TUTORIAL REVIEW

CSR
www.rsc.org/csr

Synchrotron and neutron techniques in biological crystallography

M. P. Blakeley,[a] M. Cianci,[b] J. R. Helliwell*[c] and P. J. Rizkallah[d]

[a] *6, Rue Jules Horowitz, BP181 38042 Grenoble Cedex 9, France. E-mail: blakeley@embl-grenoble.fr; Fax: 33 476.20.71.99; Tel: 33 476.20.77.96*

[b] *CCLRC Daresbury Laboratory, Keckwick Lane, Daresbury, UK WA4 4AD. E-mail: m.cianci@dl.ac.uk; Fax: 44 1925 60 3124; Tel: 44 1925 603658*

[c] *Chemistry Department, The University of Manchester, Oxford Road, Manchester, UK M13 9PL. E-mail: John.Helliwell@man.ac.uk; Fax: 44 161 275 4598; Tel: 44 161 275 4970*

[d] *CCLRC Daresbury Laboratory, Keckwick Lane, Daresbury, UK WA4 4AD. E-mail: p.j.rizkallah@dl.ac.uk; Fax: 44 1925 60 3124; Tel: 44 1925 603808*

Received 3rd February 2004
First published as an Advance Article on the web 16th September 2004

Synchrotron radiation (SR) techniques are continuously pushing the frontiers of wavelength range usage, smaller crystal sample size, larger protein molecular weight and complexity, as well as better diffraction resolution. The new research specialism of probing functional states directly in crystals, *via* time-resolved Laue and freeze trapping structural studies, has been developed, with a range of examples, based on research stretching over some 20 years. Overall, SR X-ray biological crystallography is complemented by neutron protein crystallographic studies aimed at cases where much more complete hydrogen details are needed involving synergistic developments between SR and neutron Laue methods. A big new potential exists in harnessing genome databases for targeting of new proteins for structural study. Structural examples in this *tutorial review* illustrate new chemistry learnt from biological macromolecules.

Matthew P. Blakeley is undertaking postdoctoral research at the European Molecular Biology Laboratory (EMBL) in Grenoble on the Institut Laue Langevin neutron Laue diffractometer, LADI. He took his PhD on an EPSRC/Institut Laue Langevin PhD studentship award supervised jointly by JRH and Dr Dean Myles (EMBL,Grenoble). His research interests are neutron protein crystallography methods and protein structure including protein saccharide recognition in concanavalin A and peanut lectin.

Michele Cianci is beamline scientist at CCLRC Daresbury Laboratory Synchrotron Radiation Department. He is currently commissioning the MAD10 beamline for structural genomics and proteomics briefly described in this paper. He read Chemistry at The University of Padua (Italy) and gained his PhD in Protein Crystallography at The University of Manchester with JRH. His research interests include macromolecular crystallography with synchrotron radiation, phasing methods with softer X-rays, crustacyanin from lobster, M. tuberculosis and Alzheimer's disease.

John R. Helliwell is Professor of Structural Chemistry, since 1989, at the University of Manchester and Joint Appointee with CCLRC Daresbury. He is also a Visiting Professor of Crystallography at Birkbeck College, London and is an Honorary member of the National Institute of Chemistry, Ljubljana, Slovenia. He trained in Physics (BA Hons. 1974 and then DSc 1996, both York University) and in Protein Crystallography (DPhil 1978, University of Oxford). His research interests are centred on biological crystallography including synchrotron radiation and neutron methods applied to elucidating the structure and function of proteins such as lectins, enzymes and crustacyanin.

Matthew Blakeley

Michele Cianci

John Helliwell

DOI: 10.1039/b312779f

Table 1 Synchrotron sources with facilities for SR biological crystallography around the world

Country	Facility	Stations[a]	General webpage
United Kingdom	SRS, Daresbury	5	http://www.dl.ac.uk/SRS/PX/index.html
	DIAMOND[b], Oxford	5 planned	http://www.diamond.ac.uk/Activity/I04A
France	ESRF, Grenoble	8	http://www.esrf.fr/UsersAndScience/Experiments/MX/
	SOLEIL[b], Paris	3 planned	http://www.synchrotron-soleil.fr/
Switzerland	SLS, Villigen	1	http://sls.web.psi.ch/view.php/beamlines/px/index.html
Italy	Elettra, Trieste	1 + 1 planned	http://www.elettra.trieste.it/experiments/beamlines/xrd1/index.html
Germany	Bessy II, Berlin	2	http://www.psf.bessy.de/
	DESY, Hamburg	5	http://www.embl-hamburg.de/facilities/
	ANKA, Karlsruhe	1	http://hikwww1.fzk.de/anka/
Sweden	Max II, Lund	6	http://maxsun5.maxlab.lu.se/beamlines/bli911/
Spain	LLS, Barcelona	1 planned	http://www.lls.ifae.es/
Brazil	LNLS, Campinas	1	http://www.lnls.br/principal.asp?idioma=2&conteudo=345&opcaoesq=105
USA	ALS, Berkeley	3	http://www-als.lbl.gov/als/
	APS, Chicago	16	http://www.aps.anl.gov/aps/frame_home.html
	MacChess, Cornell	3	http://www.macchess.cornell.edu/
	SSRL, Stanford	5	http://smb.slac.stanford.edu/
	NSLS, Brookhaven	5	http://www.nsls.bnl.gov/beamlines/Default.asp
	CAMD	1	http://www.camd.lsu.edu
Canada	CLS, Saskatoon[b]	1	http://www.lightsource.ca/beamlines/ProteinCrystallography/summary.php
Japan	SPRING-8, Hyogo	5	http://www.harima.riken.go.jp/eng/index.html
	Photon Factory, KEK	5	http://pfweis.kek.jp/eng/index.html
Australia	Boomerang[b]	1	http://www.synchrotron.vic.gov.au/
Taiwan	SRRC	1	http://biosrrc.nsrrc.org.tw/
Korea	PLS	1	http://pal.postech.ac.kr/kor/index.html
China	BSRF	1	http://www.ihep.ac.cn/bsrf/english/facility/beamline.htm
FSU	VEPP-3	1	http://ssrc.inp.nsk.su/

[a] Approximate numbers in some cases. [b] Under construction.

1 Introduction

This review outlines recent developments and current capabilities in biological crystal structure determination. The chemical basis of a variety of biological phenomena is elucidated in profound ways such as with macromolecular machines. As a broad selection we also highlight i) target selection in the post-genomics era; ii) the pharmaceuticals industry sector; iii) understanding the coloration chemistry of the lobster carapace; and iv) time-resolved Laue crystallography in enzyme catalysis structural studies. X-ray crystallography is less effective in finding the positions of H atoms, which cannot always be determined experimentally or fixed as riding-hydrogens in calculated positions based on the detailed structural chemistry. Although progress has been made in finding directly the H atoms with SR X-rays, neutron crystallography capabilities are evolving too in terms of technique (neutron Laue crystallography), new large area detectors (image plates) and new sources. The complementary nature of X-rays and neutrons as probes into the structure of matter is described.

