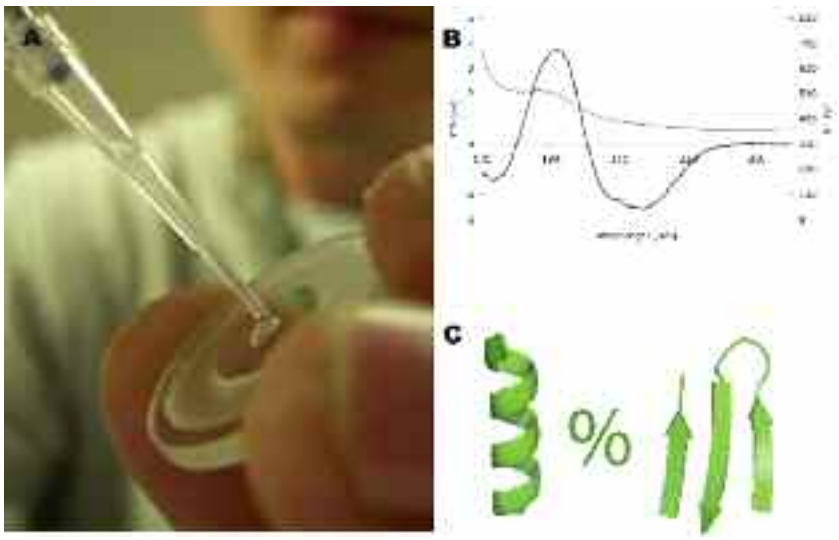

Figure 1. SRCD. **A.** A protein sample being dispensed onto the surface of a cuvette for analysis by SRCD. **B.** SRCD spectrum of a mammalian protein. The darker line shows the CD signal and the lighter line shows the HT voltage. HT can be thought of as gain compensating for absorption by the sample. When HT becomes exponential, as seen at the far left of the plot, there may be insufficient photons for accurate CD measurement. **C.** Schematic of protein secondary structural elements α-helix and β-sheet indicating that the relative proportion of these can be determined from CD data.

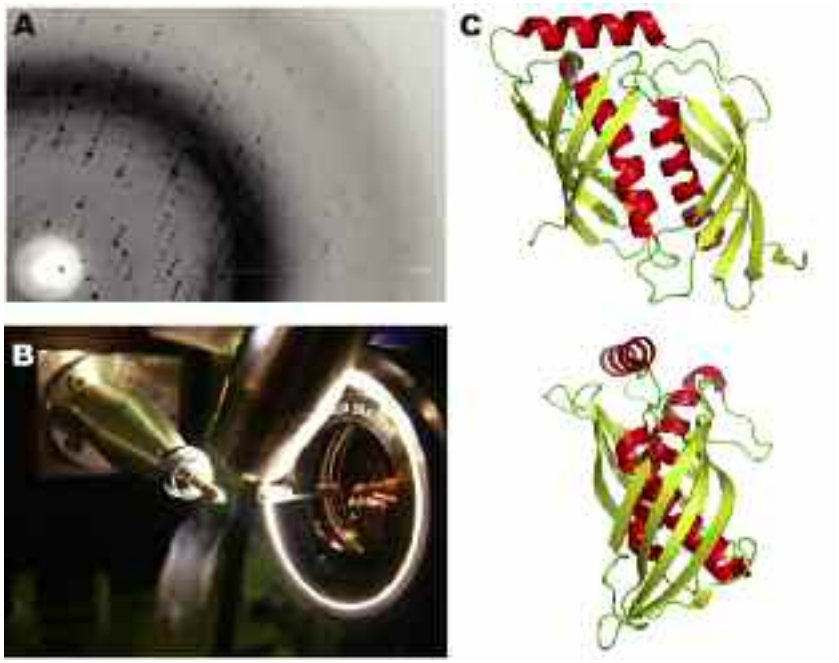

Figure 2. PX. **A.** X-ray diffraction image from a protein crystal. **B.** The sample environment at PX beamline 23-ID-B at the Advanced Photon Source synchrotron in Chicago. In this set-up, x-rays pass from left to right through the sample and then through a hole in the sample visualisation optics at the right of the photograph. The protein crystal is mounted in a small fibre loop on the end of the metal arm protruding from the rear of the image. The crystal is cooled to 100 K in a stream of nitrogen gas from the nozzle pointing downwards from the top of the image. **C.** Perpendicular views of the 3D x-ray structure of latexin,[4] solved using synchrotron radiation from the Advanced Light Source in Berkeley.

no change in secondary structure in either protein. However, the SRCD spectrum reveals changes that are attributable to complex formation. These subtle changes are not visible by conventional CD due to poorer signal-to-noise and lack of light flux at low wavelengths.[1] In addition, synchrotron radiation yields more information-rich CD spectra. The lower wavelength data in particular could be used as a fingerprint region to cluster proteins of known structure, and to assign the structure and function of novel proteins based on their SRCD spectra.[2]

Protein crystallography

One of the most complex proteins ever characterised structurally is the 650 kDa dimeric multi-subunit bacterial photosynthetic oxygen evolving centre. It comprises two copies each of 19 different polypeptides as well as nine different ligands and prosthetic groups.[3] The structure was determined by PX, allowing us to understand how the protein uses light energy to drive the oxidation of water, a fundamentally important process in bacteria. This analysis would not have been possible without a synchrotron, in this case the Swiss Light Source. The tiny protein crystals and the weakness of the diffraction data precluded measurement by anything other than a synchrotron radiation source. Indeed, although other methods like NMR and electron microscopy can be used to determine protein structure, the combination of PX and synchrotron radiation (Fig. 2) is by far the most successful method. Over 85% of about 44 000 structures deposited in the protein data bank were determined by x-ray crystallography and most of these cited synchrotron as a key word (RCSB Protein Data Bank, June 2007).