Biological crystallography has been a major development in the molecular sciences in the last 70 years. In the 1980s and 1990s there was an exponential rise in biological X-ray crystallography with thousands of structures being solved (see section 4.3). A major driver of this was SR. Protein crystallography itself was identified by Max Perutz OM FRS[1] as the technique to be developed to work out the structure and function of haemoglobin, 'the molecular lung', which he embarked on in the 1930s and solved in the 1950s in Cambridge. A portent of the potential of SR in biological diffraction studies was the seminal work of Rosenbaum, Holmes and Witz (1971)[2] and Phillips *et al.* (1976)[3] as applied to muscle fibres and protein crystallography respectively. A portent of capability in neutron protein crystallography was the work going on at Brookhaven (USA) by *e.g.* Kossiakov and Schoenborn and at the Institut Laue Langevin in Grenoble by Sax Mason from the 1970s/80s (for an overview see 'Neutrons in biology', edited by B. P. Schoenborn and R. B. Knott, Plenum, 1994).[4]

2 Overview of developments in synchrotron radiation biological crystallography

Synchrotron radiation sources have been at the forefront of biological crystallography research for the last 25 years and continue to be so (see Table 1 for a list of current facilities). The demand for ever-higher resolution from smaller crystals of ever larger unit cell volumes, and the pressure to speed up structure determination have been the driving forces for a continuous development of beamlines, optics and detectors. The intrinsic properties of SR, namely high intensity and tunability, are routinely exploited to provide the user with higher resolution data and phasing information. The special case of harnessing the full emitted spectrum in Laue crystallography has opened up time-resolved structural studies on the seconds to nano-seconds timescale, and most recently to the 100 picoseconds time scale (see section 5.2).

Pierre J. Rizkallah is a Senior Beamline Scientist at SRS Daresbury Laboratory. He has worked there since 1987, first on a joint appointment with Liverpool University, and since 1990 on a full time Daresbury appointment to supply crystallography services to the pharmaceuticals industry, and to support academic users. A BSc in Chemistry from the Lebanese University, Beirut, 1979, was followed by a year in the cosmetics industry then a PhD in Chemistry from Nottingham University, UK. His research interests are mainly in carbohydrate recognition proteins, crustacyanin and immune system proteins. He is currently the Protein Crystallography Supply Manager for DARTS, the commercial office at Daresbury.

Pierre Rizkallah

2.1 Data collection considerations

These developments were possible thanks to the low emittance of electron storage rings allowing high brightness X-ray emission to yield finely focused, collimated, beams. The high radiation density increases the dose rate by many orders of magnitude compared with home rotating anode X-ray sources. Radiation damage becomes a significant issue,[5] affecting crystals in two ways: (i) heating due to absorption of radiation; (ii) generation of free radicals by ionisation due to the high energy incident radiation. Disulfide bridges in proteins are species that show specific structural damage.

A practical method to remove the heating effect is cryo-cooling of crystals.[6] While it is efficient at removing the heat, the reactive free radicals cannot be removed easily, but certainly their mobility is much reduced by cryo-temperatures. Basically most monochromatic and MAD (see section 2.2) data sets can be collected from one crystal at cryo-temperatures before significant damage sets in.

Flash cooling, however, can produce mechanical stresses in the crystal, due to the finite rate of heat loss in the crystal. Large crystals can break up under this stress (but see section 6.3), and the diffraction pattern would be spoilt. But the finely focused beam from a high brilliance SR source allows the use of small crystals (< 0.2 mm) that can be cryo-cooled usually routinely. Small crystal size is also emerging as a regular feature of high throughput 'structural genomics' protein crystallography pilot projects (see section 4) where 'microcrystals' are cited as a frequent problem *e.g.* where there is less time to optimise crystal size. The third generation sources such as ESRF in Grenoble, APS in Chicago and SPRING-8 in Japan have reduced the usable sample size range to 20 micron or less, when using the popular 1 Å wavelength. Further wavelength optimisation is possible, to allow even smaller crystals to be studied; the diffraction energy into a reflection, E_{hkl}, equation 1, can be described by the following expression, for the monochromatic rotating crystal case.[7]

$$E_{hkl} = \frac{e^4}{m^2c^4\omega} I_0\lambda^3 LPA \frac{V_x}{V_0^2} |\underline{F}(h)|^2 \qquad (1)$$

So, with the Lorentz factor, L, being approximately proportional to $1/\lambda$ at small θ, the overall diffracted intensity is proportional to $\sim\lambda^2$. Thus, softer X-rays can enhance diffraction that is rendered weaker by a small crystal volume V_x, where the situation is assisted by a small absorption effect, with A close to 1.0. There is a small penalty of utilising a longer wavelength in having a less favourable monochromator transmission polarization correction but overall an increase from 1 Å to 2.5 Å wavelength offers an order of magnitude increase in E_{hkl}, and a smaller crystal volume possibility by an equivalent factor.[8] The variation of the noise in the data as a function of wavelength must also be considered; systematic error due to absorption would be a problem if too big a crystal were used.

At the opposite end of the available SR spectrum, the use of ultra-short wavelengths, down to 0.3 Å, was suggested, and early tests conducted, as a method to remove the systematic error of absorption in the X-ray data, and to reduce radiation damage.[5] Such ultra-short wavelengths have been found to be beneficial for experiments on proteins and viruses under high pressure, where the cell window is necessarily thick to withstand the pressure.[9]

2.2 Phasing in macromolecular crystallography

Microscopy, analogous to imaging with visible light and lenses, cannot be employed with hard X-rays, because there exists no material with an appropriate optical index that can be used as an X-ray lens. Therefore, a successful, atomic detail, novel structure determination needs a computed estimate of the phases. Many combinations of experimental and mathematical techniques have been used over the years, all based on measured reflection intensities:

1) The chemistry based phasing approach of multiple isomorphous replacement (MIR) involving heavy atom substitution of proteins and multiple data set measurements from several crystals was the mainstay of *de novo* structure determination in the 1960–1990 period.

2) This situation changed with the advent of polychromatic SR where a multiple-wavelength anomalous dispersion (MAD) approach, based on measurements from a single crystal of a protein containing a suitable anomalous scatterer, became practical. Such a formalism, based on a theoretical minimum of two wavelengths (TW) (λ_1 and λ_2) and three experimental measurements ($F\lambda_1^+$, $F\lambda_1^-$ and $F\lambda_2^+$ or $F\lambda_2^-$, where '+' refers to an *hkl* reflection and (−) refers to its Friedel counterpart, or a symmetry equivalent, was described as early as 1956 by Okaya and Pepinsky.[10] Two-wavelength phasing at SR sources, 'TW',[8] and single-wavelength anomalous techniques[11] (and refs. therein) as an even more efficient usage of beam time, are now growing in popularity. The key is the physics of the scattering process where metal atoms, and even lighter atoms like sulfur, show a measurable variation ('dispersion') with wavelength; it is 'anomalous' only because the first theories in the 1900s ignored the effect. Thus the atomic scattering factor of such an 'anomalously scattering' atom, f, has wavelength dependent ($f' + if''$) corrections to it (Fig. 1); these are exploited in phasing each and every reflection. It is readily applicable due to two factors: some 30% of all proteins are metalloproteins, where the metal anomalous scattering X-ray dispersion can be harnessed, and it is possible to prepare selenomethionine substituted proteins, where the Se K edge at $\lambda = 0.978$ Å is well matched to the spectrum output from the most intense SR sources.[12] Together with the steady, concerted, development of instruments at SR facilities capable of MAD measurements, the prospect is now open of very large numbers of protein structures being determined, even on a genome scale, *i.e.* one beam line alone delivering many hundreds of protein structures per year.[13]

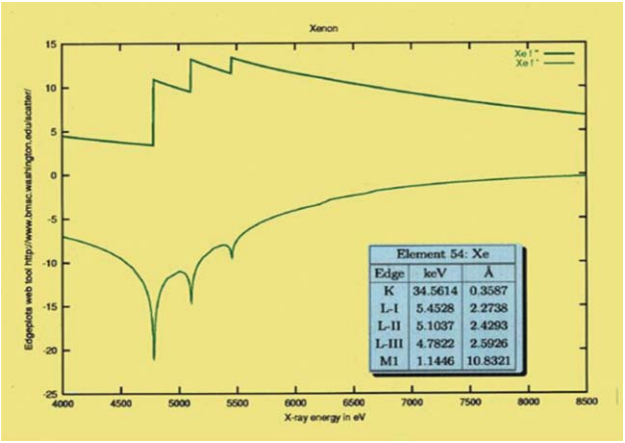

Fig. 1 The variation with wavelength at the L absorption edge of the anomalous scattering corrections (f' and f'') for xenon. Careful choice of wavelength at the SR beamline allows these effects to be varied and optimised so as to solve the crystallographic phase problem as part of the elucidation of the protein structure electron density image. Xenon can be introduced into the crystal under a high pressure and freeze trapped there. A xenon atom sits inert in a pocket on the protein surface or interior. Similar anomalous scattering effects apply to metal atom cases, including K and L absorption edges. Reproduced from ref. 8 with permission of the International Union of Crystallography (IUCr).

3) Xenon introduced under high pressure is yet another significant new method of heavy atom derivative preparation that has been developed in recent years. Xe is inert and therefore tends to occupy hydrophobic pockets without initiating a chemical reaction. Its relatively large number of electrons contributing to diffraction intensities means the Xe atoms are easy to locate in an electron density map, but low substitution occupancy can be a problem. Fig. 1 shows the variation with wavelength at the L absorption edge of the anomalous dispersion corrections for xenon.

4) Naturally occurring anomalous scatterers like sulfur or chlorine have attracted interest as a phasing tool, since it has become practical to measure highly redundant data sets,[14] and at longer wavelengths[8] (boosting their anomalous signal).

5) Direct methods of phasing diffraction amplitudes, as used routinely in chemical, small-molecule, crystallography, depend on matching theoretical and observed intensity distributions. It is successful generally at resolutions higher than 1.2 Å, but increasingly being recorded from protein crystals using SR.[15]

2.3 Detectors

For a review see chapter 5 in ref. 5. The classical detector employed by SR crystallographers in the 1980s was photographic film. The long tedious process of developing and fixing film before digitisation, spurred the development of electronic detectors. Early devices used the television screen principle, coupled with image intensifying technology, to produce real time observation of diffraction patterns. This device mainly suffered from poor spatial resolution and intrinsic noise.

Another early electronic detector was the multiwire chamber, developed for particle physics originally, which has two sets of finely separated wires with a high voltage field between them. The gap would be filled with special mixtures of gases at specific pressures. Diffracted X-rays passing through this atmosphere would ionise the gas and produce charge detectable at the wires that form the electrodes. Its shortcoming was mainly a low counting rate.

The next generation of electronic detectors was the image plate, which consisted of an X-ray absorbing material, held in a polymer matrix, which could not re-emit the absorbed energy until a laser light of the appropriate wavelength was shone on it. The response largely encompassed the dynamic range of the data without saturation. The duty cycle was relatively slow, though, requiring over 1 minute for most cases. Nevertheless they were again a major breakthrough. They are still the detector of choice in neutron protein crystallography (see section 6.1).

The current generation of electronic detectors rely on CCD imaging technology developed for the mass market. The technology was adapted by transforming the diffracted X-rays to visible light with an X-ray fluorescent layer, before delivering the pattern to the CCD. There is a mismatch between the large active areas required and the small chip sizes needed to ensure a uniform quality of CCDs. Optical demagnifying tapers are employed to bridge this gap. Larger active areas are achieved by building tiled versions of the system, such as in a $2 \times 2, 3 \times 3$ or even 4×4 matrix of CCD chips. A big advantage of CCDs is the fast duty cycle with excellent data quality.

Appearing over the horizon is the age of the 'pixel' detector. This is effectively the ultimate tiling of the CCD technology, where each pixel would be an individual detector. There remains the problem of achieving satisfactory electronic circuitry to address each pixel individually, without crosstalk, overheating or radiation damage.

2.4 A current example of a state-of-the-art beamline

Worldwide, structural genomics (see section 4) has been made possible by recent rapid progress in several related key technologies. These include genome sequencing projects, gene cloning and recombinant protein expression, bioinformatics for data mining to assign/predict function, and SR based MAD methods of phase determination. A North West Structural Genomics Consortium (NWSGC) resulted from the joint effort of research groups from several North West UK Universities, two major pharmaceutical companies (AstraZeneca and Astex-Technology) and CCLRC Daresbury Laboratory. The NWSGC was funded jointly by the UK Research Councils (BBSRC, EPSRC and MRC) to build a new MAD beamline at the SRS on a new high field Multipole Wiggler source. In order to meet the requirements of speed and tunability necessary for this project a new 2.4 tesla 10 pole wiggler beamline (MPW10) has been developed by Daresbury project team staff. The optical arrangement is characterized by a Rh coated collimating mirror, a double crystal Si(111) monochromator with horizontal sagittal focusing system, and finally a second Rh coated mirror for vertical focusing. The monochromatic beam is optimized through a 200 by 200 micron^2 collimator. The double crystal monochromator guarantees rapid tunability and the collimating mirror high energy resolution. Data can be collected from small, weakly diffracting, crystals over a wide range of wavelengths. It is dedicated to MAD techniques, operating in the 4 to 14 keV photon energy range, or 2.9 to 0.8 Å λ range. The 2θ max angular coverage will reach 110 degrees, important for measuring high diffraction resolution data even at longer wavelengths. A program of development is being undertaken with the aim of making the protein crystallography facilities as easy to use as possible. The station is already equipped with a MAR automated sample changer and semi-automatic crystal alignment. Up to 19 samples can be loaded at once on the robot, considerably reducing the time needed to change samples. In general artificial intelligence (AI) approaches, including robotics, are being developed to assess data quality with little human interaction. Such an intelligent system conducts continuous automatic data collection from cryo-cooled samples stored in the dispensing dewar in the experimental hutch. The store of samples is scanned for quality of diffraction, and priorities for full data collection set accordingly. It has become possible to collect tens of data sets or more in a 24 hour period using such systems. The computer network 'GRID' capabilities that are being developed will also enable the operator to be 'remote', for example in an office even thousands of miles away!

3 Recent applications of SR

3.1 Macromolecular machines

Major challenges were presented by crystal structures of large protein complexes such as the 'energy enzyme' F1-ATPase, a variety of viruses, or ribosomes. Ribosomes read the genetic code to perform protein biosynthesis and are assemblies of proteins and ribonucleic acids arranged in two subunits named 30S and 50S; in prokaryotes they have a molecular weight of 0.85 and 1.45 MDa respectively. The 30S unit is made up of one rRNA chain with $\sim$1500 nucleotides, and about 20 proteins, while the 50S unit is made up of two rRNA chains with $\sim$3000 nucleotides, plus 35 proteins. Ribosomes of pathogenic bacteria are important targets for different classes of antibiotics and recent high resolution crystal structures of complexes between ribosome subunits and antibiotics allowed the identification of binding sites and the design of many potential inhibitors.[16]

The work on the F1-ATPase crystal structure[17] was carried out by a team of scientists led by Professor John Walker using

the Daresbury SRS facility. The Nobel Prize for Chemistry for this effort, awarded in 1997 (shared), was the first for biological crystallography work based on SR data. The recent Chemistry award for Professor Rod MacKinnon's work is now the second recognition of SR based work by the Nobel Committee. MacKinnon[18] gave the first description of an ion channel protein and it was based on data acquired at the Cornell High Energy Synchrotron Source (CHESS), the National Synchrotron Light Source (NSLS) at Brookhaven, and at ESRF (Grenoble, France). The crystal structure was the first physical characterization of the membrane protein responsible for the selective movement of K^+ into and out of cells.