PX beamlines have evolved in the last decade to increase throughput of samples or to specialise on difficult cases. High-throughput (HT) beamlines typically allow users to

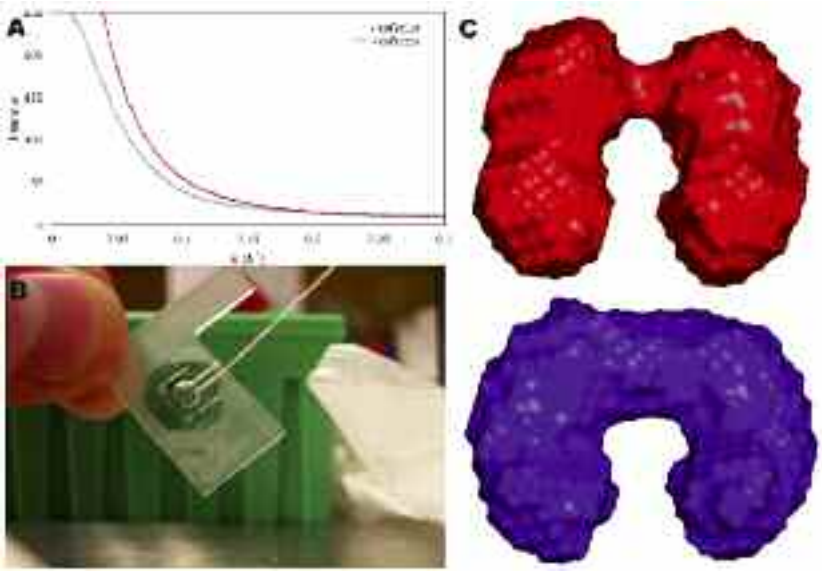

Figure 3. SAXS. **A.** Solution SAXS data from a mouse protein, with (red) and without (blue) a cofactor. The intensity (Y-axis) of scattered radiation decreases exponentially from the direct beam position (X-axis) at the left hand side of the graph to larger scattering angles at the right hand side. **B.** Photograph of a protein sample being dispensed into a cuvette for analysis by synchrotron SAXS. **C.** Envelope shapes of the two samples of the mouse protein (coloured as per panel A) calculated from the SAXS data.

change (tune) the radiation wavelength to facilitate structure determination by single or multiple anomalous diffraction (SAD/MAD) techniques. Fast charge-coupled device (CCD) detectors and robotic sample mounters increase the speed of data collection and in some cases allow users to measure data remotely. This facilitates HT projects such as structural proteomics or fragment-based screening. Other beamlines are optimised for small or weakly diffracting crystals, providing intense x-ray beams with cross-sections in the 5–30 μm range.

SAXS

Like CD, SAXS (Fig. 3) is a solution-based technique that does not rely on crystal formation. Furthermore, it is a dynamic approach that can be used to follow changes in protein structure over time. When X-rays are passed through a solution containing protein particles, the majority of the radiation passes through the sample or is absorbed. A small fraction of the beam, generally less than 1%, is scattered away from the incident beam. SAXS uses the radiation scattered in the angular range 0.1–10°, which gives information on the size and shape of the protein molecules in solution. This low-resolution shape data is a powerful means of determining the tertiary structure of proteins or the quaternary structure of protein complexes, visualising conformational changes, providing low-resolution structure factors to aid in the analysis of PX or electron microscopy data and for validating molecular modelling.[5] The benefits of SAXS at a synchrotron are similar to those for SRCD: high flux improves signal-to-noise, allowing faster measurements from smaller samples.

The Australian Synchrotron

The Australian Synchrotron will have nine beamlines initially, two earmarked for PX. Beamline 1 is a general purpose MAD PX beamline designed for HT, and will be useful for expert and non-expert alike. Beamline 2 is a fine-focus, high-flux, high-brilliance PX beamline intended for use by experts working on the most difficult problems, including membrane proteins and macromolecular complexes. A SAXS/WAXS beamline (beamline 4) will be a valuable resource for the structural biology community, and an SRCD beamline has been proposed as a high priority in the decadal plan.[6]

Summary

Synchrotron radiation provides a powerful toolkit of techniques for structural biologists to address questions of protein structure and function. These techniques can provide information at each level of protein structure from secondary to quaternary, they can be tuned to focus on ligand interactions and they can make static or dynamic measurements. The brightness of synchrotron light allows measurements to be made with a speed and sensitivity that is simply not possible with other light sources.

Acknowledgments

JLM is an NHMRC Senior Research Fellow and NPC is an ASRP Synchrotron Postdoctoral Fellow.

References

1 Cowieson N.P., Miles A.J., Robin G. et al. *Proteins: Structure, Function and Bioinformatics* 2007 (in press).
2 Wallace B.A., Janes R.W. *Curr. Opin. Chem. Biol.* 2001, **5**, 567–71.
3 Ferreira K.N., Iverson T.M., Maghlaoui K., Barber J., Iwata S. *Science* 2004, **303**,1831–8.
4 Aagaard A., Listwan P., Cowieson N. et al. *Structure* 2005, **13**, 309–17.
5 Lipfert J., Doniach S. *Ann. Rev. Biophys. Biomol. Struct.* 2007, **36**, 307–27.
6 Nugent K.A., Goodwin M. (compilers). *Australian and New Zealand Synchrotron-Based Science Strategic Plan 2007–2017*. Australian Synchrotron and Australian Synchrotron Research Program, Melbourne, 2007.