3.2 Structure-based drug discovery in industry

Biological crystallography has also become an industrial tool for structure-based drug discovery in the last ten years. Leading companies are now also incorporating robotics in large-scale initiatives to clone target genes, express them, purify the protein product and set up crystallisation trials. When applied to an identified target for the treatment of a disease, the protein crystal structure provides important starting information for designing a ligand that could affect a protein's function. Currently, a structure–activity relationship, SAR, is assessed mainly using NMR. A crystallography alternative SAR carries out the screening through soaking protein crystals in pharmacophore fragment cocktails. The actual bound fragment is developed further into lead compounds (see article in this issue by Tickle *et al.*[19]).

4 Structural genomics and bioinformatics

Synchrotron radiation biological crystallography is experiencing significant growth arising from the genome sequence projects. These applications are highlighted separately now here.

4.1 Examples of genome sequencing projects

There is now a plethora of projects reporting whole genome sequences. In the bacterial world, a great impact on world health care can be anticipated from the work on *Mycobacterium tuberculosis*, Tb, (4000 genes).[20] The first complete genome of a eukaryotic parasite is that of *Plasmodium falciparum* (5200 genes),[21] clearly a milestone in malaria research. Early comparisons with the human genome have already shown pathways that are specific to the pathogen and which could thereby be targets for drug design. Besides the human genome, where the final draft is being readied, the *Mouse Genome* (30,000 genes)[22] is a landmark in mammalian research. Work with mouse genes allows their manipulation and the determination of their function, and gives insights into equivalent basic human biology.

4.2 Protein target selection in the post genomic era

Target selection has until recent years been a matter of choosing proteins where extensive biochemistry had been undertaken prior to the crystal structure analysis. In the very early years in effect the availability of crystals was usually the sole issue. Today, *via* the power at one's fingertips of the genome sequence databases, target selection has expanded possibilities. For example, our most recent selection of targets to work on includes a set of 95 proteins derived from a Boolean algebra analysis of common proteins between Tb, *M. leprae* and chlamydia but non-intersecting with human proxy genomes of worm and fly[23] (Fig. 2). This group of 95 proteins are present in the organism but absent in the human genome, thus targeting them would hopefully avoid side effects.

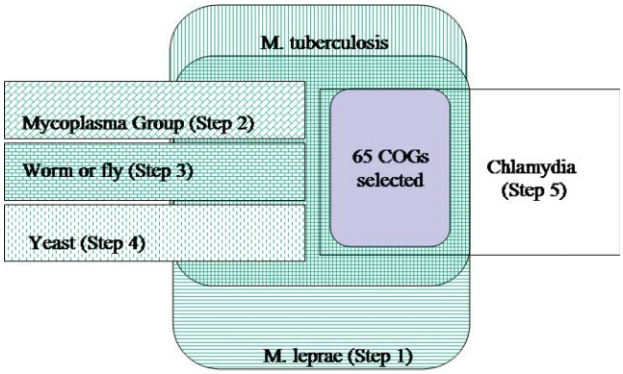

Fig. 2 Specific targeting of new proteins can now harness genome sequence data. Raftery and Helliwell, (2002)[23] performed a Boolean algebra analysis to identify "cluster of orthologous group (COG)" genes in the genomes of these pathogens that are not in the Human (proxy) genome.

4.3 Projects to fill the protein fold gaps

The Protein Data Bank, PDB, was set up ~25 years ago, and currently holds ~24,000 depositions, with a growth rate that has doubled approximately every few years. The content is made up of structures mainly determined and refined by crystallographic techniques (~90%). There are ~4000 unique protein folds in the PDB. A long standing question in the field of protein structure is: "What is the true number of unique folds?" A variety of algorithms were developed[24] with a consensus that the answer is finite, numbering 'thousands'. The answer to this question is not academic. The knowledge of a particular unique fold could perhaps serve, through homology modelling, as the template for the structure determination of many other proteins using 'molecular replacement'. This latter method is an alternative to phasing (section 2.2) whereby the position and orientation of the search motif are found in the new protein crystal. A major thrust of the new field of structural genomics is indeed to expand the representation of unique folds and thus fill the gaps.

4.4 Medical-oriented structural genomics projects

Following the sequencing of its genome, the TB Structural Genomics Consortium (http://www.doe-mbi.ucla.edu/TB/)[25] was formed to provide a structural basis for the development of therapeutics by solving the 3-D structures of proteins from Tb, and making them publicly available. The Consortium has now enlisted more than 70 laboratories worldwide.

The SPINE Program (Structural Proteomics in Europe; http://www.spineurope.org/) has assembled 11 labs within Europe, and several satellites, all dedicated to structural biology. The scientific, and coordination objectives of the SPINE project are to develop technologies permitting high throughput structure determination by X-ray crystallography and NMR not only of prokaryotic but also of eukaryotic proteins and complexes. These technologies will then be used for the determination of the atomic structures (50 in year one and 500 by the end of year three) of proteins of medical interest including viral structural genomics and MTb proteins.

5 Structural projects from our labs as detailed examples

5.1 The case of crustacyanin: learning new colouration chemistry from biology

The molecular basis of the camouflage colouration of marine crustacea is often provided by carotenoproteins. The blue colour of the lobster *Homarus gammarus* carapace, for example, is intricately associated with a 16 protein subunit

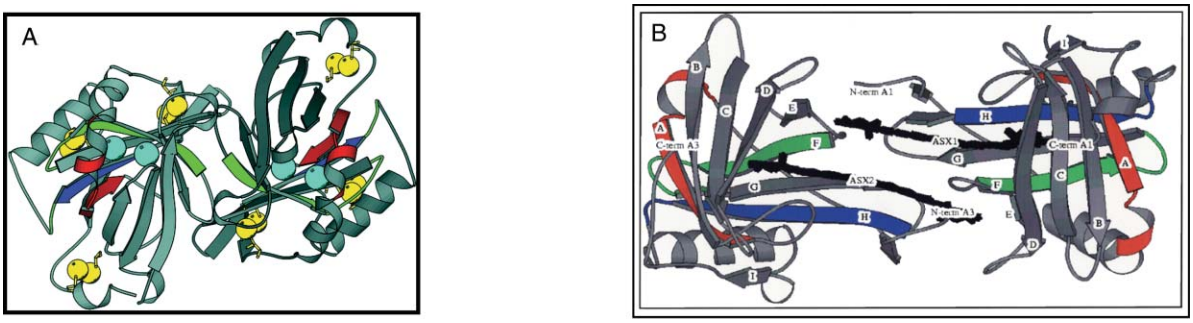

Fig. 3 (A, left) The apocrustacyanin A1 homodimer has a quite different dimerisation assembly to the β-crustacyanin (A1, A3) in (B, right), thus showing the influence of the two astaxanthins on the dimerisation assembly process. Xenon atoms are in turquoise and sulfurs in yellow. (B, right) β-Crustacyanin (A1 protein with A3 protein) dimer complexation of two astaxanthins (bold-black) AXT1, AXT2. Red, green and blue colouring highlight β-sheet strands that are the lipocalin family conserved regions of the protein amino acid sequence. The amino acid sequence identity between the A1 and A3 proteins is ~40%, but with a higher conservation in the group of carotenoid complexing amino acid side chains from A1 and correspondingly from A3. Reproduced from ref. 28 with the permission of PNAS.

complex, with 16 bound astaxanthin molecules (for a recent review see[26]).

Crystals of α-crustacyanin were first reported some 30 years ago though without any recorded diffraction. Later, single subunits of this complex were targeted, hoping that they might be more suitable for structure determination. But this proved problematic through lack of a suitable molecular replacement model (even from the same family of lipocalin proteins), lack of good heavy atom derivatives and lack of selenomethionine variants. Crystal structures were needed to provide the structural basis of the spectral shifts of the carotenoid in the complex state. The first, breakthrough, structure of apocrustacyanin A1 (Fig. 3A) was solved using 2 Å X-rays, comparing native data with a Xe derivative dataset. The relatively longer wavelength optimised the *f″* signal of Xe. The correct hand of the Xe-derived phases was determined using the S anomalous signal from a high multiplicity native data set, also enhanced by recording at 2 Å wavelength.[27] Thus the specific attribute of SR of its tunability, along with a high intensity, was harnessed.

The crystal structure solution of the apo A1 subunit (Fig. 3A) opened the way for the structural determination of β-crustacyanin in holo form (Fig. 3B).[28] These crystals were quite small (100 × 100 × 450 microns), the unit cell volume was quite large (in excess of one million cubic Å) and the solvent volume of the crystal also large (> 80%); these three factors together meant that the high brightness of SR was therefore needed in this project. The structure revealed firstly that the dimerisation of the proteins was totally different in the apo and holo forms; thus the two bound carotenoids totally determined the dimer arrangement. Secondly the binding sites of the carotenoids are both similar; the bound astaxanthins each have at one end a bound water molecule within 2.8 Å of the end ring keto oxygen and at the other end a histidine each with a nitrogen atom within 3.0 Å of the other end ring keto oxygen. Discussions on the bathochromic colour shift are based on these astaxanthin keto oxygens and possible charge polarization effects as well as coplanarisation of the end rings with respect to the plane of the astaxanthin polyene chain.

Comparison of apo A1 and holo A1 structures allowed a simulation of the conformational changes on carotenoid binding to be made (see http://journals.iucr.org/d/issues/2003/12/00/gr5000/gr5000sup1.gif).

5.2 Time-resolved crystallography: the biological chemistry of hydroxymethylbilane synthase, HMBS

HMBS catalyses the polymerisation of four molecules of porphobilinogen to form hydroxymethylbilane. The evolution of the reaction in crystals of a Lys 59 Gln mutant was studied by repeated data collection *via* Laue diffraction snapshots. A progression of Laue exposures were recorded at ESRF on the Laue beamline ID09 with pre-set, lengthening, time gaps as a substrate solution was passed over a crystal held in a flow cell. The experimental difference maps, produced with the Laue data, revealed a lengthening difference electron density feature, commencing after ≈8 minutes, grown more by 25 minutes and strongest and at its longest after 2 h, by which time the initially colourless enzyme crystal had become red/pink.[29] This elongated electron density commenced at the position of C2 of the oxidised co-factor of the enzyme, the putative binding site of the substrate, and directly above the critical carboxyl side chain of Asp 84 involved in the first ring coupling reaction step (Fig. 4). The

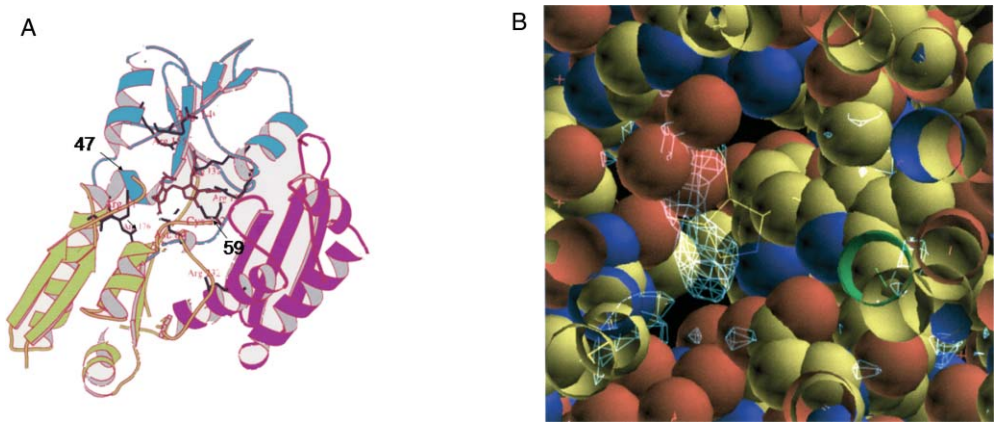

Fig. 4 (A) Hydroxymethylbilane synthase (HMBS) and active form cofactor; arrows mark the start and end of the mobile polypeptide loop 47 to 59; (B) the '2 hour' difference electron density and the HMBS oxidised cofactor 2nd pyrrole ring sit in the enzyme active site cavity. Reproduced from ref. 29 with permission of the RSC.

electron density then extended past amino acid residues that are known from protein engineering to affect later stages of the catalysis, and into open solvent. There is a disordered stretch of amino acid residues (49–57) in all current HMBS structures undertaken at ambient or cryo-temperature (time-resolved Laue or static monochromatic). The elongated density referred to above resides between where that loop is likely to be and the position of the C2 ring, in the reduced active cofactor conformation. Thus the time-evolution of this reaction in HMBS (Lys 59 Gln mutant) has been established over a period of seconds to hours, in structural and functional terms. The study of the evolution with time of structural intermediates of reactive protein molecules in enzymes like HMBS, is a modern growth trend in crystal structural analyses.[30] These developments have been made feasible due to the hugely expanded capabilities for fast, repeated, data collection with CCDs (see section 2.3). Thus a structure to function relationship can be explored directly by experiment. A particular impetus in time-resolved Laue protein structural studies is that even nanosecond studies[31,32] and most recently to the 150 picoseconds timescale[33] are possible. The range of time-resolutions reaches well beyond freeze trapping of structural intermediates. For a review see Ren *et al.*, 1999.[34]

6 Neutron biological crystallography

Neutrons have a unique role to play in determining the structure and dynamics of biological macromolecules and their complexes. The similar neutron scattering magnitude from deuterium, carbon, nitrogen and oxygen nuclei, means that the effect of atomic vibration in lowering the visibility of these atoms in Fourier maps is no worse for the deuterium atoms, unlike the X-ray case. Moreover, the negative scattering length of hydrogen and the positive scattering length of deuterium allow the well-known neutron H/D contrast variation method to be applied. Furthermore, as there is no problem with radiation damage using neutrons as the diffraction probe, unlike X-rays, room temperature neutron data collection studies are completely viable. Unfortunately, even with these advantages, the low flux of existing neutron facilities means that neutron biological crystallography is not going to be a high throughput technique due to the long measuring runs at present necessary *e.g.* typically between one to four weeks or more. However, the proteomics programmes of research will make accessible many more candidate proteins. Overall, there is a renewed and growing interest in neutron studies today. The importance rests on needing to know the details of the hydrogens and bound water structure, which are involved in virtually all the molecular processes of life. This experimental structural information is mostly incomplete when studied by X-rays alone. Also, many enzyme reactions involve hydrogen. Thus, there is great potential for wide application of neutrons, should the technical capability be found.

There are two major hurdles for wide application; firstly the size of crystals routinely available *versus* the sizes required, and secondly a molecular weight ceiling of typically 40 kDa. These limits are set by the weak scattered signal into the diffraction reflections, dictated by these parameters in equation (1). The background noise in the diffraction pattern is generated in large part by the incoherent scattering effect of the hydrogens. By soaking the crystals in D_2O the noise can be reduced but not eliminated because the hydrogens on the carbon atoms do not exchange by simple soaking. These hydrogens can be deuteriums *via* microbiological expression of proteins from bacteria grown on deuterated media. This will improve the signal to noise of diffraction reflection measurements in neutron protein crystallography by at least an order of magnitude,[35] as well as open up new contrast variation experiments. At the Institut Laue Langevin (the ILL) in Grenoble a new European funded perdeuteration laboratory has recently come on line with a focus on more routine production of fully deuterated proteins.

6.1 Neutron data collection perspectives

Synergies between neutron and SR Laue crystallography, namely a commonality of knowledge of Laue geometry irrespective of radiation type, opened up a new path in neutron crystallography data collection at the ILL, the Laue Diffractometer 'LADI'.[36] The gain in speed over monochromatic neutron techniques has been notable with the advantage also of allowing smaller crystals and bigger unit cells to be investigated. However, this is at the expense of signal to noise (S/N) in the diffraction pattern. Nevertheless, very high resolution studies ($\sim$1.5 Å) on small proteins (*e.g.* rubredoxin, triclinic and tetragonal lysozyme) have been undertaken in times of 10 to 14 days.[37,38] Narrow bandpass Laue and/or fully deuterated protein significantly improves S/N. On LADI, the geometric limit of resolution is $\sim$1.4 Å and the molecular weight ceiling is around 40 kDa in practice, finally currently limited due to spot overlap congestion for the fixed radius and the crystal cross sections commonly in use (up to 3 mm) *e.g.* see Fig. 5.

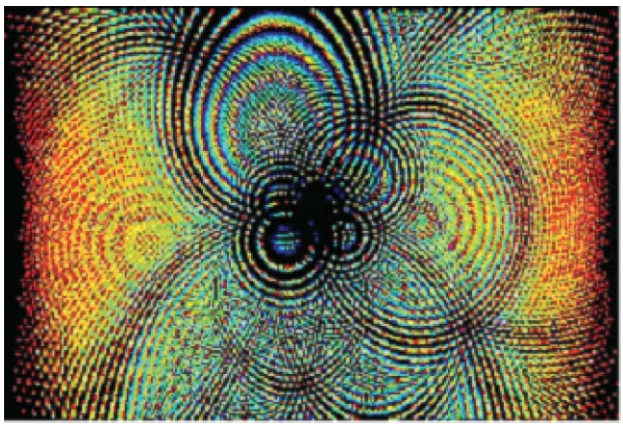

Fig. 5 Neutron Laue crystal diffraction pattern simulation from a concanavalin A/glucoside complex. Crystal space group $I213$, a = 167.8 Å. The pattern is colour coded for wavelength (red = long wavelength, 3.8 Å through the rainbow colours to blue = short wavelength, 2.8 Å; these are the typical experimental conditions in use at the Institut Laue Langevin LAue DIffractometer for biological crystallography). At far left and far right then are the 'full back scattering position' reflection data.

However, monochromatic techniques can also remain attractive at neutron reactor sources,[39] but this results in longer data collection times (typically 30 to 60 days per monochromatic data set), even with large area coverage neutron image plates, and is restricted to smaller protein unit cells.

Neutron source and further apparatus developments will make an important impact. There are enhancement plans for LADI in Grenoble; there is a proposed Large Molecule Xstallography (LMX) instrument on ISIS2; recently on-line is the Protein Crystallography System (PCS) at Los Alamos; there is the MANDI, 'MAcromolecular Neutron Diffraction Instrument', planned for the new 1.4 MW USA source 'SNS' and similarly an instrument planned at the 1 MW Japanese source, both sources under construction. A summary of current and planned instruments is given in Table 2.

6.2 Realising a complete structure: the complementary roles of X-ray and neutron protein crystallography

Whilst the determination of hydrogens in proteins is now feasible with ultra-high/atomic resolution protein SR X-ray

Table 2 Neutron sources with facilities for neutron biological crystallography around the world

Country	Facility	Stations	General webpage
United Kingdom	ISIS TS2, Oxford	LMX[b]	http://www.isis.rl.ac.uk/targetStation2/instruments/LMXweb.pdf
France	ILL, Grenoble	LADI[a]	http://www.ill.fr/YellowBook/LADI/
Japan	JAERI, Tokai	BIX3[a], BIX4[a]	[ref. 39]
	JPS[c], Tokai	nPX[b]	
USA	SNS, Tennessee[c]	MANDI[b]	http://www.sns.gov
	LANSCE	PCS[a]	http://lansce.lanl.gov/lujan/instruments/pcs/PCS.html

[a] Currently operating. [b] Planned. [c] Under construction.

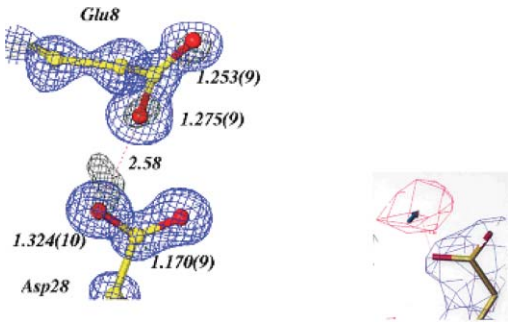

Fig. 6 Finding protein carboxy side chain hydrogens *via* (left) X-ray derived electron density maps and precise bond distances (standard uncertainty values in brackets) and (right) *via* neutron derived nuclear density maps. Reproduced from ref. 41 with permission of the RSC.

crystallography (Fig. 6), mobility of hydrogens can abolish their diffraction signal. Since neutron protein crystallography determination of deuteriums at around 2 Å resolution matches that at 1.0 Å by SR X-rays, the more mobile hydrogens are determinable by the neutron approach. Indeed the bound solvent is a whole category of deuterium atoms which are more efficiently sought by neutron techniques;[40] Fig. 7. As

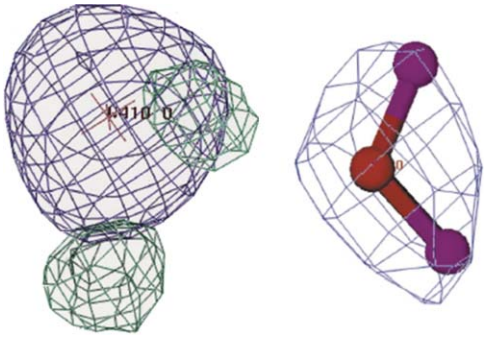

Fig. 7 (left) Finding bound water hydrogens *via* the ultra high resolution X-ray structure of saccharide-free concanavalin A at 110 K.[41] (right) A water molecule in the 2.4 Å room temperature neutron study[40] showing the orientation of the D_2O nuclear density. The 2Fo–Fc neutron map is shown in blue (at 1.5 r.m.s.). In the neutron study the oxygen atom positions of bound waters such as this one were held fixed at the X-ray position from an X-ray study of an identically prepared D_2O protein crystal. Reproduced from ref. 40 with permission of the International Union of Crystallography (IUCr).

emphasised in ref. 41 with concanavalin A studied at 0.94 Å, the X-ray data to parameter ratio was ~5. Protein model structure dictionary restraints add data, but some of these restraints are not appropriate *e.g.* for carboxyl side chains where the non-protonated dictionary assumption is not necessarily correct; having sufficient X-ray data can then override these restraints. It is in fact not possible to remove the dictionary restraints because individual atoms or groups of atoms (such as on loops on the protein) are sufficiently mobile that they diffract relatively more weakly towards the edges of the pattern. Thus the dictionary structure restraints have to be used to retain a proper overall geometry. It is in these more mobile parts of the protein structure that the role of neutron protein crystallography in completing the protein structure protonation details is most obviously needed.

6.3 Cryo-neutron protein crystallography

Since neutrons do not cause radiation damage there has not been a pressing need to cryo-protect crystals. However, there is the advantage to be gained from cryo-crystallography of the reduction of atomic mobility and enhanced nuclear-density definition. Furthermore, collecting data at cryo-temperatures can also reduce background (diffuse scattering), and therefore can aid the signal-to-noise ratio and hence the resolution limit improved. Cryo-crystallography also opens up the possibility of freeze trapping of intermediates in biological reactions triggered in a protein crystal.[42] So, the idea of combining neutron and cryo-crystallography advantages together is of considerable interest.

It seems to have been generally assumed that freezing large crystals is not possible. However, we have recently shown that it is possible to freeze and collect high resolution X-ray (1.65 Å on a rotating anode) and neutron data (2.5 Å on LADI, Fig. 8)

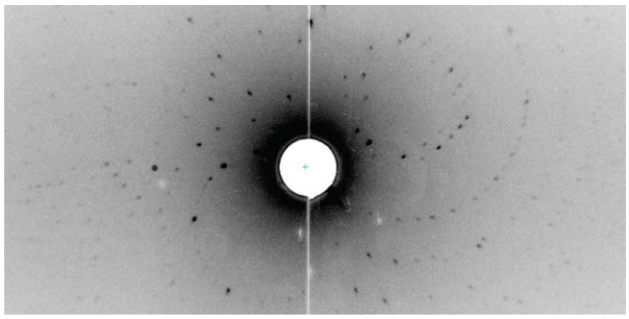

Fig. 8 A neutron Laue diffraction pattern from a large concanavalin A crystal at 15 K[43] recorded on LADI still shows good diffraction spot shape.

from large concanavalin A protein crystals (~ 2 and 5 mm^3) as examples.[43] These data have allowed a combined 'X + n' protein structure analysis to be undertaken[43] as performed previously with room temperature X-ray and neutron data sets.[40] These results demonstrate the potential of protein cryo-crystallography with neutrons, thus combining the advantages of the neutron and cryo-approaches for studying the structural details of bound water hydrogens (as deuteriums) and of protonation states of histidines and carboxyl side chains. Most exciting of all, this opens up the possibility of time-resolved neutron freeze-trap biological crystallography.

7 Conclusions

The picture that is emerging is quite complex and challenging. Biological crystallography has reached a maturity where it is now branching in many different and interrelated directions. It

has certainly become a tool to routinely study *e.g.* protein–protein interactions and protein–ligand/inhibitor/drug interactions. Genomics projects will unravel more and more proteins critical for structure and function studies. Synchrotrons will always be at the forefront of such studies providing investigators with state-of-the-art facilities and brighter and brighter X-ray beams. It is a major reason why the UK is investing so strongly in the new 3rd generation source 'diamond' (www.diamond.ac.uk). Optimal use of neutrons is still in a very active phase of development, and with expanding instrument provision worldwide. A major aspect of the placement of diamond together with ISIS on the same UK site is the strong synergy planned involving SR and neutron techniques.

Overall, the scope of the technique of biological crystallography has been transformed. As a result more complicated biologies are being understood at the chemical level. Also new chemistries are being learnt from studies of 'how nature does it'.

Acknowledgements

JRH is very grateful for research grant support to BBSRC, The Wellcome Trust and The Leverhulme Trust and for PhD studentship support from EPSRC, BBSRC, the British Council, the UK/Israel Fund and the EU (Host Institute, Network and Marie Curie Training Centre awards), as well as SR beam time awards at SRS Daresbury, ESRF and Cornell and neutron beamtime at the ILL, Grenoble. The MPW10 beamline on SRS is funded by BBSRC (with EPSRC and MRC) in a grant to the North West Structural Genomics Consortium, PIs Professor S. S. Hasnain (Daresbury) and JRH. JRH most warmly thanks all his collaborators in the publications referenced. PJR is grateful for the technical support by the staff at SRS Daresbury Laboratory, and for the supply of beamtime by the Joint Biology Programme of the UK Research Councils. MC wishes to acknowledge all the CCLRC staff at Daresbury Laboratory involved in the development of the MPW10 beamline. He is also grateful to The University of Manchester and its research staff in the Chemistry Department for the support throughout his PhD studies funded by EPSRC and the University of Manchester. MPB and JRH thank Dr Dean Myles for discussions and collaborations in neutron protein crystallography. MPB was an Institut Laue Langevin (ILL) and EPSRC PhD student with JRH and DM; ILL and EPSRC are gratefully thanked for this studentship.

References

1 M. F. Perutz, 1992, *Protein Structure; New approaches to disease and therapy*, W. H. Freeman & Co., New York.
2 G. Rosenbaum, K. C. Holmes and J. Witz, *Nature*, 1971, **317**, 145.
3 J. C. Phillips, A. Wlodawer, M. M. Yevitz and K. O. Hodgson, *Proc. Natl. Acad. Sci. U. S. A.*, 1976, **73**, 128.
4 *Neutrons in biology*, ed. B. P. Schoenborn and R. B. Knott, Plenum, New York, 1994.
5 J. R. Helliwell, 1992, *Macromolecular Crystallography with Synchrotron Radiation*, Cambridge University Press, Paperback Edition 2004.
6 E. J. Garman and T. R. Schneider, *J. Appl. Crystallogr.*, 1997, **30**, 211–237.
7 M. M. Woolfson, 1970, *Xray Crystallography*, Cambridge University Press, 2nd Edition 2002.
8 J. R. Helliwell, *J. Synchrotron Radiat.*, 2004, **11**, 1.
9 R. Fourme, E. Girard, R. Kahn, I. Ascone, M. Mezouar, A. C. Dhaussy, T. Lin and J. E. Johnson, *Acta Crystallogr., Sect. D*, 2003, **59**, 1767.
10 Y. Okaya and R. Pepinsky, *Phys. Rev.*, 1956, **103**, 1645.
11 A. Olczak, M. Cianci, Q. Hao, P. J. Rizkallah, J. Raftery and J. R. Helliwell, *Acta Crystallogr., Sect. A*, 2003, **59**, 327.
12 W. A. Hendrickson, J. R. Horton and D. M. LeMaster, *EMBO J.*, 1990, **9**, 1665.
13 A. Cassetta, A. M. Deacon, S. E. Ealick, J. R. Helliwell and A. W. Thompson, *J. Synchrotron Radiat.*, 1999, **6**, 822.
14 Z. Dauter, M. Dauter and E. Dodson, *Acta Crystallogr., Sect. D*, 2002, **58**, 494.
15 J. X. Yao, *Acta Crystallogr., Sect. D*, 2002, **58**, 1941–1947 and refs. therein.
16 A. P. Carter, W. M. Clemons, D. E. Brodersen, R. J. Morgan-Warren, B. T. Wimberly and V. Ramakrishnan, *Nature*, 2000, **407**, 340.
17 J. P. Abrahams, A. G. Leslie, R. Lutter and J. E. Walker, *Nature*, 1994, **370**, 621.
18 D. A. Doyle, J. Morais Cabral, R. A. Pfuetzner, A. Kuo, J. M. Gulbis, S. L. Cohen, B. T. Chait and R. MacKinnon, *Science*, 1998, **280**, 69.
19 I. Tickle, A. Sharff, M. Vinković, J. Yon and H. Jhoti, *Chem. Soc. Rev.*, 2004, this issue (DOI: 10.1039/b314510g).
20 S. T. Cole, R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier, S. Gas, C. E. Barry III, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels and B. G. Barrell, *Nature*, 1998, **393**, 537.
21 M. J. Gardner, N. Hall, E. Fung, O. White, M. Berriman, R. W. Hyman, J. M. Carlton, A. Pain, K. E. Nelson, S. Bowman, I. T. Paulsen, K. James, J. A. Eisen, K. Rutherford, S. L. Salzberg, A. Craig, S. Kyes, M. S. Chan, V. Nene, S. J. Shallom, B. Suh, J. Peterson, S. Angiuoli, M. Pertea, J. Allen, J. Selengut, D. Haft, M. W. Mather, A. B. Vaidya, D. M. Martin, A. H. Fairlamb, M. J. Fraunholz, D. S. Roos, S. A. Ralph, G. I. McFadden, L. M. Cummings, G. M. Subramanian, C. Mungall, J. C. Venter, D. J. Carucci, S. L. Hoffman, C. Newbold, R. W. Davis, C. M. Fraser and B. Barrell, *Nature*, 2002, **419**, 498.
22 Mouse Genome Sequencing Consortium, *Nature*, 2002, **420**, 520.
23 J. Raftery and J. R. Helliwell, *Acta Crystallogr., Sect. D*, 2002, **58**, 875.
24 D. Vitkup, E. Melamud, J. Moult and C. Sander, *Nat. Struct. Biol.*, 2001, **8**, 559.
25 T. C. Terwilliger, M. S. Park, G. S. Waldo, J. Berendzen, L. W. Hung, C. Y. Kim, C. V. Smith, J. C. Sacchettini, M. Bellinzoni, R. Bossi, E. De Rossi, A. Mattevi, A. Milano, G. Riccardi, M. Rizzi, M. M. Roberts, A. R. Coker, G. Fossati, P. Mascagni, A. R. Coates, S. P. Wood, C. W. Goulding, M. I. Apostol, D. H. Anderson, H. S. Gill, D. S. Eisenberg, B. Taneja, S. Mande, E. Pohl, V. Lamzin, P. Tucker, M. Wilmanns, C. Colovos, W. Meyer-Klaucke, A. W. Munro, K. J. McLean, K. R. Marshall, D. Leys, J. K. Yang, H. J. Yoon, B. I. Lee, M. G. Lee, J. E. Kwak, B. W. Han, J. Y. Lee, S. H. Baek, S. W. Suh, M. M. Komen, V. L. Arcus, E. N. Baker, J. S. Lott, W. Jacobs, Jr, T. Alber and B. Rupp, *Tuberculosis (Edinb.)*, 2003, **83**(4), 223.
26 N. E. Chayen, M. Cianci, J. G. Grossmann, J. Habash, J. R. Helliwell, G. A. Nneji, J. Raftery, P. J. Rizkallah and P. F. Zagalsky, *Acta Crystallogr., Sect. D*, 2003, **59**, 2072.
27 M. Cianci, P. J. Rizkallah, A. Olczak, J. Raftery, N. E. Chayen, P. F. Zagalsky and J. R. Helliwell, *Acta Crystallogr., Sect. D*, 2001, **57**, 1219.
28 M. Cianci, P. J. Rizkallah, J. Raftery, N. E. Chayen, P. F. Zagalsky. and J. R. Helliwell, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 9795–9800.
29 J. R. Helliwell, Y. P. Nieh, J. Habash, P. F. Faulder, J. Raftery, M. Cianci, M. Wulff and A. Hädener, *Faraday Discuss.*, 2002, **122**, 131.
30 J. R. Helliwell and C. C. Wilson (Eds.), *Faraday Discuss.*, 2002, **122** (Time-resolved Chemistry).
31 Z. Ren, B. Perman, V. Srajer, T. Y. Teng, C. Pradervand, D. Bourgeois, F. Schotte, T. Ursby, R. Kort, M. Wulff and K. Moffat, *Biochemistry*, 2001, **40**(46), 13788–13801.
32 V. Srajer, Z. Ren, T. Y. Teng, M. Schmidt, T. Ursby, D. Bourgeois, C. Pradervand, W. Schildkamp, M. Wulff and K. Moffat, *Biochemistry*, 2001, **40**(46), 13802–13815.
33 F. Schotte, M. H. Lim, T. A. Jackson, A. V. Smirnov, J. Soman, J. S. Olson, G. N. Phillips, M. Wulff and P. A. Anfinrud, *Science*, 2003, **300**(5627), 1944–1947.
34 Z. Ren, D. Bourgeois, J. R. Helliwell, K. Moffat, V. Srajer and B. L. Stoddard, *J. Synchrotron Radiat.*, 1999, **6**, 891.
35 F. Shu, V. Ramakrishnan and B. P. Schoenborn, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 3872.
36 J. R. Helliwell, C. Wilkinson, 1994, *Hercules Course*, **Volume III**, chapter XII, Edition de Physique, Springer Verlag, Berlin, Heidelberg.
37 R. Bau, M. P. Blakeley, D. A. A. Myles, 2004, to be published.

38 C. Bon, M. S. Lehmann and C. Wilkinson, *Acta Crystallogr., Sect. D*, 1999, **55**, 978.
39 I. Tanaka, K. Kurihara, T. Chatake and N. Niimura, *J. Appl. Crystallogr.*, 2002, **35**, 34–40 and refs. therein.
40 J. Habash, J. Raftery, R. Nuttall, H. J. Price, C. Wilkinson, A. J. Kalb (Gilboa) and J. R. Helliwell, *Acta Crystallogr., Sect. D*, 2000, **56**, 541.
41 A. Deacon, T. Gleichmann, A. J. Kalb (Gilboa), H. Price, J. Raftery, G. Bradbrook, J. Yariv and J. R. Helliwell, *J. Chem. Soc., Faraday Trans.*, 1997, **93**, 4305.
42 K. Moffat and R. Henderson, *Curr. Opin. Struct. Biol.*, 1995, **5**, 656–663.
43 M. P. Blakeley, A. J. Kalb (Gilboa), J. R. Helliwell and D. A. A. Myles, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, submitted.